



JOINT WMO/IOC TECHNICAL COMMISSION FOR
OCEANOGRAPHY AND MARINE METEOROLOGY

ADVANCES IN THE APPLICATIONS OF MARINE CLIMATOLOGY

**The Dynamic Part of the WMO Guide to the Applications
of Marine Meteorology**

WMO/TD-No. 1081

2003

JCOMM Technical Report No. 13

WORLD METEOROLOGICAL ORGANIZATION

INTERGOVERNMENTAL OCEANOGRAPHIC
COMMISSION (OF UNESCO)

ADVANCES IN THE APPLICATIONS OF MARINE CLIMATOLOGY

**The Dynamic Part of the WMO Guide to the Applications
of Marine Meteorology**

WMO/TD-No. 1081

2003

JCOMM Technical Report No. 13

NOTE

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariats of the Intergovernmental Oceanographic Commission (of UNESCO), and the World Meteorological Organization concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

TABLE OF CONTENTS

	FOREWORD	v
	ACKNOWLEDGEMENTS	vi
	INTRODUCTION	vii
SECTION 1	MARINE DATABASE ENHANCEMENTS	1
	COADS updates and the blend with the UK Met Office Marine Data Bank ..	3
	The Kobe Collection: newly digitized Japanese historical surface marine meteorological observations	11
	An archive of underway surface meteorology data from WOCE	20
SECTION 2:	EVALUATION OF MARINE DATA SOURCES	26
	The accuracy of marine surface winds from ships and buoys	27
	Report on Beaufort equivalent scales	41
	Evaluation of ocean winds and waves from voluntary observing ship data	53
	Evaluation of NCEP reanalysis surface marine wind fields for ocean wave hindcasts	68
SECTION 3:	METADATA AND DATA QUALITY	87
	Improving global flux climatology: the role of metadata	89
	Establishing more truth in true winds	98
	Quality control in recent and pending COADS releases	116
SECTION 4:	DEVELOPMENT AND USE OF SATELLITE MARINE DATABASES	125
	An intercomparison of voluntary observing, satellite data, and modelling wave climatologies	127
	The joint calibration of altimeter and in situ wave heights	139
	On the use of in situ and satellite wave measurements for evaluation of wave hindcasts	149
	Scatterometry data sets: high quality winds over water	159
SECTION 5:	ANALYSIS OF CLIMATE VARIABILITY AND CHANGE	175
	Outlier detection in gridded ship data sets	177
	A methodology for integrating wave data from different sources permitting a multiscale description of wave climate variability	187
	Reduced space approach to the optimal analysis of historical marine observations: accomplishments, difficulties, and prospects	199
	Analysis of wave climate trends and variability	217
SECTION 6:	USER REQUIREMENTS FOR CLIMATE INFORMATION	227
	Offshore industry requirements and recent metocean technology developments	229
	Specific contributions to the observing system: sea surface temperatures	234
	Importance of marine data to seasonal forecasting in Australia	242

FOREWORD

The Commission for Marine Meteorology (CMM), now known as the Joint WMO/IOC Technical Commission for Oceanography and Marine Meteorology (JCOMM), agreed at its twelfth session (Havana, Cuba, 10–20 March 1997) that the *Guide to the Applications of Marine Climatology* (WMO-No. 781) should, in future, comprise two distinct sections, namely a static part and a dynamic part, to enhance its utility and facilitate the updating process. The static part, the existing Guide, WMO-No. 781, would remain unchanged, while the dynamic part would be updated approximately every four years. In keeping with this agreement, the session supported a proposal to convene a self-funding workshop to provide input for the dynamic part of the *Guide*.

The workshop, CLIMAR99 – WMO International Workshop on Advances in Marine Climatology, took place in September 1999 in Vancouver, British Columbia, Canada, and was hosted by the Meteorological Service of Canada, with additional sponsorship from WMO, NOAA's Office of Global Programs (OGP) and the National Weather Service (NWS), NOAA, USA. Several papers presented to the workshop were subsequently peer-reviewed. The first session of JCOMM (Akureyri, Iceland, 19–29 June 2001) recommended that these papers, as well as a paper on Beaufort equivalent scales by Mr Ralf Lindau (Germany), which was requested by the twelfth session of CMM, be published as the dynamic part of the *Guide*.

The present technical document contains all of these papers, reproduced in their revised form following the peer review process. It thus essentially represents the dynamic part of the *Guide to the Applications of Marine Climatology*. It is intended that this dynamic component should be updated again on the basis of papers presented at a second CLIMAR workshop.

The coordination and organization of CLIMAR99, as well as the compilation of this report, were undertaken largely by Mr Val Swail (Canada). On behalf of WMO, I would like to express my gratitude and appreciation to Mr Swail and to all those who contributed to the workshop and this report.



(G.O.P. Obasi)
Secretary-General

ACKNOWLEDGEMENTS

Among the many to whom thanks are due, first and foremost are those who have contributed their time and effort in preparing for the CLIMAR99 Workshop and compiling this report, for example, Val Swail (Canada), Joe Elms (USA), Henry Diaz (USA) and all the authors of the papers contained in this report.

Also to be thanked are the individuals who have contributed without being named specifically. Workshop participants and session chairs are thanked for their important contribution to the workshop discussions and/or effective management of the workshop programme. Members of the former WMO Commission for Marine Meteorology and its Subgroup on Marine Climatology are also thanked for their support.

The cooperation and support given to CLIMAR99 by the following organizations is gratefully acknowledged: the Meteorological Service of Canada; the Office of Global Programs, NOAA (USA), and its Environmental Services Data and Information Management program; the National Weather Service, NOAA (USA); and the University Corporation for Atmospheric Research (USA).

Last but not least, thanks are due to the support staff of the above-mentioned organizations for their patient efforts in preparing for the Workshop and compiling this report, in particular Eileen Jeffers (Canada), Maria Carney (Canada) and the technical editing staff of the WMO Secretariat.

INTRODUCTION

PURPOSE AND BACKGROUND OF THIS REPORT

The purpose of this report is to provide all Members with access to new technology related to marine climatology and emerging issues such as climate change.

The WMO *Guide to the Applications of Marine Climatology* (WMO-No. 781) was published in early 1995. However, much of the material in the *Guide* was written based on knowledge and information available up until around 1992. Since then, there has been a significant amount of research carried out by national agencies in a number of Member countries. Initiatives such as the Comprehensive Ocean-Atmosphere Data Set (COADS) Wind Workshop in 1994 (held in Kiel, Germany) have also contributed much valuable information on marine climatology.

As a result, the Subgroup on Marine Climatology of the former WMO Commission for Marine Meteorology (CMM), in developing a plan and procedures to maintain the *Guide* as up-to-date as possible, decided to adopt the basic approach now being implemented for the *Guide to Climatological Practices*, namely that it would comprise two parts:

- (i) A static part, which would be expected to remain valid over a relatively long time-frame, and which could be maintained as a hard-copy publication;
- (ii) A second, more dynamic part covering matters relating to new technologies and emerging issues such as climate change which could be made available in digital form on the World Wide Web.

The Subgroup on Marine Climatology proposed that the most effective way to provide up-to-date information for the dynamic part of the *Guide* would be through a WMO-sponsored workshop, to be held in 1999, focusing on specific technologies and issues, the Proceedings of which would be made available initially on the World Wide Web, and eventually in hard copy. The issues of particular importance at that time were deemed to be climate change and impacts on the marine industry, the use of satellite data in climate applications, and issues related to retrospective data, metadata and the production of long-term homogeneous data sets for climate analysis.

At the twelfth session of CMM (Havana, Cuba, 10-20 March 1997), a recommendation was made for WMO to organize a self-funding workshop, to serve primarily as a means for generating appropriate input for the dynamic part of the *Guide to the Applications of Marine Climatology*.

The Commission requested the president of CMM, and the chairman of the Subgroup, in consultation with the Secretary-General, to take the necessary actions for convening the Workshop, in particular the task of finding a host country and potential external sponsors. Subsequently, Canada volunteered to host the Workshop and appointed Mr Val Swail to take the lead in its organization.

An informal meeting was held on 18 September 1997 in Toledo, Spain, immediately after the Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, to initiate the planning of the Workshop. A Workshop Organizing Committee, consisting of Mr Val Swail (Canada, Chair), Mr Joe Elms (USA), Mr Henry Diaz (USA) and Mr Fernando Guzman (WMO Secretariat) was established to organize the Workshop, which was named CLIMAR99 – WMO Workshop on Advances in Marine Climatology. Subsequent planning meetings were held in Boulder, Colorado (USA), on 12-14 June 1998 and Toronto, Ontario (Canada), on 19-21 April 1999. At the kind invitation of the Government of Canada, the Workshop was scheduled to take place on 8-15 September 1999, in Vancouver, British Columbia (Canada).

The Workshop was organized in conjunction with a workshop on the NOAA COADS Project. Sponsorship was obtained from WMO, NOAA's Office of Global Programs (OGP) and its Environmental Services Data and Information

Management (ESDIM) program, the National Weather Service, NOAA, and the Meteorological Service of Canada.

The objectives of the workshop were defined as follows:

- To receive appropriate input for the dynamic part of the new version of the WMO *Guide to the Applications of Marine Climatology*, with particular emphasis on new technologies;
- To review the requirements of users for new marine climate products and enhanced climate information;
- To provide guidance and technical support for those National Meteorological Services with responsibilities under the Marine Climatological Summaries Scheme (MCSS);
- To make a further contribution to the data and metadata of COADS.

A 'call for papers' was distributed to WMO Members and the general marine climate community. The format of the Workshop called for selected invited presentations from experts in the respective fields. Shorter, relevant contributions were also accepted from the general scientific community. This resulted in more than 70 abstracts being submitted by experts from every Regional Association of WMO for consideration by the Organizing Committee. The final programme was developed from these abstracts. The Workshop itself was a huge success with more than 80 participants, and it generated a great deal of interest in holding in the such meetings in the future. The final versions of all the papers will be published by WMO in the JCOMM report series.

A subset of the papers that addressed the primary objectives of the Workshop was subsequently identified by an editorial committee. These papers were then subjected to a peer review process. The first session of JCOMM (Akureyri, Iceland, 19-29 June 2001) agreed to publish the revised versions of these papers as the dynamic part of the WMO *Guide to the Applications of Marine Climatology*. This report thus constitutes the dynamic part of the *Guide*

The content of this report is organized as follows. Section 1 describes recent enhancements to marine climate databases. Section 2 contains information on the evaluation of various marine data sources; also included in this section is a report for the Subgroup on Marine Climatology by the rapporteur Mr R. Lindau on Beaufort equivalent scales. Section 3 describes data quality and metadata issues. Section 4 deals with the development and use of satellite marine databases for winds and waves. Section 5 concerns various aspects relating to the analysis of climate variability and change, while Section 6 describes some user requirements for marine climate information and its applications.

SECTION 1

MARINE DATABASE ENHANCEMENTS

COADS updates and the blend with the UK Met Office Marine Data Bank	3
The Kobe Collection: newly digitized Japanese historical surface marine meteorological observations	11
An archive of underway surface meteorology data from WOCE	20

COADS UPDATES AND THE BLEND WITH THE UK MET OFFICE MARINE DATA BANK

S.D. Woodruff⁽¹⁾, S.J. Worley⁽²⁾, J.A. Arnott⁽³⁾, H.F. Diaz⁽¹⁾,
J.D. Elms⁽⁴⁾, M. Jackson⁽³⁾, S.J. Lubker⁽¹⁾, and D.E. Parker⁽³⁾

ABSTRACT Ongoing data and metadata enhancements to the Comprehensive Ocean-Atmosphere Data Set (COADS) are described, including the blend with the UK Met Office Marine Data Bank (MDB). MDB data available since 1854 are being used, together with other new or improved sources, to enhance data coverage and quality within the presently available period-of-record (1854-1997). In addition, some newly available historical data will be used to extend coverage back to 1784. Data composition and coverage are discussed, and future plans outlined, including improved products to help address data continuity problems arising from observational, instrumental, and processing changes. Improved and expanded metadata also are becoming available as part of an upgrade of the COADS web site (<http://www.cdc.noaa.gov/coads>), which includes details about how to request data products.

INTRODUCTION The Comprehensive Ocean-Atmosphere Data Set (COADS) is the most extensive set of surface marine meteorological data presently available for the world ocean, now covering the 1854–1997 period. Surface meteorological observations from ships of opportunity are available for the entire period-of-record. These have been supplemented in more recent years by increasing amounts of data from moored and drifting buoys, oceanographic Research Vessels (R/Vs) and fishing vessels.

Extensive efforts also are underway to enhance the quality and completeness of earlier ship records. These include a blend of COADS with the UK Met Office Marine Data Bank (MDB), and national and international efforts to digitize additional logbook data and metadata (Diaz and Woodruff, 1999). The updated data are providing crucial input for the Intergovernmental Panel on Climate Change (IPCC) Scientific Assessments, and internationally for several centres that compute global atmospheric reanalyses.

Sections 2 and 3 of this paper describe efforts over the last several years toward a complete replacement and update of COADS Release 1 (1854–1979) (Slutz *et al.*, 1985; Woodruff *et al.*, 1987), including improved observational and summary products.

COADS provides a relatively uniform database for a wide variety of scientific investigations, and its products are distributed openly and without restrictions. These characteristics have been critical in developing broad international participation.

Section 4 highlights variations in data composition and coverage of the presently available 1854–1997 data. These variations are compounded by other data inhomogeneities arising, for example, from instrumental, observational and processing changes. Section 5 outlines future plans, including data and metadata improvements targeted to help address these issues.

RECENT AND ONGOING UPDATES RELEASE 1A (1980–97)

Release 1a data, originally completed for 1980–92 (Woodruff *et al.*, 1993), have been updated and extended several times in response to requirements for reanalysis projects and demands from the user community for updated products. The most recent update, completed in June 1999, involved a complete reprocessing of

(1) NOAA/OAR Climate Diagnostics Center, Boulder, Colorado, USA

(2) National Center for Atmospheric Research, Boulder, Colorado, USA

(3) Hadley Centre, Meteorological Office, Bracknell, UK

(4) NOAA/NCDC, Asheville, North Carolina, USA

the previously available data for the 1980–95 period, plus an extension through to 1997.

A major element of the 1980–97 update was the blend with the MDB for 1980–94. The Met Office implemented and documented a conversion from MDB ‘flatfile’ formats into the Long Marine Report (LMR) format used for COADS production processing (a fixed-length LMRF format is generally distributed to users). The Met Office provided 17.9 million MDB reports for 1980–94, and 7 per cent of these data were retained as unique or judged to be of preferable quality compared with data already in COADS (Figure 1), and British Navy decks were retained at relatively high rates compared to other MDB sources (Table 1).

For earlier periods we estimate that gains from the blend will be higher: Woodruff (1990) estimated 63 per cent duplicate and 28 per cent/9 per cent unique from COADS/MDB based on tests against Release 1 data for six sample 10° latitude \times 10° longitude boxes.

Some major enhancements were made as part of the 1980–97 update to data from moored and drifting buoys. Global Telecommunication System (GTS) receipts after 1990 from Tropical Atmosphere–Ocean (TAO) moorings were replaced by ‘standard archive’ data obtained directly from NOAA’s Pacific Marine Environmental Laboratory (PMEL) which offer improved diurnal coverage and quality controls; the Pilot Research Moored Array in the Tropical Atlantic (PIRATA) is also now part of this archive. Canada’s Marine Environmental Data Service (MEDS) corrected 1980–85 drifting buoy data for a day misassignment problem (+1, only impacting buoys reporting in the last quarter of the day). Also, for 1993–97, processing changes at Service Argos necessitated some modifications in the handling of MEDS quality control information to obtain increases in the available drifting buoy data.

Other new or improved data sources were included in this update. Sea surface temperature (SST) estimates derived from the uppermost levels of oceanographic profiles, and some surface meteorological fields, were added through 1996 from the World Ocean Database 1998 (WOD98; Levitus *et al.*, 1998). Additional data from Russia’s large marine archive of ship data (MARMET) were included in 1995, as well as Russian Arctic and Antarctic Research Institute (AARI) North Pole (NP) Station (manned drifting ice floe) data through 1991, obtained through the University of Washington’s Polar Science Center.

RELEASE 1B (1950–79)

This update (completed in November 1996) pre-dated our work on the blend with the MDB (to be fully implemented in a future update of this time period), but provided improvements in data quality and coverage in comparison to previously available Release 1 data for 1950–79. Data additions included Russian MARMET

Figure 1—Monthly bars show the numbers of ship reports (from LMRF) output for 1980–97, received via the Global Telecommunication System (GTS) or in delayed mode (generally keyed logbook data).

Within the logbook category, curves show the total retention of MDB data (dark), and within that the number of unique MDB reports (light); i.e. “total retention” includes unique MDB reports plus others that were considered preferable duplicates (compared with COADS, plus possibly with other MDB reports).

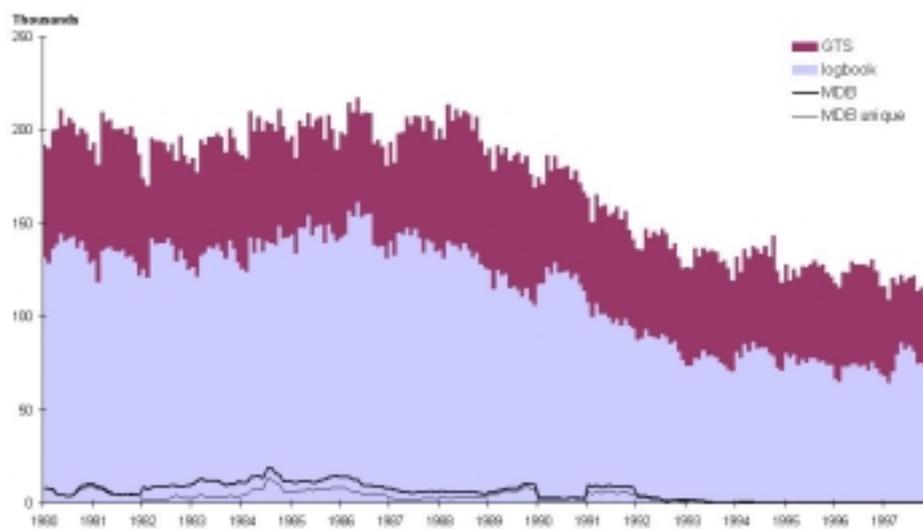


Table 1—Numbers (K=thousands and M=millions) and percentages of MDB data by deck (“series” number) output (retained, e.g. as unique or “best” duplicate) in the blend with COADS, 1980–97. Overall, 93 per cent of the reports were eliminated from the MDB as duplicates.

Deck	Description	Reports	Output LMR: %
221	“MARIDS” and trawlers	1.55K	6
223	Selected ships	1.52K	<1
224	Ocean Weather Stations	59	6
229	British Navy (HM) ships	10.6K	92
233	Selected ships	48.8K	2
234	Ocean Weather Stations	781	4
239	British Navy (HM) ships	42.6K	34
254	Int. Maritime Met. (IMM)	1.18M	8
255	Undocumented sources	2.78K	8
Total		1.29M	7

and NP data and an earlier oceanographic archive designated as the World Ocean Atlas 1994 (WOA94; Levitus and Boyer, 1994).

Significant data corrections were also made during Release 1b processing, including the correction of widespread temperature biases in GTS records (see Woodruff *et al.*, 1998). In 1999, a minor correction to Release 1b October–November 1970 data also was made to remove a small number of mislocated GTS observations.

RELEASE 1C (1784–1949)

We are nearing completion of the reprocessing of this time period (planned by early 2001). The update will blend MDB data for 1854–1949 (including 0.5M newly keyed 1935–39 UK merchant data), Russian MARMET data back to 1888, and about one million recently keyed reports from Japan’s Kobe Collection (concentrated in the Pacific) to enrich the data sparse period around the First World War (Manabe, 1999).

We will also include data from several recently digitized collections (described in more detail in Diaz and Woodruff, 1999 and Elms *et al.*, 1999): the US Maury Collection (covering 1784–1863, but concentrated around 1830–60; 1.3M reports), the Norwegian Logbook Collection (1867–89; 201K), the US Merchant Marine 1912–46 Collection (3.5M), Arctic Drift Stations (1893–1938; 16K) and the Russian S.O. Makarov Collection (1804–91; 3,500 reports).

However, conversions of early logbook data to modern units in the LMR format involve complex and scientifically important translation issues, and resources to implement the conversions are limited. Therefore, some conversions may have to be delayed until a future update, or be scaled back to key data elements, which are considered to be SST, sea level pressure (SLP), air temperature and wind.

PRODUCTS: OBSERVATIONS, STATISTICS, METADATA

Releases 1a and 1b already offer significant improvements in comparison to the original Release 1 products, as discussed in sections 3.1 and 3.2. Completion of Release 1c will make the entire archive available in uniform observational and summary formats. Improved metadata are also being made available as part of these updates, as discussed in section 3.3. The COADS web site (<http://www.cdc.noaa.gov/coads>) provides links to, or information on how to request, data and metadata products.

INDIVIDUAL MARINE REPORTS (OBSERVATIONS)

To develop Release 1a, the LMR format used for COADS production processing was updated. Also, a fixed-length LMRF format was developed to satisfy the majority of current user requirements for individual observations and to replace the Release 1 Compressed Marine Report (CMR) format. The number of data fields was expanded in LMRF, compared to CMR, and critical new metadata, such as platform type and identification, were added to track the increasing number and diversity of data sources.

The LMRF archive is maintained in a packed-binary format, which offers computational and storage-volume efficiencies and is appropriate for long-term global studies with multiple variable requirements. However, for smaller scale studies (temporally or spatially), a simple ASCII format is more portable between computers and is easier to use. As part of the blend with the MDB we have started to design an abbreviated ASCII format that would include basic data elements and most probably some pre-applied quality controls (i.e. suspect data elements would be eliminated).

We note that a highly abbreviated ASCII format is also available to meet requirements for observational data that are more recent than those available in COADS (i.e. now later than 1997). These near-real-time data are updated on a monthly basis by NOAA's National Centers for Environmental Prediction (NCEP) and can be accessed through the COADS web site.

MONTHLY SUMMARY
STATISTICS

The COADS observations are statistically summarized for each year-month and in 2° or 1° latitude × longitude boxes. Table 2 summarizes the temporal and geographical coverage, and box resolution, for the most popular statistical products. Detailed descriptions of these and secondary products are available from the COADS web site.

The Monthly Summary Trimmed Groups (MSTG) product was developed during Release 1 and is available for the full period-of-record. User suggestions, and our experience, resulted in the development of an improved product, the Monthly Summary Groups (MSG), which contains more statistics and variables (see Woodruff *et al.*, 1998) and will fully replace MSTG upon the completion of Release 1c.

Studies of events such as the 1982–83 El Niño revealed that the “trimming” (quality control) processing developed for Release 1 was too conservative and resulted in the distortion or elimination of some large climate signals (Wolter, 1997). Concerns have also been raised about the effects of mixing ship data with data from other platform types such as drifting and moored buoys (e.g. Woodruff *et al.*, 1993).

To help mitigate these problems, and so that researchers can study the effects, two separate sets of statistics were computed for Releases 1a and 1b. The ‘standard’ statistics are derived from ship data only, using the restricted Release 1 trimming limits. In contrast, the ‘enhanced’ statistics are derived from ship and other platform types (e.g. drifting and moored buoys), using relaxed trimming limits to better preserve climate anomalies. Similar strategies are under consideration for Release 1c (1784–1949) statistics (discussed in Wolter *et al.*, 1999).

METADATA

Major improvements to the COADS web site were completed in June 1999. These included software, electronic documentation (e-doc) and inventories for currently available products; selected on-line publications (see the references for examples); and annual ship instrumental metadata available in digital form since 1973 gathered in WMO publication No. 47 (1955-). The WMO-No. 47 files from 1973–94 were reprocessed by Elizabeth Kent of the UK Southampton Oceanography Centre (discussed in Kent *et al.*, 1999).

<i>Product</i>	<i>Period</i>	<i>Domain</i>	<i>Resolution</i>
MSTG	1854–1997	global	2° × 2°
MSG	1950–1997	global	2° × 2°
MSG	1960–1997	global	1° × 1°
MSG	1960–1997	equatorial belt [†]	1° × 1°

Table 2—COADS year-month summary statistics: product abbreviations, present temporal and geographic availability, and latitude × longitude box resolution.

* Products: Monthly Summary Trimmed Groups (MSTG) and Monthly Summary Groups (MSG). Each MSTG (MSG) contains eight (10) statistics for each of four variables, and overall the products comprise 19 (22) observed and derived variables.

† 10.5°N–10.5°S, with gridding offset 0.5° from the global product and a row of 1° boxes straddling the equator.

The metadata available on the web site will continue to grow in the future. We plan to add early UK and US documentation that is not readily available, along with discussions and descriptions of the data problems that we analyze and answers to frequently asked questions.

DATA COMPOSITION AND COVERAGE

Ship data volume has been declining since peaks in the 1980s (Figures 1-2). This trend is presently influenced by delays in receiving keyed ship logbook data (in future, delayed-mode ship data will be provided by electronic means such as on diskettes or through telecommunication services like Inmarsat). However, most of the likely changes will stem from factors such as the decline in global ship traffic, increases in ship size and shifting shipboard priorities for reporting weather observations.

Conversely, there has been substantial growth in the number of drifting and moored buoy reports since around 1979 (Figure 2). The moored TAO array has significantly improved the quantity and quality of reports in the tropics, and drifting buoys provide vital reports, for example, in the far southern latitudes (although only SST and SLP in general).

As shown in Figure 1, the GTS contributed a relatively stable amount, but increasing fraction, to the total ship data mixture in recent decades; in 1997, delayed-mode ship data still composed over half of the mixture. Figure 3 shows leading contributions of the International Maritime Meteorological (IMM) keyed ship logbook data exchanged internationally since WMO Resolution 35 (Cg-IV) was adopted in 1963.

The pre-1950 data mixture (Figure 4) will be significantly enhanced to the extent that we will be able to include newly digitized data as part of Release 1c. Also, as illustrated in Figure 4, Dutch (deck 193) data predominate prior to the 1880s. The recovery and gravity correction of deck 193 SLP data that are available among the LMR supplementary data form an important goal for Release 1c processing to enrich SLP coverage (Figure 5).

FUTURE PLANS

Analyses of the COADS historical record have been affected by many changes since the 19th century as regards instrumentation, observing and reporting practices and processing (e.g. quality controls). Historical metadata to help address these heterogeneities are often nonexistent or incomplete. Many of these issues have been discussed in past COADS workshops (e.g. Diaz and Isemer, 1995) and in the open literature (e.g. Parker *et al.*, 1995).

In more recent decades, data from drifting and moored buoys have expanded data coverage. However, in a trend that will continue as future global ocean observing systems are implemented (Molinari, 1999), new platform types have further complicated the data mixture, and metadata are still frequently inadequate (e.g. information on past buoy arrays, although efforts have been

Figure 2—Annual bars show reports (LMRF) output for 1950–97, with the numbers from different platform types plotted in descending order, from bottom to top, of total magnitude: ship, moored and drifting buoy, Coastal-Marine Automated Network (C-MAN) and ocean R/Vs reports. The LMRF total line is sometimes larger due to miscellaneous additional platform types or reports for which platform type was not determined.

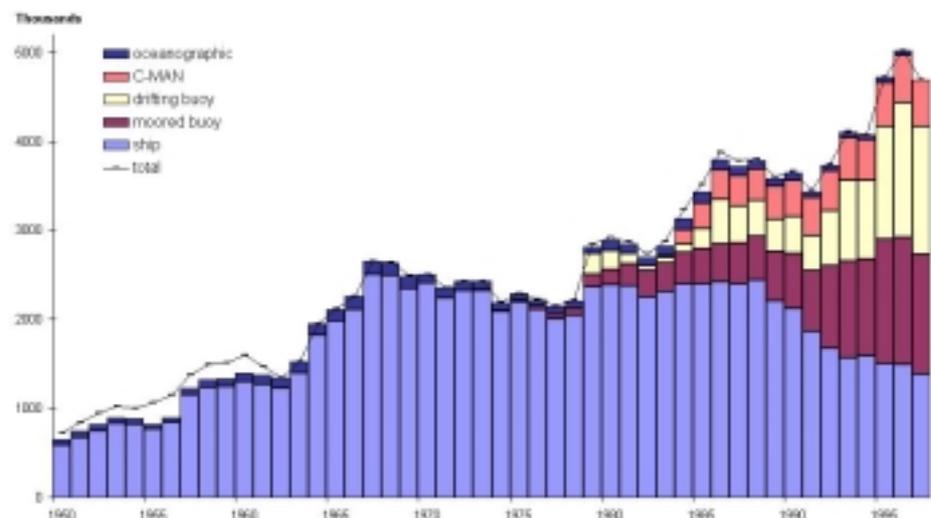


Figure 3—Annual bars show the six largest national contributions of IMM (digitized ship logbook) reports (from LMRF) output for 1960-97, plotted in descending order, from bottom to top, of total magnitude. These, plus IMM contributions from other countries, yield the total line.

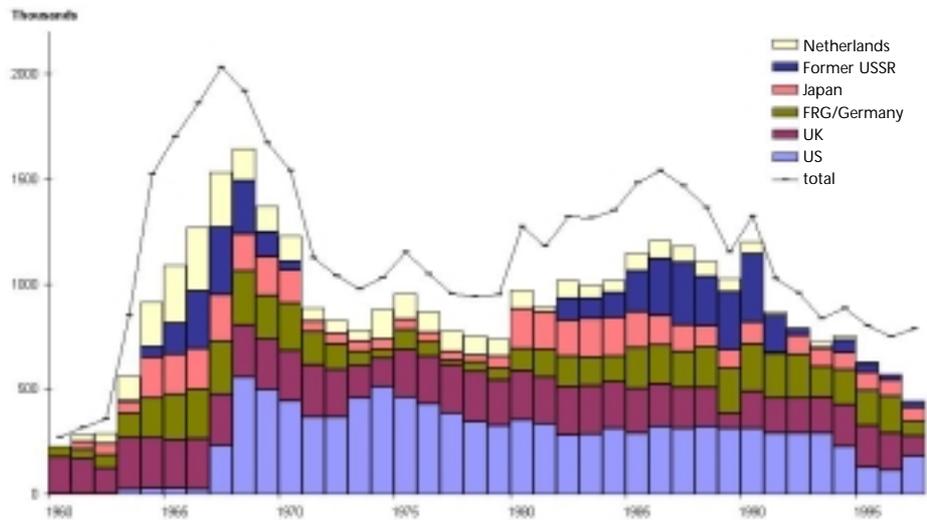
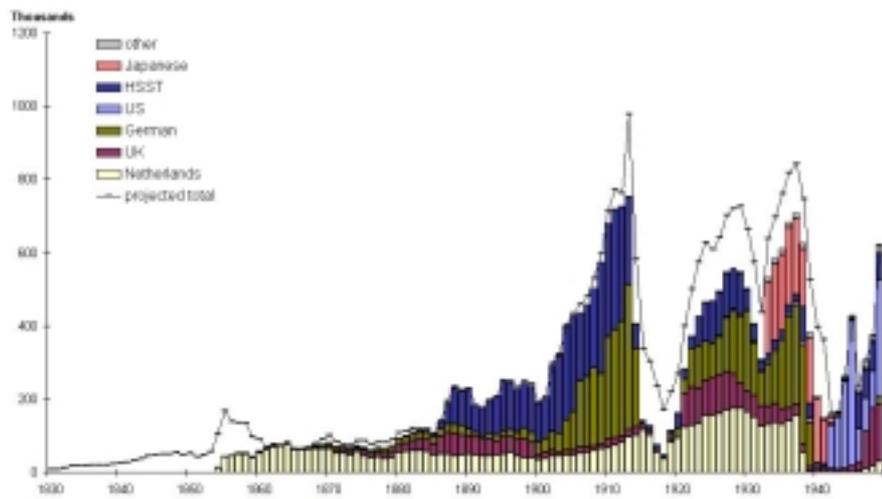


Figure 4—Annual bars show the global source deck make up of COADS Release 1 (CMR) for 1854-1949, roughly grouped according to national categories, plus Historical Sea Surface Temperature (HSST) Project data (after Figure 3 in Woodruff et al., 1998). Note the virtual absence of US data prior to 1941; and the predominance of Japanese data during 1940-41, of US data during 1942-45, and of Dutch data prior to around 1880 (it is undocumented whether the Dutch and some other early decks might actually be international compilations). The projected total line indicates the gains expected dating back to 1830 over the present COADS data from available Kobe, Maury, Norwegian, US Merchant Marine (1912-46) and Arctic Station collections (as discussed in section 2.3).



initiated to archive buoy metadata in accordance with WMO-No. 47, 1955). Even within a category such as ship data, we have the contemporary mixture of merchant and Navy, ocean R/Vs and fishing vessels which ply the oceans for different reasons and in doing so may collect data with different biases (e.g. some fishing vessels seek calm tropical regions resulting in a low wind bias).

So far we have dealt with these problems using the simplest approach. As discussed in section 3.2, separate sets of statistics are computed in an attempt to examine platform heterogeneity questions and to reduce invalid data-exclusions during climatic events with large variability.

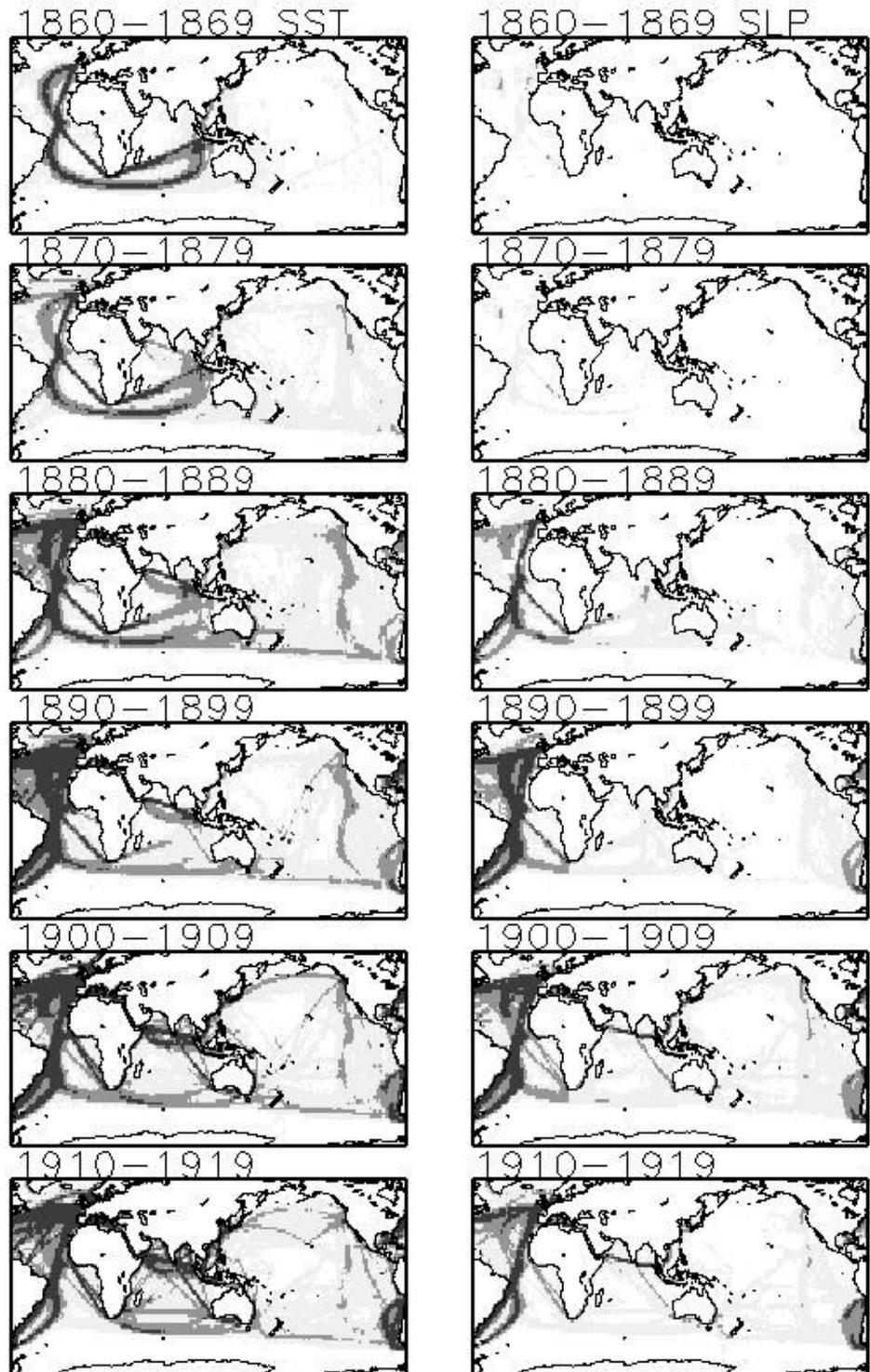
When Release 1c and the COADS-MDB blend are completed we will be in a position to process the full archive (1784 through to the late 1990s) and better address these problems. This is an important development phase for COADS. We are hoping to complete what will be known as Release 2 towards the end of 2001.

To be successful, Release 2 will require better user access to metadata and data and improved quality controls. In conjunction with new ASCII format products, temporal and spatial subsetting of the global long-term archive will be provided. As discussed in Woodruff *et al.* (1998), we plan to concentrate metadata improvements and bias adjustments on individual observations. However, calculations of some new $2^\circ \times 2^\circ$ and $1^\circ \times 1^\circ$ statistical products, including night-time air temperatures, will be developed; further separations, such as platform types and source decks, may also be desirable.

ACKNOWLEDGEMENTS

COADS is the result of an ongoing cooperative project between the National Oceanic and Atmospheric Administration (NOAA)—specifically its Office of

Figure 5—Decadal totals of SST (left) and SLP (right) observations (70°N – 78°S ; 68°W – 68°W) for six early decades of Release 1 data (after Figure 5 in Woodruff et al., 1987). Three increasingly dark shadings show at least 10, 100, or 400 observations in a $2^{\circ} \times 2^{\circ}$ box per decade, i.e. respective averages over 120 months of 0.08, 0.83, or 3.33 observations per month (note: all the observations for a decade could fall into as few as one year-month).



Oceanic and Atmospheric Research (OAR)/Climate Diagnostics Center (CDC), its National Environmental Satellite, Data and Information Service (NESDIS)/National Climatic Data Center (NCDC), and the Cooperative Institute for Research in Environmental Sciences (CIRES, conducted jointly with the University of Colorado)—and the National Science Foundation's National Center for Atmospheric Research (NCAR). The NOAA portion of COADS is currently supported by the NOAA Climate and Global Change (C&GC) programme and the NOAA Environmental Services Data and Information Management (ESDIM) programme.

REFERENCES

Diaz, H.F. and H.-J. Isemer (Eds.), 1995: *Proceedings of the International COADS Winds Workshop* (Kiel, Germany, 31 May-2 June 1994). NOAA Environmental Research Laboratories, Boulder, Colo., 301 pp.

- Diaz, H.F. and S.D. Woodruff (Eds.), 1999: *Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata* (15-17 September 1997, Toledo, Spain). WMO/TD-No.957, MMROA Report No. 43.
- Elms, J.D., S.D. Woodruff, D.E. Parker and S.J. Worley, 1999: Newly digitized historical marine data and metadata becoming available for COADS and the UK Met Office Marine Data Bank. *Preprints: WMO Workshop on Advances in Marine Climatology—CLIMAR99* (Vancouver, B.C., Canada, September 8-15, 1999). Environment Canada, Downsview, Ontario, 11-22.
- Kent, E.C., P.K. Taylor and S.A. Josey, 1999: Improving global flux climatology: the role of metadata. *Preprints: The Workshop on Advances in Marine Climatology—CLIMAR99* (Vancouver, B.C., Canada, September 8-15, 1999). Environment Canada, Downsview, Ontario, 292-299.
- Levitus, S. and T.P. Boyer, 1994: *NOAA Atlas NESDIS 4, World Ocean Atlas 1994, Volume 4: Temperature*. National Oceanographic Data Center, Ocean Climate Laboratory, Washington, DC, 117 pp.
- Levitus, S., T.P. Boyer, M.E. Conkright, T. O'Brien, J. Antonov, C. Stephens, L. Stathoplos, D. Johnson and R. Gelfeld, 1998: *NOAA Atlas NESDIS 18, World Ocean Database 1998, Volume 1: Introduction*. National Oceanographic Data Center, Ocean Climate Laboratory, Silver Spring, MD, 346 pp.
- Manabe, T., 1999: The digitized Kobe Collection, phase I: historical surface marine meteorological observations in the archive of the Japan Meteorological Agency. *Bull. Amer. Meteor. Soc.*, 80, 2703-2715.
- Molinari, R.L., 1999: Lessons learned from operating global ocean observing networks. *Bull. Amer. Meteor. Soc.*, 80, 1413-1419.
- Parker, D.E., C.K. Folland and M. Jackson, 1995: Marine surface temperature: observed variations and data requirements. *Climatic Change*, 31, 559-600.
- Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set; Release 1*. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 268 pp.
- World Meteorological Organization, 1955-: *International List of Selected, Supplementary and Auxiliary Ships*. WMO-No. 47, Geneva, Switzerland.
- World Meteorological Organization, 1963: *Fourth World Meteorological Congress Abridged Report with Resolutions*. WMO-No. 142, Geneva, Switzerland.
- Wolter, K., 1997: Trimming problems and remedies in COADS. *J. Climate*, 10, 1,980-1,997.
- Wolter, K., S.J. Lubker, and S.D. Woodruff, 1999: Quality control in recent and pending COADS releases. *Preprints: WMO Workshop on Advances in Marine Climatology—CLIMAR99* (Vancouver, B.C., Canada, September 8-15, 1999). Environment Canada, Downsview, Ontario, 312-318.
- Woodruff, S.D., 1990: Preliminary comparison of COADS (US) and MDB (UK) ship reports. *Observed Climate Variations and Change: Contributions in Support of Section 7 of the 1990 IPCC Scientific Assessment*, Parker, D. (Ed.), WMO/UNEP publication.
- Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, 68, 1239-1250.
- Woodruff, S.D., S.J. Lubker, K. Wolter, S.J. Worley and J.D. Elms, 1993: Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. *Earth System Monitor*, 4, No. 1, 1-8 [On-line at: <http://www.cdc.noaa.gov/coads/coads1a.html>].
- Woodruff, S.D., H.F. Diaz, J.D. Elms, and S.J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, 23, 517-527 [On-line at: http://www.cdc.noaa.gov/coads/egs_paper.html].

THE KOBE COLLECTION: NEWLY DIGITIZED JAPANESE HISTORICAL SURFACE MARINE METEOROLOGICAL OBSERVATIONS

Teruko Manabe, Japan Meteorological Agency, Tokyo, Japan

1. INTRODUCTION

The Kobe Marine Observatory (formerly the Imperial Marine Observatory), a field office of the Japan Meteorological Agency (JMA), collected and stored surface marine meteorological observations reported by ships in log sheets over the period from 1890 to 1961. This data set is called the Kobe Collection (Komura and Uwai, 1992; Uwai and Komura, 1992; Manabe, 1999a).

In the Collection, the data obtained by merchant ships, fishing boats and research vessels amount to around 6.8 million observations. Figure 1 shows annual numbers of ships and reports of these data. In addition, reports by Japanese Imperial Navy ships, which cover the period from 1903 to 1944, amount to around 5 million. Unfortunately, the annual number of reports by Navy ships were not counted and, therefore, could not be included in Figure 1.

In 1960 and 1961, log sheets of the data obtained by merchant ships, fishing boats and research vessels were copied onto microfilm (364 rolls in total) under the JMA-NOAA (National Oceanic and Atmospheric Administration) joint project. In this project all the data taken after 1933 (about 2.7 million in 185 rolls of microfilm) were digitized. These digitized data were already included in the Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1 (Slutz *et al.*, 1985). However, until recently, pre-1933 data and Navy data had not been digitized.

2. DIGITIZATION PROJECTS SINCE 1995

At present, COADS is one of the most complete marine meteorological data sets. However, even COADS does not contain a large amount of observations made before the 1950s, especially during the two World Wars (Woodruff *et al.*, 1987). Furthermore, since most data are from US and European ships, which mainly have ship routes in the Atlantic Ocean, there are fewer data for the Pacific than the Atlantic. The Kobe Collection covers the period of the First World War, and the main ship routes of Japanese vessels are in the Pacific. Therefore, digitization of the Kobe Collection was one of the most urgent and important projects amongst the many anticipated projects related to data archeology and rescue.

In the fiscal year (FY) 1995, JMA began to digitize the pre-1933 surface marine observations made by merchant ships in the Kobe Collection, with with cooperation from the Japan Weather Association (JWA) and the Nippon Foundation. So far, the series of efforts can be divided into three phases: Phase I is digitization in FY

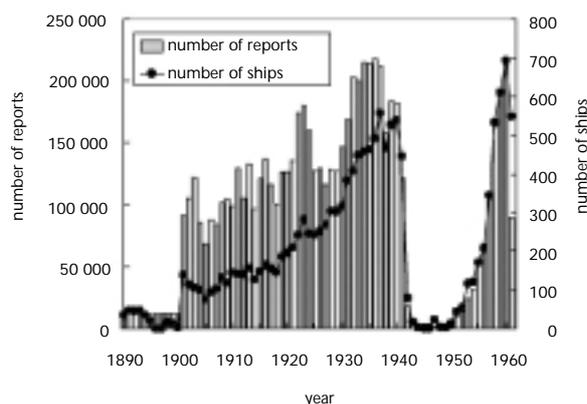


Figure 1—Yearly distribution of reports and ships in the Kobe Collection (excluding Navy data from 1890 to 1960). For the period from 1890 to 1900 annual bars were estimated by dividing by the number of years per period since only a total figure was available.

1995/96 and its quality check in FY 1997; Phase II is digitization in FY 1997/98 and its quality check in FY 1998/99; and Phase III is digitization in FY 1999/2000 and its quality check planned for 2001. In Phase I, a total of 1 045 682 reports were digitized and quality checked, and these data are now available on CD-ROM (1998 edition). In Phase II, a total of 571 472 observations were digitized in 1998/99 and quality checked, and these observations will soon be available on CD-ROM (2000 edition). Phase III was carried out from 1999 to 2000, and it is estimated that a total of around 600 000 observations will be digitized.

Because of a particular shortage of data during the First World War in the currently available COADS, the JMA devoted its efforts to digitizing the reports for the period from 1915 to 1917, which is dealt with in Phase I.

3. THE DATA SOURCE AND REMARKS ABOUT THE DIGITIZATION

Figure 2 shows an example of one of the oldest log sheets, which was reported in 1890. The observational elements in the first format include: date and time (local) of the observation; ship location; wind direction and Beaufort force; air pressure; temperature indicated by a thermometer attached to a barometer; dry-bulb and wet-bulb temperatures (Fahrenheit); cloud; present weather; direction and height of wind waves; sea surface temperature (SST); specific gravity of sea surface water; direction and speed of sea surface current; and any remarks. The format of log sheets was changed several times and the biggest change was made in 1923. Following this change, columns were added to describe the type of barometer and its instrumental correction value, as well as columns for new observational elements such as the direction and height of swell. Figure 3 shows an example of a log sheet of 1930.

During the data period (1890-1933), the number of observations taken each day varied according to the ship. Before 1923, most ships made observations six times a day at 0200, 0600, 1000, 1400, 1800, 2200, and after 1923 most ships made them four times a day at 0000, 0600, 1200, 1800. Some ships made observations four times a day at 0000, 0400, 1200, 1600, while some made them three times a day at 0400, 1200, 2000 or 0800, 1200, 2000 and others twice a day at 0000, 1200. All times were local.

Often, the ship's exact location (latitude and longitude) was logged only once a day, especially in early data. In Phase I digitization, reports without latitude/longitude information were not keyed. No reports were digitized for ships which navigated relatively close to Japan and made observations mostly at port (ships were requested to make observations even at port in the early days) and/or reported latitude/longitude once a day or less (e.g. every three or seven days).

Figure 2—Sample of one of the oldest log sheets (1890). Yamashiro Maru left Yokohama on 9 May heading for Honolulu. The observational elements are: date and time (local); ship position (latitude and longitude); wind (direction and force); barometer (and attached thermometer reading); thermometer readings (dry-bulb and wet-bulb); clouds; weather; waves (direction and height); sea water (temperature and density); and sea surface current (direction and speed).

Figure 3— Log sheet of 1930. Buenos Aires Maru was en route from Galveston to Los Angeles via the Panama Canal as part of a round the world trip. Observational elements are: date (date and time); location (latitude and longitude); wind (direction and force); barometer (corrected reading and attached thermometer reading); temperature (air-temperature, sea surface temperature); clouds; weather; visibility; wind waves (direction and height); swell (direction and height). There are five questions at the bottom of the log sheet: (1) Is the barometer mercury or aneroid? (2) Date of last check of the barometer? (3) Value of instrumental correction? (4) Height of barometer? (5) Was the instrumental correction applied?

This is the main reason why a lot of data, especially those prior to 1901, were not entered into the final digitized data set. Also, many reports in the early days were not keyed because it was difficult to read the handwriting. No reports from the the log sheets of 1891, 1895, 1896 and 1897 were digitized in the Phase I project.

The digitization procedures are as follows. First, the microfilms to be digitized were selected and printed out to produce hard-copies of the original log sheets. Besides the period from 1915 to 1917, microfilms were selected evenly for the whole period (1890–1932) so that various types of log sheets can be seen. Secondly, some elements, such as weather and units of temperature and air pressure, were coded before the keying process was initiated. After coding the information from the log sheets, the data on the coded log sheets were keyed to make an interim-file format designed to include nearly all the information contained in the log sheets. The interim-file format retains the original values (e.g. 32-point scale directions, Beaufort numbers, temperatures in Fahrenheit, weather and visibility). Finally, the digitized data in the interim file were converted into the International Maritime Meteorological Tape format, Version 1 (IMMT-1) (WMO, 1990), which does not retain the original values, but it can be easily handled and is widely distributed.

Then, quality controls were conducted on the digitized data in accordance with the minimum quality control standards of the WMO Marine Climatological Summaries Scheme (MCSS) (WMO, 1990). Each ship's track was examined by checking the ship's speed and land/sea information in the global ocean. Furthermore, in the North Pacific, air temperature, SST and dew point temperature were compared with JMA climatology (JMA, 1993a). No data comparisons with climatology were made in other ocean basins. Approximately 5 per cent of all the checked data were manually corrected. The errors appeared to be mainly caused by misinterpretations of the handwritten logs.

Specific remarks on several observational elements can be summarized as follows.

3.1 SHIP TIME AND LOCATION

Reports without location were not keyed. The date and location were checked by examining each ship's track and were manually corrected by referring to the original log sheets, where possible.

3.2
SEA SURFACE TEMPERATURE

It has often been pointed out that instrumental bias is quite influential when dealing with historical SST observations (e.g. Folland and Parker, 1995). It was assumed that all the observations were obtained using buckets since the data are pre-1933 and engine room intake measurements appeared for the first time in the *Guide to Weather Observations for Ships*, 1956 edition (JMA, 1956). However, the material (e.g. canvas, wood or rubber) of each bucket cannot be identified because, unfortunately, there is no remaining documentation on the SST measurement method used on each ship.

The readings were written in Fahrenheit or Celsius in the log sheets, and observations in Fahrenheit were converted into Celsius (Appendix A).

3.3
AIR TEMPERATURE AND DEW
POINT TEMPERATURE

Ships were requested to report dry-bulb and wet-bulb temperatures. The readings were made in either Fahrenheit or Celsius, and all the observations were converted into Celsius in the IMMT format (Appendix A).

Provided that both dry-bulb and wet-bulb temperatures were reported and the reported values were consistent, the dew point temperature was calculated when converting data into the final IMMT format (Appendix B).

3.4
AIR PRESSURE

Air pressure was measured using aneroid or mercury barometers. The readings were reported in inchHg or mmHg and all observations were converted into hPa in the final format (Appendix C). When mercury barometers were used, the reading of the attached thermometer was also written in the log sheets. It is assumed that temperature, gravity and scale corrections were applied before reporting, in accordance with the *Manual on the Marine Meteorological Observation* (Imperial Marine Observatory, 1921).

In the conversion process from the interim-file format to the final IMMT format, height corrections were added when a height was written in the log sheet (Appendix C).

3.5
AIR PRESSURE

Wind direction was reported and keyed in 32-point scale and was then converted into 36-point scale according to Table 1. Wind speed estimated visually was reported based on the old Beaufort scale as shown in the *Manual on the Marine Meteorological Observation* (Imperial Marine Observatory, 1921). Each scale was converted into knot according to Table 2.

3.6
WIND WAVE AND SWELL

Wave direction and wave height were reported. Direction was reported and keyed in 32-point scale and then converted into 36-point scale (Table 1). The height of wind waves and swell height were reported according to the JMA wind wave scale and the JMA swell scale, respectively. The wind wave scale and swell scale were converted into units of 0.5 m according to Tables 3 and 4, respectively.

Table 1—Conversion of 32-point scale to 36-point scale.

32-point scale		36-point scale		32-point scale		36-point scale	
Code	Description	Code	Description	Code	Description	Code	Description
00	Calm	00	Calm	17	S by W	19	185°-194°
01	N by E	01	5°-14°	18	SSW	20	195°-204°
02	NNE	02	15°-24°	19	SW by S	21	205°-214°
03	NE by N	03	25°-34°	20	SW	23	225°-234°
04	NE	05	45°-54°	21	SW by W	24	235°-244°
05	NE by E	06	55°-64°	22	WSW	25	245°-254°
06	ENE	07	65°-74°	23	W by S	26	255°-264°
07	E by N	08	75°-84°	24	W	27	265°-274°
08	E	09	85°-94°	25	W by N	28	275°-284°
09	E by S	10	95°-104°	26	WNW	29	285°-294°
10	ESE	11	105°-114°	27	NW by W	30	295°-304°
11	SE by E	12	115°-124°	28	NW	32	315°-324°
12	SE	14	135°-144°	29	NW by N	33	325°-334°
13	SE by S	15	145°-154°	30	NNW	34	335°-344°
14	SSE	16	155°-164°	31	N by W	35	345°-354°
15	S by E	17	165°-174°	32	N	36	355°-4°
16	S	18	175°-184°	99	Unknown	99	Unknown

Table 2—Conversion of Beaufort scale into knot.

<i>Beaufort number</i>	<i>Beaufort scale Description term</i>	<i>Wind speed (metres per second)</i>	<i>knot Wind speed (knot)</i>
00	Calm	=<0.3	00
01	Light air	0.3 - 1.5	02
02	Slight breeze	1.6 - 3.3	05
03	Gentle breeze	3.4 - 5.4	08
04	Moderate breeze	5.5 - 7.9	13
05	Fresh breeze	8.0 - 10.7	18
06	Strong breeze	10.8 - 13.8	24
07	High wind	13.9 - 17.1	30
08	Gale	17.2 - 20.7	37
09	Strong gale	20.8 - 24.4	44
10	Whole gale	24.5 - 28.4	51
11	Storm	28.5 - 33.5	59
12	Hurricane	33.6=<	68

Table 3—Conversion of wave height in JMA wind wave scale into units used in IMMT.

<i>JMA wind wave scale number</i>	<i>JMA wind wave scale Description</i>	<i>Equivalent wave height (feet)</i>	<i>IMMT Wave height (units of 0.5 metre)</i>
0	Dead calm	0	00
1	Very smooth	<1	00
2	Smooth	1-2 (1 ≤ <2)	01
3	Slight	2-3	02
4	Moderate	3-5	02
5	Rather rough	5-8	04
6	Rough	8-12	06
7	High	12-20	10
8	Very high	20-40	18
9	Phenomenal	40≤	24

Table 4—Conversion of wave height in JMA swell scale into units used in IMMT.

<i>JMA swell scale number</i>	<i>JMA swell scale Description</i>	<i>Equivalent height (feet)</i>	<i>IMMT Height of swell (units of 0.5 metre)</i>
0	No swell	0	00
1	Slight swell	1-3 (1=< <3)	01
2	Moderate swell	3-5	02
3	Rather rough	5-8	04
4	Rough swell	8-12	06
5	Heavy swell	12-20	10
6	Very heavy swell	20-40	18
7	Abnormal swell	40=<	24

Table 5—Conversion from clouds in tenth into Oktas.

<i>Clouds in tenth</i>	<i>Clouds in Oktas</i>
0	0
1	1
2	2
3	2
4	3
5	4
6	5
7	6
8	6
9	7
10	8

3.7 In accordance with the conversion table based on present and old manuals (WMO, 1995; Imperial Marine Observatory, 1921), total cloud amount in tenths was converted into Oktas according to Table 5.

3.8 These elements are not included in the data on the CD-ROM (1998 edition).

WEATHER AND VISIBILITY

3.9 SHIP IDENTIFIER

Since we did not have the 'call signs' used by ships during the data collection period, each ship was allocated a 'ship number'. The 'Ship number' has five digits; the first two digits correspond to the last two numbers of the year (e.g. 1910 corresponds to 10xxx); the last three digits correspond to the ship's position in the alphabetical listing of the names of every ship that reported data during that particular year. 'Ship numbers' and 'Ship names' are catalogued in the *Guide Book of the Japanese Marine Surface Data* (United States Weather Bureau and Japan Meteorological Agency, 1960).

4. TEMPORAL AND SPATIAL DISTRIBUTION OF THE DIGITIZED DATA

Figure 4 shows the geographical distribution of the data available on the CD-ROM (1998 edition). The data are mainly distributed in the North Pacific, especially along the main ship routes: Japan-northern America, Japan-Hawaii-California, and so forth. For all the digitized reports, 82.8 per cent, 11.5 per cent and 5.7 per cent are in the Pacific, Indian and Atlantic Oceans, respectively (each basin corresponds to the information contained in Figure 3.2 of Slutz *et al.*, 1985).

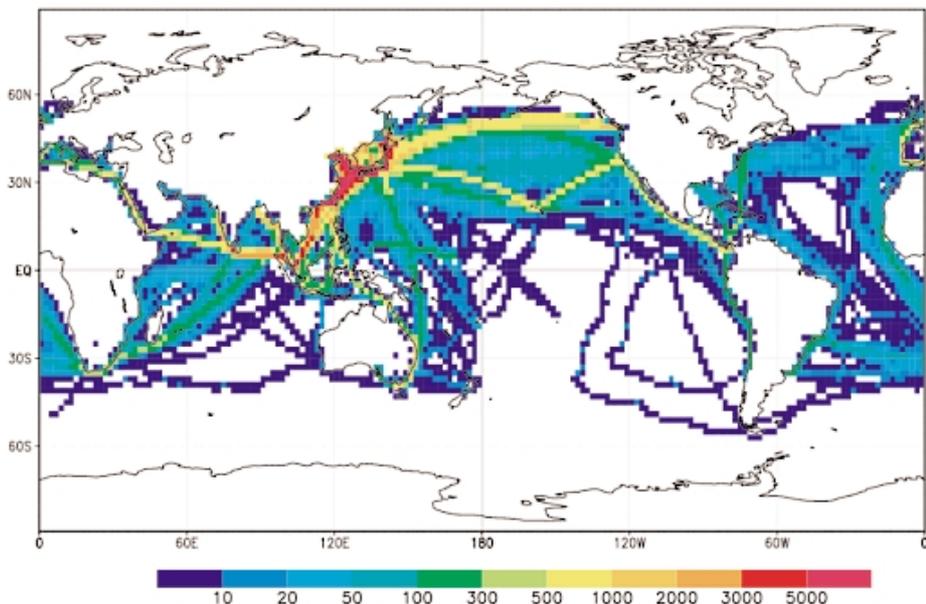
Figures 5 (a) and 5 (b) show the yearly distribution of the data of the Kobe Collection, excluding the Japanese Imperial Navy data, and COADS in the global ocean and the Pacific Ocean, where a large part of the Kobe Collection exists, respectively. In the Pacific Ocean, the presently available COADS, shown by dotted and light hatched areas, has a significant jump in the amount of data between 1932 and 1933. This is because COADS already includes the Kobe Collection data (1933–1961) which were digitized until 1961. The newly digitized Kobe Collection data (1890–1932) significantly increased the amount of available data, especially in the Pacific Ocean. It was also discovered that the amount of data covering the First World War was greatly increased by adding the newly digitized Kobe Collection data to COADS.

To show the effectiveness of the newly digitized data, a preliminary analysis was carried out (Manabe, 1999b). By using the newly digitized data along with the presently available COADS, $2^\circ \times 2^\circ$ monthly, seasonal and annual SST anomalies were calculated. In the North Pacific, data coverage (per cent of the number of grid boxes with data) increased from 5-40 per cent to 20-60 per cent from 1910 to 1933 compared with those made from COADS alone. Thanks to the increase of grid boxes with data, it was possible to apply empirical orthogonal function (EOF) analysis to data from before the Second World War as well as that from after the war. This showed that the Pacific decadal oscillation founded by Tanimoto *et al.*, 1993 can be observed back to the beginning of this century.

5. FUTURE PLAN ON DIGITIZATION

During the series of digitization projects supported by the Nippon Foundation, more than one million marine meteorological observations taken for the period from 1890 to 1932 in the global ocean, especially in the North Pacific, have been

Figure 4—Geographical distribution of reports for the whole data period (from 1890 to 1932) available on the CD-ROM (1998 edition). Each $2^\circ \times 2^\circ$ box is shaded according to the number of reports. The total number of reports appearing on this figure is 1 045 682.



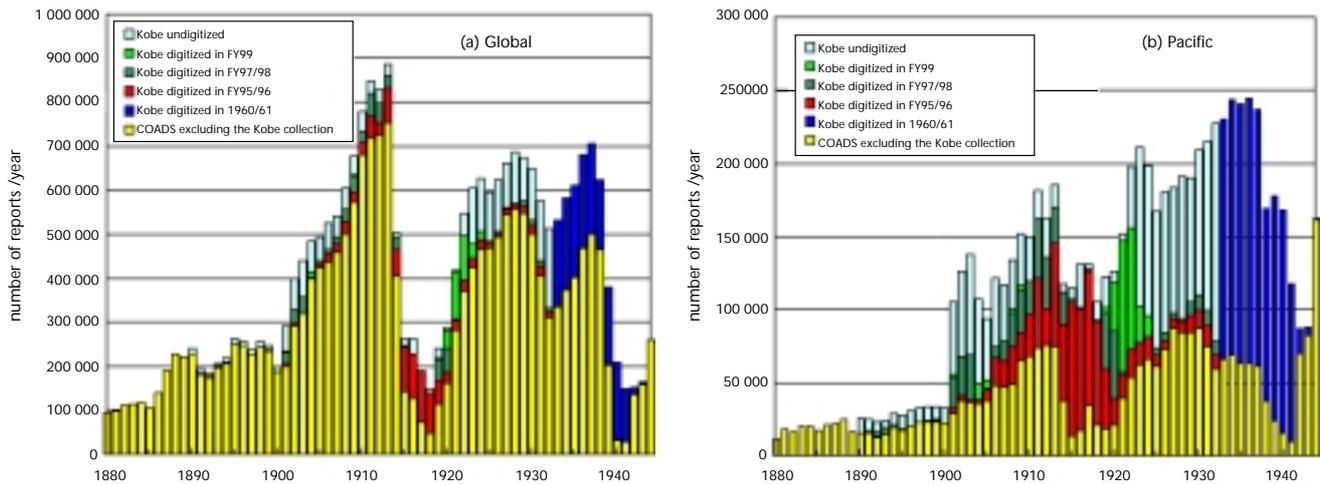


Figure 5—Yearly distribution of the reports in the Kobe Collection, excluding the Japanese Imperial Navy data, and COADS in the global ocean (a) and in the Pacific (b). The CD-ROM (1998 edition) contains the Kobe digitized in 1995/96, which is shown in black. The presently available COADS is shown by “COADS, excluding the Kobe Collection” and “Kobe digitized in 1960/61”. Because the data digitized in FY 97/98 have not been quality checked, the number of these data have not been fixed.

5.1 Digitization projects

newly made available. The biggest contribution of the digitization projects has been to increase the usable marine meteorological observational data in the North Pacific for the period around the First World War. It is expected that the newly digitized Kobe Collection will be widely used.

Following the publication of the CD-ROM (1998 edition) in 1999, a CD-ROM (2000 edition), which contains a total of 5 71 472 observations digitized in Phase II, will be published at the end of FY 2001. Phase III digitization was carried out in FY 1999 and FY 2000. It is expected that around 600 000 observations will be made available in Phase III. However, over 1 million records still needed to be digitized at the end of FY 2000. In cooperation with JWA, the JMA is making efforts to continue to digitize as many data as possible in the period following 2000. To make more historical data available, it is expected that Phase IV of this series of digitization projects will start in 2001.

Also, JMA is exploring the possibility of increasing the number of observations in the data set by including many of the ‘no-position’ reports, for which a ship’s position can be estimated by interpolation. This could substantially increase the number of observations in some regions.

5.2 DATA FORMAT FOR HISTORICAL MARINE DATA

For the distribution of the digitized Kobe Collection in Phase I, the IMMT format was adopted since it is easy to handle. However, because this format is designed for the storage and exchange of contemporary marine data (from 1961 onwards) in WMO, it is not well suited to historical data. For example, there are no columns for the thermometer attached to the barometer, the original Beaufort number, or the original units of data. Furthermore, as regards wind and wave direction, a 36-point scale is used in the IMMT, whereas a 32-point scale was often used in historical data, and the conversion from 32-point to 36-point scale could cause problems. It would be very helpful if a data format could be agreed upon that is well suited to historical and modern data and is easy to handle.

In WMO, it is recognized that while efforts have intensified to digitize the additional historical ship data that exist in many national log book collections, such as the Kobe Collection, there is no effective internationally agreed format for the exchange of keyed historical data. Efforts are being made to develop an International Marine Meteorological Archive (IMMA) format that is well suited to historical and modern data and is easy to handle (WMO, 2000).

Once the series of digitization projects are complete, all the digitized data are planned to be made available in a new format (IMMA, if possible) which is more suitable for historical data than the IMMT.

5.3
JAPANESE IMPERIAL NAVY
DATA

With regard to the Japanese Imperial Navy data which covers the period from 1903 to 1944, according to a preliminary investigation on the data, it seems that only about 10 per cent of all the reports (about 5 million reports) include location information. However, considering that these data cover the data sparse period, which includes the two World Wars, digitizing the navy data would be a valuable exercise. Thus, JMA is trying to find a way to rescue these data which will make it possible to interpolate missing location data so that as many data as possible will become available.

ACKNOWLEDGEMENT

K. Nishiyama and JMA staff assisted in many ways. The author wishes to express her thanks to all these people, especially O. Shimada, A. Shoji, M. Kaneda, Y. Yokote and A. Wada.

The digitization projects have been supported by the Nippon Foundation, with the cooperation of JWA. R. Yamamoto, K. Hanawa, K. Kutsuwada and Y. Matsuyama were instrumental in the promotion of the project.

With the kind permission of the American Meteorological Society, portions of this manuscript have been reprinted from the "Digitized Kobe Collection, Phase I: Historical Surface Marine Meteorological Observations in the Archive of the Japan Meteorological Agency", which appeared in the Bulletin of the American Meteorological Society, No.80.

APPENDICES
CONVERSION ALGORITHMS

A.

Conversion of temperature

Temperatures observed in Fahrenheit were converted into Celsius using the following equation:

$$T_c = (T_f - 32) / 1.8$$

where T_c is the temperature in Celsius, and T_f is the temperature in Fahrenheit.

B.

Calculation of dew point
temperature

When dry-bulb and wet-bulb temperatures and air pressure (station pressure) were known, dew point temperature was calculated in accordance with the *Guide to Surface Meteorological Observations* (JMA, 1993b). First, saturation vapour pressure at the wet-bulb temperature was obtained using the Goff-Gratch formulae described in the WMO *Technical Regulations* (WMO, 1988). Secondly, vapour pressure at the observed wet-bulb and dry-bulb temperatures and air pressure were calculated using the following Sprung formula using the obtained saturate vapour pressure:

$$e = E' - P(T - T_w)A / 755$$

where $A = 0.5$ when the dry-bulb is not iced, $A = 0.44$ when the dry-bulb is iced, e is vapour pressure in hPa, E' is saturated vapour pressure in hPa at the wet-bulb temperature, T is the dry-bulb temperature in Celsius, and T_w is the wet-bulb temperature in Celsius.

Finally, dew-point temperature is extracted from a table on the *JMA Surface Meteorological Tables* (JMA, 1959 (see 1986 edition for amendments)) produced on the basis of the above-mentioned Goff-Gratch formulae.

C.

Conversion of air pressure

Air pressure was observed either in mmHg or inchHg. In the interim-file, barometer readings in mmHg or inchHg remain. The interim-file format has a column to show the units of air pressure (mmHg or inchHg). Air pressure in mmHg or inchHg was converted into hPa and a height correction was also made during the conversion process. Thus, air pressure at sea level in hPa was written in the IMMT format. However, if a height was not written in the log sheet, the pressure values that were simply converted into hPa were written in the IMMT format without the height correction. Readings were converted into hPa in accordance with the following equations and then rounded off to the closest hPa:

when air pressure was observed in inchHg:

$$P = b \times 33.8639 + 1013.25 \times (\exp(1.17972 \times 10^{-4} \times H) - 1)$$

when air pressure was observed in mmHg:

$$P=b \times 1.33322 + 1013.25 \times (\exp(1.17972 \times 10^{-4} \times H) - 1)$$

where P is air pressure at sea level (hPa), b is the barometer reading, and H is height of the instrument (metres).

REFERENCES

- Folland, C.K. and D.E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, 121, 319-367.
- Imperial Marine Observatory, 1921: *Manual on the Marine Meteorological Observation* (in Japanese). Imperial Marine Observatory, 82 pp.
- Japan Meteorological Agency, 1956: *Guide to Weather Observations for Ships* (in Japanese), Marine Department, Japan Meteorological Agency, 119 pp.
- Japan Meteorological Agency, 1959: *JMA Surface Meteorological Tables* (in Japanese), Japan Meteorological Agency, 144 pp. (eighth amendment was made in 1986)
- Japan Meteorological Agency, 1993a: *Marine Climatological Charts of the North Pacific Ocean for a Thirty-Year Period (1961-1990)*. Marine Department, Japan Meteorological Agency, 191 pp.
- Japan Meteorological Agency, 1993b: *Guide to Surface Meteorological Observations* (in Japanese), Japan Meteorological Agency, 167 pp. (fourth amendment was made in 1999)
- Komura, K. and T. Uwai, 1992: The collection of historical ships' data in Kobe Marine Observatory. *Bull. Kobe. Mar. Obs.*, 211, 19-29.
- Manabe, T., 1999a: Digitization of the Kobe Collection: The historical surface marine meteorological data collected by the Kobe Marine Observatory, Japan Meteorological Agency. *Proceedings of the International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata* (Toledo, Spain). H.F. Diaz and S.D. Woodruff, Eds., WMO-Technical Document No. 957.
- Manabe, T., 1999b: The Digitized Kobe Collection, Phase I: Historical surface marine meteorological observations in the archive of the Japan Meteorological Agency. *Bull. Amer. Meteor. Soc.*, 80, 2703-2715.
- Sultz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set (COADS); Release 1*. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, Colorado, 268 pp. (NTIS PB86-105723)
- Tanimoto, Y., N. Iwasaka, K. Hanawa and Y. Toba, 1993: Characteristic variations of sea surface temperature with multiple time scales in the North Pacific. *J. Climate*, 6, 1153-1160.
- United States Weather Bureau and Japan Meteorological Agency, 1960: *Guide Book of the Japanese Marine Surface Data*, United States Weather Bureau and Japan Meteorological Agency, 527 pp.
- Uwai, T. and K. Komura, 1992: The collection of historical ships' data in Kobe Marine Observatory. *Proceedings of the International COADS Workshop* (Boulder), H.F. Diaz, K. Wolter and S.D. Woodruff, (eds.), NOAA Environmental Research Laboratories.
- Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, 68, 1239-1250.
- World Meteorological Organization, 1988: *Technical Regulations*, Volume I, General meteorological standards and recommended practices, WMO-No.49.
- World Meteorological Organization, 1990: *Manual on Marine Meteorological Services*, Volume I, Global aspects, WMO-No.558.
- World Meteorological Organization, 1995: *Manual on Codes – International Codes*, Volume I.1, Part A: Alphanumeric codes, WMO-No.306.
- World Meteorological Organization, 2000: *Final Report, Subgroup on Marine Climatology, Eighth Session*, JCOMM Meeting Report No. 2.

AN ARCHIVE OF UNDERWAY SURFACE METEOROLOGY DATA FROM WOCE

David M. Legler*, Shawn R. Smith, James J. O'Brien, Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida 32306-2840 USA

INTRODUCTION The World Ocean Circulation Experiment (WOCE) involved nearly 100 Research Vessels (R/Vs) and the participation of over 40 countries during a 10-year programme to measure the general circulation of the ocean, as well as to improve our understanding of the role of the ocean in climate. The WOCE planning process included the establishment of several distributed data centres to develop reporting methodologies and criteria for each observing system (centered primarily around measurement type) and to assemble and quality control all relevant WOCE data (WOCE International Project Office, 1997). A Data Assembly Center (DAC) for underway and moored surface meteorological data was established in the Center for Ocean-Atmospheric Prediction Studies (COAPS) at Florida State University (FSU) in support of WOCE. The mission of the FSU DAC is to collect, check, archive and distribute all surface meteorology data from the international R/Vs that participated in the WOCE programme as well as surface meteorological data from moored and drifting buoys deployed under WOCE. The FSU DAC has now established a unique archive of quality-reviewed surface meteorological data from WOCE cruises. The types of surface meteorology data processed include data from automated systems that record a wide variety of data at much higher frequencies that are not found in other data sets. We will highlight our assembly, quality-review, and management methodologies. The contents of the archive will be discussed as well as potential applications such as validating remotely-sensed data/products and identifying errors in atmospheric model fields over the ocean. Finally, questions regarding the incorporation of these data into the Comprehensive Ocean-Atmosphere Data Set (COADS) will be discussed.

DATA ASSEMBLY Surface meteorological data were recorded during most WOCE cruises. Data reporting requirements were established for reporting WOCE surface meteorological data (e.g. Joyce and Corry, 1994), but were not widely followed. We relied on cruise reports to indicate whether meteorological data were routinely recorded. Data were then pursued through contact with scientists in charge of the cruise and/or through ship support groups at home institutions. Most data were obtained through exhaustive efforts using telephones, facsimilies, the post and the e-mail as contact means. The collection process has been very successful for nearly 70 per cent of the pre-1998 WOCE-specific data at FSU (Table 1). For a modest number of cruises no information has been forthcoming, even after several attempts to confirm reports of 'meteorological data recorded'. Data from some cruises were lost due to a variety of legacy problems such as file formats written by 'someone who no longer works for us', media degradation, etc. The collection of metadata on how the observations were recorded was equally important and includes instrument type (if any), installation height (depth) and other information. The metadata were equally difficult to obtain owing to a lack of reporting standards and difficulties in locating sources of knowledge about the instrument systems. Our experience indicates that reporting standards and requirements should be updated to reflect technological advances, particularly for automated systems. Additionally, these requirements should be more widely distributed to the research vessel community.

* Corresponding author now at U.S. CLIVAR Office, 400 Virginia Ave, SW, Suite 750, Washington, DC 20024, USA [legler@usclivar.org]

Table 1—Data collection status at FSU DAC for surface meteorology. Entries reflect the number of segments (and percentage of the total number of segments). A segment is a subset of a cruise by one R/V and reflects the organization of data from cruises during the WOCE programme.

<i>Date of WOCE cruises (years)</i>	<i>Number of WOCE cruise segments for which data are available at FSU DAC</i>	<i>Number of segments from previous column with high resolution (15 minute means or faster) data</i>
pre-1989	6	0
1989	18	3
1990	35	10
1991	70	24
1992	81	32
1993	62	21
1994	73	36
1995	52	46
1996	12	12
1997	7	7
1998	2	2
Total	418	193 (47%)

The typical surface meteorological data set in the archive includes values for wind speed and direction, barometric pressure, humidity, air temperature, sea temperature, and for some installations, precipitation, and various radiation components. On some ships there may be more than one set of instruments. In this case, data from all instruments are included in the files. There are primarily two types of data in our archive. The first type are relatively low temporal resolution data that may be based on bridge observations. These observations are normally reported every several hours and are similar in nature to those found in COADS. We have focused especially on the second type of R/V data, i.e. those from automated instrument systems that record observations much more frequently. A typical automated system records one minute means of wind speed and direction, barometric pressure, humidity, air temperature, sea temperature, precipitation and short wave radiation (long-wave is optional), as well as several supporting variables. Examples of these automated systems are the IMET systems installed on several US platforms (Hosom *et al.*, 1995). Similar installations are found on ships from the UK, Germany, and Australia. Data from moored platforms are also becoming part of the archives, e.g. the WOCE Subduction Experiment had four moored buoys equipped with IMET systems for a two-year period.

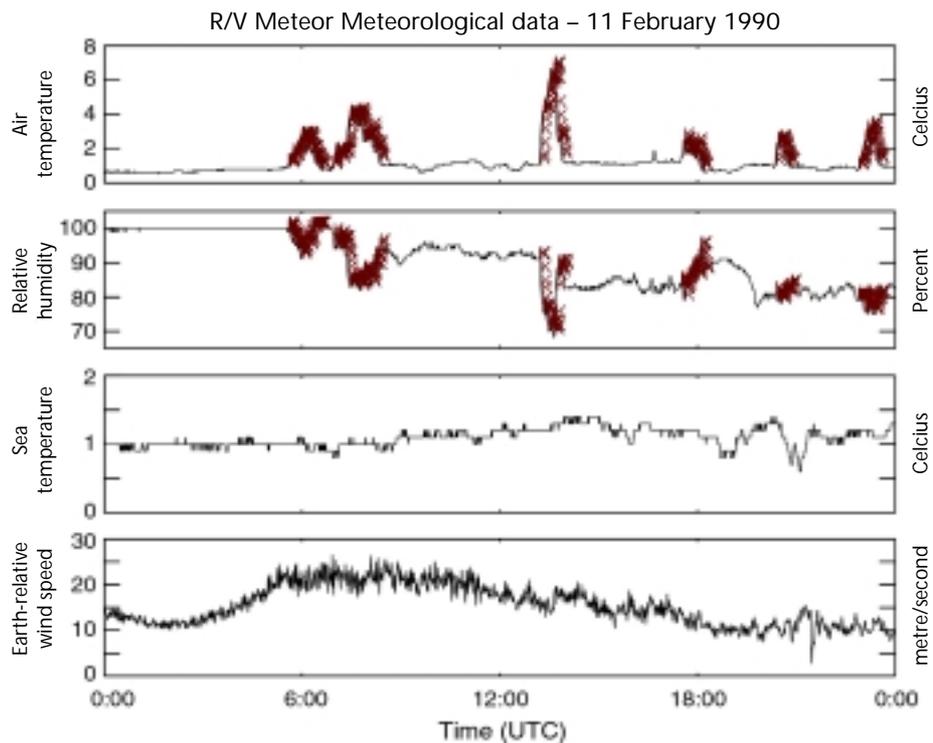
QUALITY CONTROL

The data were quality controlled using a series of statistical and graphical analysis software tools to identify suspect data (e.g. spurious data, time shifts, gaps, biases and instrument drifts). Suspect data were brought to the attention of the upstream data supplier and flagged according to the nature of the identified problem or error (Smith *et al.*, 1996). Figure 1 shows that an example of the types of errors found were spurious jumps in temperature and humidity data records of the research vessel Meteor (from our QC reports, Smith *et al.*, 1996). We received confirmation from the ship operator that these errors are caused by instrumentation mounted near the stacks such that during select periods when the orientation of the ship and the wind are aligned, the warm moist conditions over the stack pollute downstream instruments. The subsequent errors in sensible heat flux for this case are 300 W m⁻², thus demonstrating the importance of flagging suspect observations. Note that the high temporal resolution (reports every minute) of the data made it possible to confidently identify this problem. Comparable data from GTS and/or COADS (available each ~6 hours) would never indicate a problem even though one might be present. Quality control flags (including one for an interesting value, such as an extreme value verified through independent data) are included in the data files with explanations and descriptions of various data problems discussed in a quality control report that is written for each WOCE cruise.

DATA DISTRIBUTION

The WOCE community can access data through a wide variety of distribution media (e.g. electronic networks, magnetic media, interactive requests, printed

Figure 1—Underway data from the R/V Meteor with suspect data indicated by overlaid alphabetic code. Note that the Meteor was in the North Atlantic in February 1990.



reports, etc.). Documentation (i.e. metadata) on observational data and processing by DAC is also available. Data for just under half of the WOCE cruises have been published on a series of CD-ROMs, i.e. Version 2 of the WOCE Global Data Set (WOCE Data Products Committee, 2000a,b) (Figure 2). Updates are available on our web site (www.coaps.fsu.edu/WOCE). More complete versions of the WOCE Global Data Set will be produced. Although data from numerous WOCE cruises have yet to be delivered to FSU, to date nearly 50 million observations have been obtained, quality controlled and distributed through electronic (WWW/FTP) and CD-ROM media, thereby making this the largest uniformly formatted collection of surface meteorology data from research vessels.

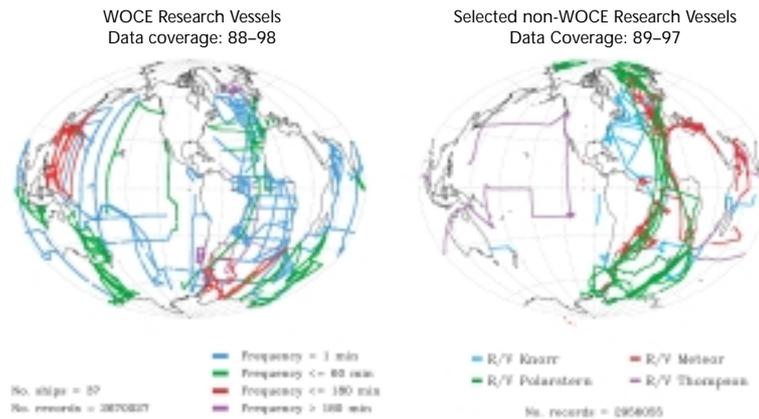
UNIQUENESS

This data archive of underway surface meteorological data has several valuable attributes. Much of the data are unique in that they originate from automated instrument systems that record observations at a relatively high time resolution (averaging periods of seconds to minutes) and are not reported via the GTS. The high-frequency recording aspect allows a more rigorous quality control review of the data and thus results in data with a higher level of quality. Additionally, the higher resolution data enables additional applications to be considered that would not be possible (i.e. more difficult to complete) with ordinary reporting at synoptic hours. In our review of data from several of the WOCE R/Vs, many observations do not appear in the COADS; neither are they reported via the GTS. High time resolution data are certainly not available in COADS or through other sources. Some R/Vs do report surface meteorological values at synoptic hours, but again the data from automated systems are not typically reported. In cases where coincident data from both automated systems and ordinary observations are reported for a single ship, the WOCE DAC data should by default be considered to be of a higher quality. Lastly, the high-quality metadata (instrument type, placement and height) make these data unique and valuable for climate studies.

APPLICATIONS

The applications of these data are varied. Process-oriented studies often require coincidental measurements of the water column as well as the surface air-sea fluxes. Some of the surface meteorological data are used to validate NSCAT surface winds (Bourassa *et al.*, 1997) and will be used to validate other remotely sensed data. Other validation work with these data is underway and will address issues

Figure 2—(left) Coverage of FSU DAC archive as of May 1998. Shades indicate temporal frequency of data recordings. (right) Coverage from non-WOCE data to be processed in the future. Shades indicate respective research vessels. All data on this plot are recorded at greater than 5-minute frequency.



such as optimal averaging times for recording anemometers so that remotely-sensed and in situ winds can be blended.

We have taken an active role in improving the reporting standards of some of the ships that have provided us with data. Feedback from our quality control review of the data has led to improved data recording practices, particularly for wind reporting (Smith *et al.*, 1999).

Because these meteorological data are high resolution, air-sea fluxes can be computed more confidently and with more accuracy. Consequently, we are developing methodologies to use the DAC data as an independent means for validating surface meteorology and flux products. We have begun to compare select WOCE surface meteorology observations from our DAC with surface reanalysis products from NCEP. There are certain advantages to this approach. First, the high time resolution data from our archive produce more accurate flux estimates because we can average over the same six-hour time period that is represented in the reanalysis. Additionally, we can remove suspect data and make proper adjustments to height measurements and observing methods. As previously discussed, many of the WOCE data are independent of the data stream used as input for reanalysis. These data were observed over a relatively wide range of locations; they consequently represent a wider distribution of environmental conditions under which the reanalysis and other flux products may be evaluated. Such an analysis could be completed with individual ship reports from COADS, but given the substantial errors associated with COADS data and the tremendous difficulties in gaining knowledge of how each ship observes and records these data, numerous questions would arise and compromise quantitative results (Smith *et al.*, 2001).

FUTURE PLANS AND DISCUSSIONS

Our WOCE centre has focused on completing the processing of data only from WOCE cruises. We have assembled a substantial collection of high time resolution surface meteorological data from non-WOCE cruises from many of the same ships. We have started to process these data for selected ships with the best data and coverage during the WOCE period. On the basis of the initial processing of some of these data, we estimate that we will expand our high time resolution data volume for the selected ships by nearly three million observations (Figure 2). These will provide additional surface flux data for a variety of studies. Additionally, they will increase our pool of WOCE high time resolution data for potential matches to evaluate flux products by three-fold. These additional data will also supplement the general pool of in situ data for other purposes such as remote sensor validation. It should be noted that most of the additional cruises from this collection are in the Atlantic, with quite a few being from rarely sampled regions of the Southern Hemisphere.

All of our data will be made available to NODC for final archiving. Additionally, these (and all other relevant) WOCE data are being melded into a single WOCE Global Data Resource (its final composition and structure have yet to be completely defined). Questions remain concerning the inclusion of these data in COADS and other such data collections. Much work should precede this decision to address questions of representativeness.

ACKNOWLEDGEMENTS

Thanks are due to the NSF Physical Oceanography programme for its support of the WOCE DAC for Surface Meteorology and SAC for Surface Fluxes (NSF Grants OCE9314515, 9529384 and 9818727) and the Office of Naval Research, Physical Oceanography Section, which provided COAPS base funding.

REFERENCES

- Bourassa, M.A., M.H. Freilich, D.M. Legler, W.T. Liu, and J.J. O'Brien, 1997: Wind observations from new satellite and research vessels agree. *EOS Trans. Amer. Geophys. Soc.*, 78, 597-602.
- Hosom, D.S., R.A. Weller, R.E. Payne and K.E. Prada, 1995: The IMET (Improved METeorology) ship and buoy system. *J. Atmos. Oceanic Tech.*, 12, 527-540.
- Joyce, T., and C. Corry, 1994: *Requirements for WOCE Hydrographic Programme Data Reporting*. WHPO Publication 90-1, Revision 2, WOCE Hydrographic Programme Office, WHOI, Woods Hole, MA 02543, USA, 144 pp.
- Smith, S.R., M.A. Bourassa and R.J. Sharp, 1999: Establishing more truth in true winds. *J. Atm. and Oceanic Tech.*, 16, 939-952.
- Smith, S.R., C. Harvey and D.M. Legler, 1996: *Handbook of Quality Control Procedures and Methods for Surface Meteorology Data*. WOCE Report 141/96, COAPS Report 96-1, WOCE Data Assembly Center, COAPS, Florida State University, Tallahassee, FL 32306-2840, USA, 56 pp.
- Smith, S.R., D.M. Legler and K.V. Verzone, 2001: Quantifying uncertainties in NCEP reanalysis using high-quality research vessel observations. *J. Phys. Oceanogr.*, accepted pending revision.
- WOCE Data Products Committee, 2000a: *WOCE Global Data, Version 2.0, Surface Fluxes*. WOCE Report No. 171/00, World Ocean Circulation Experiment International Project Office, Southampton, UK, Available from COAPS, Florida State University, Tallahassee, FL 32306, USA.
- WOCE Data Products Committee, 2000b: *WOCE Global Data, Version 2.0, Surface Meteorology*. WOCE Report No. 171/00, World Ocean Circulation Experiment International Project Office, Southampton, UK, Available from COAPS, Florida State University, Tallahassee, FL 32306, USA.
- WOCE International Project Office, 1997: *WOCE Data Guide 1997*. WOCE Report No. 150/97, WOCE International Project Office, Southampton, UK, 12 pp.

SECTION 2

EVALUATION OF MARINE DATA SOURCES

The accuracy of marine surface winds from ships and buoys	27
Report on Beaufort equivalent scales	41
Evaluation of ocean winds and waves from voluntary observing ship data	53
Evaluation of NCEP reanalysis surface marine wind fields for ocean wave hindcasts	68

THE ACCURACY OF MARINE SURFACE WINDS FROM SHIPS AND BUOYS

Peter K. Taylor*, Elizabeth C. Kent, Margaret J. Yelland, and Ben I. Moat, Southampton Oceanography Centre, Southampton, SO14 3ZH, UK

1. INTRODUCTION

In this paper we will review the progress made in determining the accuracy of marine wind observations since the International COADS Winds Workshop, held in Kiel, Germany in 1994 (Diaz and Isemer, 1995). Accurate marine wind data are important because, as the sea surface roughness increases with wind speed, wind stress increases roughly as $(\text{wind speed})^{2.7}$ and mixed layer deepening with $(\text{wind speed})^4$. However, a major problem is that we do not have an error free source of wind data over the ocean. Whilst it might be expected that the best data sources would be anemometer measurements from research ships, ocean weather ships (OWSs) or meteorological buoys, we shall demonstrate in section 2 that there are potential biases in each of these data types. In section 3, we will discuss the methods of wind determination used by the Voluntary Observing ships (VOSs) and then consider random errors (section 4) and systematic errors (section 5). We will demonstrate that quantitative knowledge of the errors is vital in order, for example, to compare ship and satellite winds. We shall consider how future developments may improve the accuracy of VOS winds (section 7) before summarising our conclusions and providing some recommendations (section 8).

2. LACK OF AN ABSOLUTE STANDARD

2.1 RESEARCH SHIPS

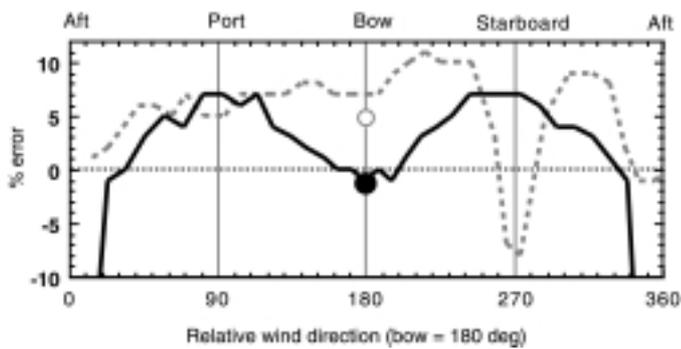
It should not be assumed that anemometer measurements on research ships are necessarily accurate. For example, before the World Ocean Circulation Experiment (WOCE), Taylor and Weller (1991) carefully specified the required underway meteorological measurements. Despite this, only one in five of the vessels recorded all the parameters needed to compute true wind, and for less than one ship in seven that calculation was applied correctly (Smith *et al.*, 1999). On ships like research ships, which are frequently moving slowly, possibly sideways or backwards, it is particularly important to log both the ship's head and the ship's course separately; this is not always appreciated.

Many research ships have a ship's anemometer which is permanently mounted, often over the wheelhouse, to give an indication of the meteorological conditions. Only for specific air-sea interaction experiments might they be equipped with accurately calibrated research anemometers, usually mounted on a special mast in the bow. Like all ships, research ships disturb the wind flow and the effect varies according to location. The results of a wind tunnel study using a model of a small research ship, CSS Dawson, are shown in Figure 1 (Thiebaut, 1990). At the ship's mainmast anemometer site the airflow is generally accelerated by 5 to 10 per cent except when the wind is from starboard (when it is in the wake of part of the mast) or from astern. Results from a computational fluid dynamics (CFD) study for bow on flow were in reasonable quantitative agreement and showed (Figure 2) that there is a large region of accelerated flow over the main accommodation block - this is typical of ships in general (section 5.3).

At the bow anemometer site the wind speed was close to the free stream value when the ship was pointed into the wind. However, for wind from either beam the wind would have been overestimated, and for winds from astern the anemometer was in the wake of the accommodation block. Had this anemometer

* Corresponding author's address: Peter K. Taylor, James Rennell Division for Ocean Circulation and Climate, Southampton Oceanography Centre, European Way, Southampton, SO14 3ZH, UK

Figure 1—Wind speed errors at the bow (solid lines) and main mast (dashed lines) anemometer sites as measured in wind tunnel studies of the CSS Dawson (Thiebaux, 1990). CFD model results are shown as a shaded circle for the bow anemometer site and a open circle for the main mast site. “Aft”, “Port”, “Bow”, “Stbd” indicates that the stern, port side, bow or starboard side of the ship is facing the wind (adapted from Yelland *et al.*, 1998b).



been mounted lower, it would have measured accelerated flow. On many ships the accommodation is nearer the bow and in such cases the bow anemometer would be in a region of decelerated flow.

Further examples of the computed flow around research ships are given by Yelland *et al.* (1998b). It is clear that obtaining accurate measurements of the mean wind requires considerable care, and that almost all ship wind data will be biased unless the airflow disturbance is taken into account.

2.2 OCEAN WEATHER SHIPS

Most OWSs were a similar size and shape to research ships. Typically they maintained their station by drifting beam on to the wind until the limit of their station ‘box’ was reached when they would steam back into the windward limit. In higher winds they would be ‘hove to’, i.e. heading into the wind at a speed just sufficient to maintain steerage way. These different operating modes would cause varying wind flow errors at the anemometer sites which were, in any case, not necessarily ideal. For example, the aft mast was used on the OWS Cumulus (which was studied by Taylor *et al.*, 1995 for the period 1987-1994 when the ship operated at 57°N 20°W). This was considered acceptable because the ship’s main purpose was to make weather observations for forecasting purposes (and now-casting and navigation for aviation) rather than to provide a climatological wind standard.

For the same reason, it is likely that corrections were not applied to the ship’s velocity through the water unless the ship was actually steaming. Taylor *et al.* (1995) used a sonic anemometer and GPS system on the OWS Cumulus to show that when the ship was drifting, the reported wind speed was too low and by slightly more than the expected amount - possibly due to flow distortion (Figure 3). When the ship was ‘hove to’ wind speeds were overestimated by approximately the expected amount. The difficulty of constructing a time series of weather ship data has been well illustrated by Isemer (1994); careful consideration of the history of observations at the OWS sites has resulted in a data set that is more consistent through time compared to VOS data (Isemer, 1995), but within which there are significant discontinuities at some sites.

Figure 2—CFD calculations for bow-on flow over the CSS Dawson. The shading indicates wind speed error, as a percentage of the undisturbed value, on a vertical fore-aft plane through the bow-mast anemometer position (shown by a cross). The numbers indicate the percentage error in each region.

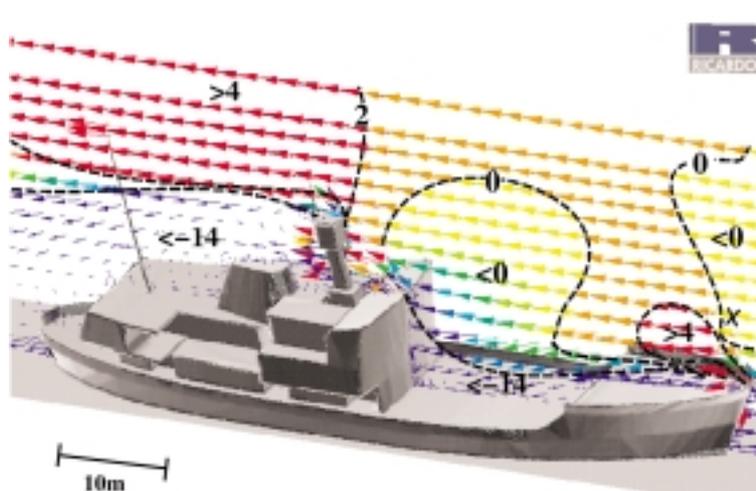
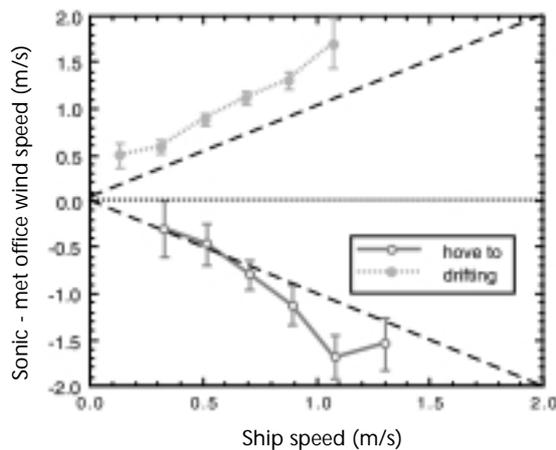


Figure 3—Average difference between wind speeds measured by research instrumentation (sonic anemometer plus GPS navigation package) and the standard WMO reports from the OWS Cumulus plotted against the ship speed from the navigation package. Cases where the ship was hove to or drifting are shown separately, the diagonal lines indicate agreement between the wind speed error and the ship's speed (from Taylor *et al.*, 1995).

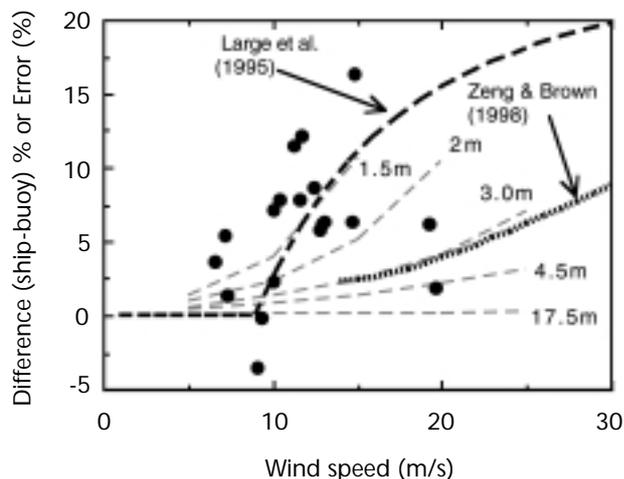


2.3 BUOY DATA

Wind speeds from meteorological buoys are believed to be biased low in strong winds (Large *et al.*, 1995; Weller and Taylor, 1998; Zeng and Brown, 1998). During the Storm Wind Study 2 experiment, SWS-2 (Dobson *et al.*, 1999; Taylor *et al.*, 1999), 10 m neutral equivalent winds were estimated using sonic anemometers on a buoy (at 4.5 m) and a nearby research ship (at 17.5 m). The comparison of the measured wind speed values is shown in Figure 4. The data are very scattered, but on average the buoy appears to underestimate the wind by about 5 per cent. There are two possible mechanisms. Firstly, assuming that the mean wind profile is logarithmic, an instrument being moved up and down vertically by the waves will measure an average wind which is less than the wind at the mean measurement height. Using the observed wave height to wind relationship for SWS-2, this effect has been crudely estimated for different anemometer heights (light dashed lines on Figure 4). Zeng and Brown (1998) noted that there were a lack of high wind speed data in buoy observations used for scatterometer calibration. They used surface air pressure data to infer a low bias for buoy winds at higher wind speeds. Their polynomial relationship (Figure 4) appears very similar to what might be expected due to the logarithmic averaging for a 3 m anemometer height - not an unreasonable mean anemometer height for the mix of buoy data that they used.

The second mechanism is that the instrument may enter regions where the vertical wind profile is distorted due to the sheltering effect of the waves. Large *et al.* (1995) suggested that the effect is to significantly bias buoy wind data for wind speeds above some threshold. Their predicted error for a 5 m anemometer height is also shown in Figure 4 and is much greater than that predicted by Zeng and Brown (1998). The preliminary SWS-2 results shown on Figure 4 appear to be of a similar order to the Large *et al.* (1995) prediction. However, the measured friction velocity values suggested that the wind error in the 20 to 25 m/s region

Figure 4—The difference (closed circles: (ship - buoy) per cent) between values of the 10 m neutral wind from buoy and ship data during the SWS-2 experiment for cases where the separation was less than 10 km (anemometer heights 4.5 m and 17.5 m respectively). Also shown are the calculated effects of vertical movement through logarithmic wind profiles for instruments at heights between 1.5 m and 17.5 m (light dashed lines); the mean error curve reported by Large *et al.* (1995) and the polynomial of Zeng and Brown (1998).



was 3 per cent to 5 per cent (similar to Zeng and Brown) rather than 15 per cent or more. The high frequency (2 Hz) data logged on the SWS-2 buoy became available just recently. These include buoy motions and wind velocities and will hopefully lead to a greater understanding of the problems related to wind measurements taken by buoys.

2.4
SATELLITE DATA

The physics of radar backscatter or microwave emission is not known well enough to allow an absolute calibration of satellite instruments so they are calibrated and verified against buoy data. Thus, if, as discussed above, the buoy data are biased, the satellite retrievals will also be biased (e.g. Zeng and Brown, 1998).

3.
METHOD OF
OBSERVATION FOR VOS
WINDS

VOS winds are either visually estimated or determined using an anemometer. In the Pacific most reports are anemometer-based (Table 1). The fraction of anemometer measurements has increased with time as has the average height of the anemometer. Because of the preference of some European meteorological agencies for visually estimated winds, the fraction of anemometer reports is significantly lower in the North Atlantic, and the anemometers are on average mounted lower. As might be expected, the anemometer height tends to be higher in the trans-oceanic shipping routes and lower in coastal regions (Kent and Taylor, 1997).

Table 1—Mean and standard deviation of the distribution of anemometer heights in January of each year indicated for the North Pacific and the North Atlantic. Also shown is the fraction of wind observations measured by anemometer (after Kent and Taylor, 1997).

Year	Mean Height (m)	Standard deviation (m)	Fraction (per cent)
<i>North Pacific (30° to 50°N, 180° to 150°W)</i>			
1980	28.7	5.9	69
1986	33.7	6.4	81
1990	35.2	8.4	82
<i>North Atlantic (30° to 50°N, 40° to 20°W)</i>			
1980	18.4	7.3	35
1986	21.5	8.9	44
1990	24.2	10.9	38

4.
RANDOM ERRORS IN VOS
WINDS

The random errors in VOS observations may be determined by the semivariogram technique which was described at this conference (Kent *et al.*, 1999b). Observations from pairs of ships are compared and the squared differences in the reported wind value are ranked according to the distance separating the ships. If enough observations are available, then the mean difference at zero separation may be determined by extrapolation. This represents twice the random error variance for a single ship observation.

4.1
METHOD OF DETERMINATION

4.2
TYPICAL ERROR VALUES

Kent *et al.*, (1999) analysed VOS observations from four months (January and July in 1980 and 1993) which they assumed to be typical of the period from 1980 to 1993 (the large computing resources needed for the calculations prevented more months from being examined). The results for wind speed are shown in Figures 5 and 6. A typical root mean square (RMS) error for a single wind speed observation was about 2.2 m/s. However, this was after instrumental observations had been corrected for the height of the anemometer above the sea surface (using the data from WMO-No. 47 and Kent *et al.*, 1999b) and visual observations corrected using the Lindau (1995) version of the Beaufort scale. For the observations as reported, the errors were about 15 per cent greater - about 2.5 m/s. This demonstrates that, despite the varying effects of air flow distortion around the ship, correcting the data for anemometer height does reduce the errors. The RMS wind speed errors appeared to be lower than average in tropical regions, however no significant dependence on wind speed was found.

Figure 5—Random observation errors for VOS wind speed reports. The upper figure is the number of report pairs used to make the estimate, the central figure is the rms error (m/s) for each 30° region, and the lower figure is the estimated uncertainty in the rms error estimate (from Kent *et al.*, 1999).

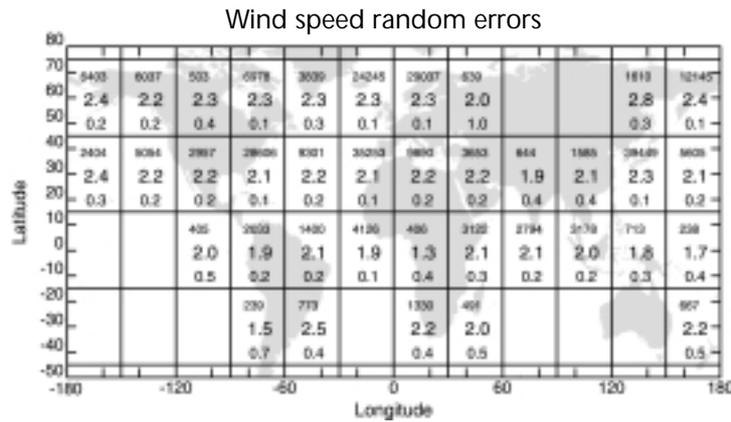
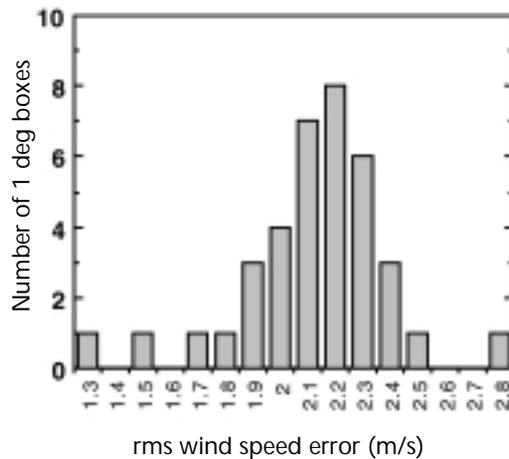


Figure 6—Histogram of the rms error estimates shown in Figure 5.



4.3 POSITION ERRORS

About 2 to 3 per cent of the VOS weather reports in the COADS (Woodruff *et al.*, 1993) collection of VOS weather reports can be identified as having incorrect position information. Typically the position is incorrect by 10° or is in the wrong quadrant. Often these data exist in COADS as duplicates, with one report having the correct position. Position errors are detected in operational forecast centres by tracking individual ships, but this is rarely done for climate studies. However, position errors are potentially very serious because the ship might be erroneously placed away from the shipping lanes in a data sparse region. Such a report may thus be given undue weight. For example, in January 1984, ship reports from near Iceland appeared as a group of erroneous duplicates in the COADS data set, positioned near Antarctica. Therefore, position errors may introduce significant errors into calculated wind fields (along with the fields of other variables).

5. SYSTEMATIC ERRORS IN VOS WINDS

5.1 METHOD OF DETERMINATION

Owing to the lack of an absolute standard, determining the systematic errors in VOS observations is difficult. The VSOP-NA (Voluntary Observing Ship Special Observing Programme - North Atlantic) project (Kent *et al.*, 1991, 1993) was designed to identify and, if possible, quantify systematic errors in the VOS data. A subset of 46 VOS was chosen, the instrumentation used on each of the participating ships documented (Kent and Taylor, 1991), and extra information was obtained with each report, for example, the relative wind at the time of the observation. The output from an atmospheric forecast model was used to compare one ship observation against another. The results were then analysed according to instrument type and exposure, ship size and nationality, and other factors.

5.2. ACCURACY OF ANEMOMETER WINDS

The VSOP-NA results showed that speed estimates from hand-held anemometers were very scattered at wind speeds above about 7m/s and that there was also a larger scatter in the direction estimates compared to other methods. The use of hand-held anemometers was therefore to be discouraged.

The VOSs in the VSOP-NA project reported the anemometer estimated relative wind speed in addition to the calculated true wind speed (only the latter is

transmitted in the standard ships weather observation). Kent *et al.* (1991) showed that a major cause of error was the calculation of the true wind speed. Only 50 per cent of the reported winds were within 1 m/s of the correct value and 30 per cent of the reports were more than 2.5 m/s incorrect (Figure 7). For wind direction, only 70 per cent were within $\pm 10^\circ$ of the correct direction and 13 percent were outside $\pm 50^\circ$. These are substantial needless errors which significantly degrade the quality of anemometer winds. A similar conclusion was reached by Gulev (1999). Results from a questionnaire distributed to 300 ships' officers showed that only 27 per cent of them used the correct method to compute true wind, 19 per cent did not know how to do the calculation, 21 per cent usually did not do the calculation and 33 per cent did it either episodically or approximately. This is perhaps not surprising given the problems in obtaining accurate true wind data from research ships (Smith *et al.*, 1999; see section 2.1 above).

Wind speed reports from VOSs are accompanied by a wind speed indicator flag which establishes whether the wind observations are a visual or anemometer report, and whether the units are knots or m/s. Any error in the indicator flag, for example resulting from miscoding or transmission, may lead to a large error in the accompanying wind report.

We have already noted (section 4.2) that correcting for the height of the anemometer above the sea demonstrably improved the data set. This correction should be done on a ship-by-ship basis since the average height of anemometers varies both geographically and with time (section 3).

For a 10 m/s wind and neutral stratification, an anemometer at 35 m will read about 10 per cent higher than one mounted at 20 m. For unstable conditions this ratio decreases. For very stable conditions one or both anemometers may be outside the near surface boundary layer, in which case the error would be indeterminate. Fortunately, very stable conditions are relatively rare over most of the ocean. For the VSOP-NA ships which used anemometers, the mean difference between the ship and model wind speed estimates increased with anemometer height even more than might have been expected due to the vertical wind profile (Figure 8).

Taylor *et al.* (1995) reanalysed the VSOP-NA results for wind speed. They found that having corrected OWS Cumulus data for ship motion and the VOS data for anemometer height, there appeared to be agreement between the OWS and VOS data for winds below 10 m/s. For higher wind speeds the VOS winds were biased high - by about 1.5 m/s to 2 m/s at 20 m/s wind speed. If this bias is real, the reasons might include misreading of the anemometer dial (gust values rather than mean winds being reported) and the air flow distortion caused by the ship.

5.3
CFD STUDIES OF AIRFLOW
DISTORTION FOR VOSs

We have noted above (section 2.1) that for ship mounted anemometers a major consideration is the air-flow disturbance caused by the ships' hull and superstructure. We have also shown that this may be determined using CFD simulation. The CFD results have been verified for wind speeds within 30° of the bow by comparisons with data from an array of anemometers on the research ships RRS Darwin and RRS Discovery. Both ships were instrumented with up to 10 anemometers located at various sites, including some regions of high flow distortion. These comparisons showed good agreement between the ships' data and the

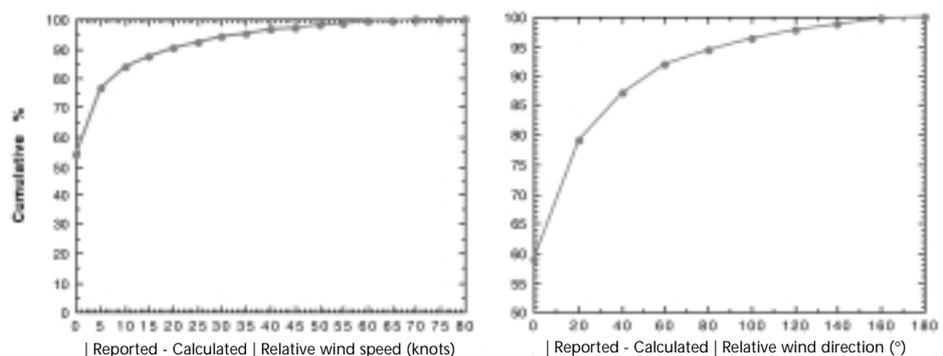
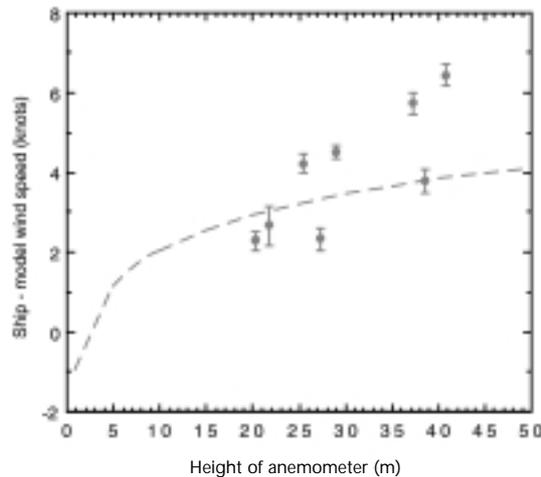


Figure 7—Cumulative histograms of the difference between the value calculated by the ship's officers and the correct value for true wind speed (right) and true wind direction (far right) (from Kent *et al.*, 1991).

Figure 8—Mean difference between the ship and model wind speed estimates for those VSOP-NA ships which used anemometers plotted against the anemometer height. Also shown is the expected variation of wind speed with height for a neutral boundary layer. This has been offset by the estimated mean error in the model winds (2 knots).



CFD results in all cases, except where the anemometers were in the wake of an upstream obstruction - a situation in which the CFD code is expected to perform poorly.

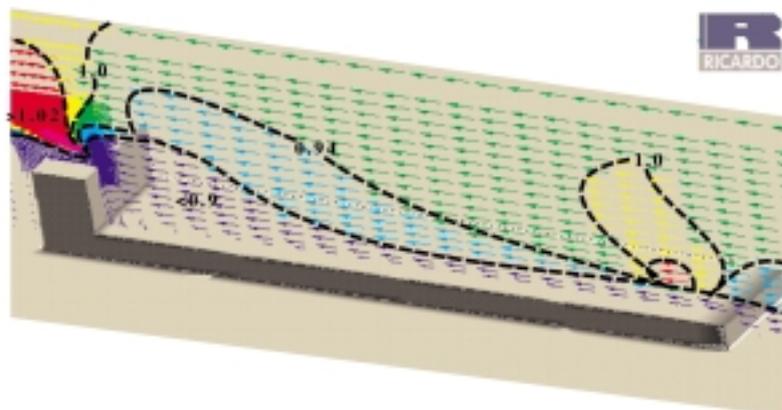
The obvious problem in applying CFD modelling to the VOS is the almost infinite variety of the size and shape of merchant ships. However, two ship types, container ships and tankers (the results of which may also be applicable to Oil Bulk Ore, or OBO ships), are believed to account for around 70 per cent of the deep-ocean merchant fleet. Since the effective shape and roughness for container ships will vary according to the degree of loading, we have chosen to study first of all the flow over tankers. Based on a sample of 36 tankers and 8 bulk carriers, three representative models were created (Table 2 and Figure 9). Tanker 1 was modelled with a close mesh to resolve the accelerated 'plume' region above the bridge; for tankers 2 and 3, a coarser mesh was used for computational efficiency.

Using the fluid dynamics analogy of flow past a rectangular block, we would expect the bridge-to-deck height (D) to be an important scaling factor. For example, the comparison between tanker 2 and tanker 3 showed a similar pattern of wind speed error for heights of less than around 8 m, but the magnitude of the decelerations differed by up to 20 per cent in profiles obtained near (i.e. within 5 m of) the front edge of the bridge. When distances were scaled by the bridge-to-deck height, these differences reduced to around 5 per cent. Indeed, all three models showed that at a height above the wheelhouse top of greater than $0.5D$ any anemometer site would give an

Table 2—Dimensions (metres) for the three tanker/bulk carrier models used in the CFD studies.

Tanker model number =	(1)	(2)	(3)
Length overall	170	250	330
Beam	27	42	62
Freeboard	6	8	10
Deck to Bridge top (D)	14	16	18
Bridge length	14	15	23

Figure 9—A three-dimensional view of a simple 'two block' tanker model. Model results of the wind speed error, expressed as the wind speed at a point divided by the free stream (undistorted) wind speed, are shown for a vertical plane intersecting the ship (Moat et al., 1998).



overestimate of the wind speed of up to 5 per cent (Figure 10). This held for all sites up to 10 m back from the front edge of the bridge (Figure 11) and would not vary with a moderate displacement to port or starboard of the centre line of the bridge.

Below a height above the wheelhouse top of 0.5D the results vary according to both anemometer position and the mesh density used in the model. The tanker 1 model (fine mesh) shows a ‘plume’ of accelerated flow, with a maximum acceleration of around 13 per cent at a height of about 4 m above the bridge (and about 4 m from the bridge front) and large decelerations below this height (Figures 11 and 12). The other two tankers do not resolve the plume and both show decelerations at heights of less than 5 or 6 m. Here we have used dimensions in metres to emphasize that an anemometer mounted above the wheelhouse may be below, in, or above the plume maximum depending on how high and how far aft it is mounted. Below the plume the wind will be significantly underestimated, above the plume an overestimate will occur. If the anemometer is in the plume the overestimate may be significant and vary rapidly with relative wind direction.

5.4 ACCURACY OF VISUAL WIND ESTIMATES

Kent and Taylor (1997) reviewed the various Beaufort equivalent scales and found that the Lindau scale (1995) was the most effective at giving similar wind speed distributions for both anemometer estimated and visual monthly mean wind data. They also confirmed Lindau’s suggestion that the characteristic biases of the earlier Beaufort scales could be explained by the statistical method by which they were derived. The ‘UWM’ scale (developed by da Silva *et al.*, 1995 at the University of Wisconsin, Milwaukee), which is similar to the Lindau scale, also performed well. It should be noted that the Lindau scale is more similar to the WMO code 1100 scale used for the observations than the so-called ‘scientific scale’ recommended by CMM-IV (see WMO, 1970).

Figure 10—The fractional wind speed error for each of the three tanker models at a distance (x) from the front of the wheelhouse where $x/D = 0.6$. The vertical scale is z/D where z is the height of the anemometer above the wheelhouse and D is the height of the bridge top above the deck.

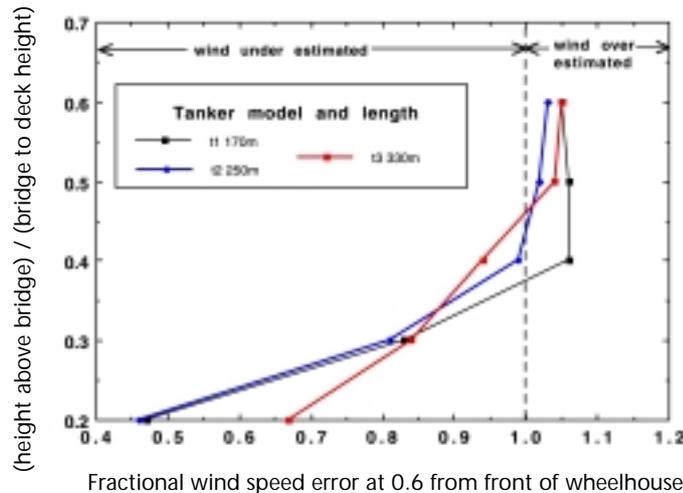
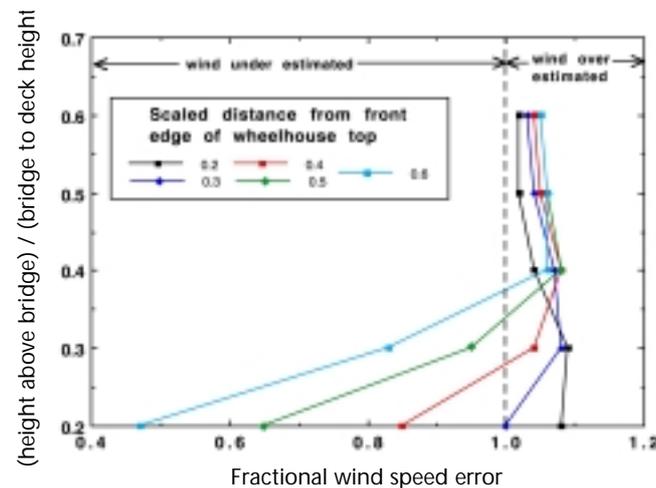


Figure 11—As Figure 10, but for tanker (1) at different scaled distances from the front of the wheelhouse (x/D).



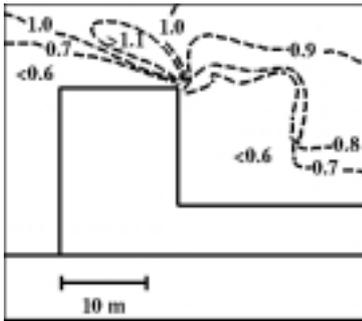


Figure 12—Detailed view showing airflow distortion over the stern section of a typical tanker as determined by CFD modelling (after Yelland *et al.*, 1998a). The wind is blowing from right to left.

6 COMPARISON WITH SATELLITE DATA

6.1 IMPORTANCE OF CORRECT ERROR TREATMENT

However, Gulev (1999) showed that the use of the Lindau scale degrades the agreement between VOS winds and a data set of Russian research ship winds. The reason for this is that the Lindau and UWM scales are calculated to bring VOS visual and VOS anemometer winds into agreement. The anemometer winds from the Russian research ships used by Gulev were similar in magnitude to unadjusted VOS visual winds and significantly higher in magnitude compared to VOS anemometer reports. Thus, converting the VOS winds to the Lindau scale decreased the stronger wind values, thereby improving the comparison with the VOS anemometer data as expected, but degrading the agreement with the research ship data.

If Gulev's research vessel data are correct, the implication is that VOS winds are on average underestimated. However, Isemer (1994) noted that when weather station C began to be manned by ships which provided Gulev's data set, there appeared to be an increase in the measured winds. This does not prove that the Russian winds are necessarily too high; we repeat that, in our view, there is not an absolute standard for wind measurement.

Finally, in discussing visual winds, we would stress that it is important that the ships' officers do not change from the present WMO code 1100 scale. Any adjustment should be left to those preparing climatological data sets.

Kent *et al.* (1998; henceforth K98) compared VOS winds with those measured by the scatterometer on ERS-1. The VOS winds had been quality controlled and corrected for anemometer height, or adjusted to the Lindau scale, as appropriate. The study demonstrated very clearly the importance of properly accounting for the observation errors in each of the data sets which are compared. Thus, Figure 13 shows the results of different comparison strategies. If the (satellite-ship) differences were averaged as a function of the ship winds it appeared that, compared to the ships, the scatterometer was biased high at low wind speeds and high at high wind speeds. Similar plots showing similar apparent bias can often be found in the literature (e.g. Liu, 1984; Offiler, 1994; Boutin and Etcheto, 1996).

However, if the same differences were binned using the satellite data as the independent variable then the conclusions appeared different. The satellite data were apparently low at lower winds but in agreement with the ship data over much of the wind speed range. K98 demonstrated that this was due to the different variance for the two data sets; a problem that has been recently discussed by Tolman (1998; see also Kent & Taylor, 1999).

To simulate the effect, K98 used a single wind speed data set obtained from a moored buoy. The simulated data sets were calculated by adding to the buoy wind data random errors, normally distributed with an rms of 2.0 m/s to represent the ship winds and 0.5 m/s to represent the scatterometer winds. These rms values had previously been obtained by semivariogram analysis. The two simulated data sets were then analysed in a similar manner to the actual data sets. Apart from a small offset when using the simulated satellite data as the independent variable, the results of the simulation (also shown in Figure 13) showed

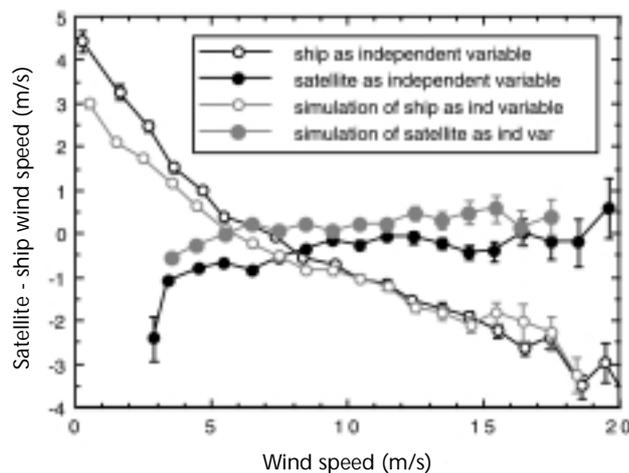


Figure 13—Comparison of the ERS-1 scatterometer with ship wind speeds showing different results depending on which data set is used as the independent variable. Also shown (lighter lines) are the results of the simulated comparison described in the text (from Kent *et al.*, 1998).

the same behaviour as the real data. K98 proceeded to demonstrate that the same effect could result in a stability dependent bias being erroneously ascribed to the scatterometer data.

Using a regression method which correctly allows for the different error characteristics for each regression variable (e.g. Graybill, 1961), K98 showed that the ship winds were slightly higher than those from the scatterometer:

$$U_{10m}(ship) = 1.025U_{10m}(scat) + 0.255 \quad (1)$$

A very different result would be obtained by regressing the satellite winds on ship winds without considering the errors. The ship values are around 0.5 m/s higher at 10 m/s and 1 m/s higher at 30 m/s. This could be due to the buoy measured winds, used to develop the scatterometer algorithm, underestimating the wind speed; it may be due to airflow disturbance biasing the ship winds; we do not know if either is correct.

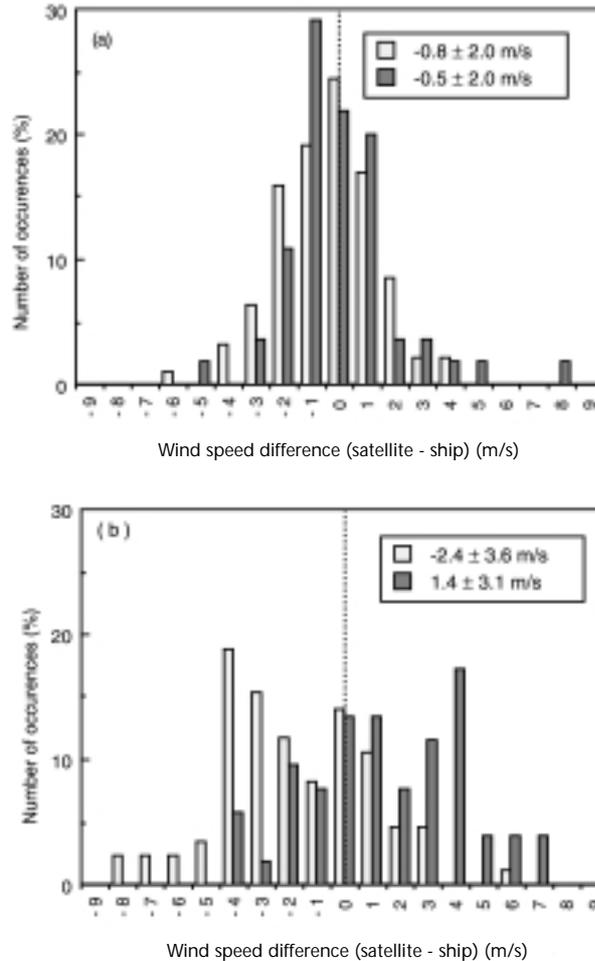
6.2 QUALITY CONTROL OF VOS DATA K98 also showed that the scatterometer data could be used to identify ships whose wind reports showed large biases or error variability. Thus, Figure 14(a) shows the distribution of satellite-ship comparisons for two ships reporting reliable winds. The rms scatter is typical of the overall data set from the ships, and the mean bias is similar to that predicted by (1). In contrast, Figure 14(b) shows the distribution for two ships whose wind estimates were less reliable. Although both histograms showed a number of observations close to the scatterometer values, secondary peaks occurred at about 4 m/s difference. Since these ships were reporting visual winds, correction to true wind should not have been a problem. Rather, it suggests that a Beaufort force two intervals away from the true value was sometimes chosen.

7. FUTURE DEVELOPMENTS
7.1 AUTOMATIC CODING The use of automatic coding of ships' weather messages using a personal computer system and form filling techniques is becoming more common. A popular system is TurboWin developed at KNMI in the Netherlands. Such a system should ensure that position is correctly coded (and compatible with the last reported position) and remove a major source of error by automatically computing true wind.

7.2 AUTOMATIC DATA ACQUISITION Computer-based systems can also be used to automate data acquisition. For example, the Improved Meteorological System (IMET) has been installed on a number of US Research Vessels and is now being placed on US VOS (Weller and Taylor, 1998). IMET uses sensors chosen (based on laboratory and field studies) for accuracy, reliability, low power consumption and their ability to stay in calibration during unattended operation. The sensors are combined with front end digital electronics to make a module which is digitally addressable (RS-232 or RS-485), stores its calibration information and provides either raw data or data in meteorological units. The present set of IMET modules includes wind velocity and most other meteorological variables.

7.3 AIR-SEA FLUX DETERMINATION Using European Union funding under the MAST programme, the AutoFlux Group (1997) is developing an autonomous system for monitoring air-sea fluxes using the inertial dissipation method and ship-mounted instrumentation. It aims to develop and test a prototype system, called AutoFlux, which will measure surface stress, sensible and latent heat flux, and also carbon dioxide flux. The system is aimed primarily at unattended use on VOSs and on unmanned buoys. The fluxes are derived from the turbulence spectra using the 'inertial dissipation' method. This technique minimizes the effects of flow distortion and platform motion. The system software will manage data conversion, storage and transmission, including the necessary navigational information. The present project should be regarded as 'proof of concept', but, if successful, AutoFlux-type systems might be installed on selected VOSs in a few years time. Transmitting flux data over the GTS will require a new code format.

Figure 14—Comparison of satellite-ship wind speed differences for individual ships: (a) two ‘good’ ships (one anemometer, one visual); (b) two ‘bad’ ships (both visual).



7.4 SATELLITE TRANSMISSION

The recent introduction of relatively inexpensive global data transmission systems via satellites suggests the possibility of transmitting a more comprehensive weather observation message that includes information such as the method of SST measurement, the relative wind observation, etc. The full message could be archived for use in climate studies with the standard GTS message being extracted and transmitted by the land station for weather prediction purposes.

7.5 AN IMPROVED SUBSET OF THE VOSS

While these various improvements to VOS observations are highly desirable, systems such as IMET or AutoFlux are much more expensive and require more shore-side support compared to the instrumentation typically provided to the VOSSs. It will not be practicable to supply such instrumentation to a substantial fraction of the VOS fleet. However, the establishment of an improved subset of VOSSs would provide a verification standard which would allow the biases in the standard VOS data to be quantified. As a result, all VOS observations would be improved in value. A subset of about 100 to 300 selected VOSSs could provide a significant contribution (e.g. Taylor, 1984).

8. SUMMARY

We have emphasised the lack of an absolute calibration standard for marine wind measurements. Wind data obtained from ships are affected by the air flow distortion around the ship. This is true for all practicable anemometer sites. Positions can be found where for some relative wind directions the disturbed wind speed matches the free stream wind speed, but this is unlikely to hold for all wind directions. We have demonstrated success in correcting these errors using CFD or wind tunnel data but there are very few data sets for which this has been done. Data from buoys are suspect at higher wind speeds because of the sheltering effect of waves. The error in buoy winds may have also caused bias in scatterometer data.

The fraction of anemometer-based winds has increased with time, particularly in the Pacific. The average height of anemometers is higher in the Pacific compared to the Atlantic. Correcting for anemometer height (on a ship-by-ship basis) and adjusting winds to the Lindau scale reduces the rms scatter in the wind speed data set by about 15 per cent.

A major source of error in anemometer-derived winds is the calculation of true wind speed and direction from the measured wind speed; an automatic method of calculation is required. CFD studies on the airflow over simple generic tanker models show that it is important that the anemometer be mounted above the plume of accelerated air which occurs over the wheelhouse top.

In comparing ship and scatterometer data we have emphasised the importance of taking the different error characteristics into account. When this is done, it appears that the ships are biased high compared to the scatterometer by around 4 per cent; we do not know which is the most correct. The scatterometer data can be used to identify ships whose wind reports are less reliable.

In the future it is expected that VOS meteorological reports will be increasingly automated, thereby removing errors in calculating true winds or in coding the ship's position. An improved subset of the VOSs would be valuable as a standard for improving the VOS data set as a whole.

Finally we make the following recommendations:

- For ships reporting anemometer winds, the ship's officers should be provided with an automated method of calculating the true wind.
- Anemometer read-outs should automatically average the winds.
- Hand-held wind sensors should not be used.
- The position of the anemometer must be documented. This must include height above sea level and also measurements indicating the location of the anemometer in relation to the overall shape of the ship. In future this will allow average CFD corrections to be calculated for typical VOSs.
- Visual wind observations should continue to be based on the WMO code 1100 scale. For scientific analysis the Lindau scale is to be preferred over other versions (such as that recommended by CMM-IV).
- That a high quality subset of the VOSs be developed and used to verify the data from the VOS fleet as a whole.

ACKNOWLEDGEMENTS

Sections of this paper originally appeared in Taylor & Kent (1999). The satellite scatterometer work was a contribution to the Joint Grant Scheme project "Coastal and Open Ocean Wind Stress". The CFD studies would not have achieved their present prominence without partial funding from the Meteorological Service of Canada, which also supported the SWS-2 campaign.

REFERENCES

- AutoFlux, 1997: AutoFlux: an autonomous system for monitoring air-sea fluxes using the inertial dissipation method and ship-mounted instrumentation (MAST Project) <http://www.soc.soton.ac.uk/JRD/MET/AUTOFLUX>
- Boutin, J. and J. Etcheto, 1996: Consistency of Geosat, SSM/I, and ERS-1 global surface wind speeds – comparison with in situ data. *J. Atmos. & Oceanic Tech.*, 13, 183–197.
- da Silva, A.M., C.C. Young and S. Levitus, 1995: Towards a revised Beaufort equivalent scale, in Diaz & Isemer (1995), 270–286.
- Diaz, H.F. and H.-J. Isemer (ed.), 1995: *Proc. Internat. COADS Winds Workshop* (Kiel, Germany 31 May–2 June 1994). NOAA/ERL Boulder, Colorado & Institut für Meereskunde, Kiel., 301 pp.
- Dobson, F.W., R.J. Anderson, P.K. Taylor and M.J. Yelland, 1999: Storm Wind Study II: Open ocean wind and sea state measurements. *Air Sea Interface Symposium* (Sydney, Australia, 11-15 January 1999).
- Graybill, F., 1961: *An Introduction to Linear Statistical Models, Volume 1*, New York: McGraw-Hill, 463 pp.
- Gulev, S.K., 1999: Comparison of COADS Release 1a winds with instrumental measurements in the NorthWest Atlantic. *J. Atmos. & Oceanic Tech.*, 16(1), 133–145.

- Isemer, H.-J., 1994: The homogeneity of surface wind speed records from ocean weather stations, Inst. für Meereskunde, Kiel, Germany, 43 pp. + Figs.
- Isemer, H.-J., 1995: Trends in marine surface wind speed: ocean weather stations versus voluntary observing ships, *in* Diaz & Isemer (1995), 68–84.
- Kent, E.C. and P.K. Taylor, 1999: Accounting for random errors in linear regression: A practical guide. *Q. J. Roy. Met. Soc.*, (accepted).
- Kent, E.C. and P.K. Taylor, 1997: Choice of a Beaufort equivalent scale. *J. Atmos. & Oceanic Tech.*, 14(2), 228–242.
- Kent, E.C., P.K. Taylor and P. Challenor, 1998: A comparison of ship and scatterometer-derived wind speed data in open ocean and coastal areas. *Int. J. Remote Sensing*, 19(17), 3361–3381.
- Kent, E.C., P. Challenor and P. Taylor, 1999a: A statistical determination of the random observational errors present in VOS meteorological reports. *J. Atmos. & Oceanic Tech.* (accepted), 16(7), 905–914.
- Kent, E.C., P.K. Taylor and S.A. Josey, 1999b: Improving global flux climatology: The role of metadata. *CLIMAR99, WMO Workshop on Advances in Marine Climatology* (Vancouver, 8–15 Sept. 1999).
- Kent, E.C., P.K. Taylor, B.S. Truscott and J.A. Hopkins, 1993: The accuracy of Voluntary Observing Ships' Meteorological Observations. *J. Atmos. & Oceanic Tech.*, 10(4), 591–608.
- Kent, E.C., B.S. Truscott, J.S. Hopkins and P.K. Taylor, 1991: *The Accuracy of ship's Meteorological Observation – results of the VSOP-NA*. Marine Meteorology and Related Oceanographic Activities 26, WMO, Geneva, 86 pp.
- Large, W.G., J. Morzel and G.B. Crawford, 1995: Accounting for surface wave distortion of the marine wind profile in low-level ocean storms wind measurements. *J. Phys. Oceanogr.*, 25, 2959–2971.
- Lindau, R., 1995: A new Beaufort equivalent scale, *in* Diaz & Isemer (1995), 232–252.
- Liu, W.T., 1984: The effects of the variation in sea surface temperature and atmospheric stability in the estimation of average wind speed by SeaSat-SASS. *J. Phys. Oceanogr.*, 14, 392–401.
- Moat, B.I., M.J. Yelland and P.K. Taylor, 1998: The impact of airflow distortion on in situ meteorological wind speed measurements. *Proc. 23rd European Geophysical Society General Assembly* (Nice, 20–24th April, 1998), Annales Geophysicae, C646.
- Offiler, D., 1994: The calibration of ERS-1 satellite scatterometer winds. *J. Atmos. & Oceanic Tech.*, 11, 1002–1017.
- Smith, S.R., M.A. Bourassa and R.J. Sharp, 1999: Establishing more truth in true winds. *J. Atmos. & Oceanic Tech.*, 16(7), 939–952.
- Taylor, P.K., 1984: The determination of surface fluxes of heat and water by satellite microwave radiometry and in situ measurements. *NATO ASI, Series C: Math. and Phys. Sci.: Large-scale Oceanographic Experiments and Satellites*, C. Gautier and M. Fieux, Ed., Vol.128, D.Reidel, 223–246.
- Taylor, P.K. and E.C. Kent, 1999: The Accuracy of Meteorological Observations from Voluntary Observing Ships – Present Status and Future Requirements. *First session of the CMM Subgroup on Voluntary Observing Ships* (Athens, 8–12 March, 1999) (to be published by WMO).
- Taylor, P.K., E.C. Kent, M.J. Yelland and B.I. Moat, 1995: The accuracy of wind observations from ships, *in* Diaz & Isemer (1995), 132–155.
- Taylor, P.K. and R.A. Weller, 1991: *Meteorological Measurements from WOCE Research Ships*. WOCE Operations Manual, Vol. 3: Section 3: Underway Measurements WHPO 91-1, WOCE Report 68/91, WOCE, Woods Hole, Mass. USA, 26 pp.
- Taylor, P.K., M.J. Yelland, F.W. Dobson and R.J. Anderson, 1999: Storm Wind Study II: Wind stress estimates from buoy and ship. *Air Sea Interface Symposium* (Sydney, Australia, 11–15 January 1999).
- Thiebaux, M.L., 1990: Wind tunnel experiments to determine correction functions for shipborne anemometers. *Canadian Contractor Rep. Hydrog. & Ocean Sci.* 36, BIO, Dartmouth, Nova Scotia, 57 pp.

- Tolman, H.L., 1998: Effects of observational errors in linear regression and bin-average analysis. *Q. J. Roy. Met. Soc.*, 124(547), 897–917.
- WMO, 1970: *The Beaufort Scale of Wind Force (technical and operational aspects)*, World Meteorological Organization, Geneva, 22 pp.
- Weller, R.A. and P.K. Taylor, 1998: A strategy for marine meteorological and air-sea flux observation. *Preprint volume: 3rd Symp. on Integrated Observing Systems* (Dallas, Texas, Jan., 1999), AMS, Boston.
- Woodruff, S.D., S.J. Lubker, K. Wolter, S.J. Worley and J.D. Elms, 1993: Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. *Earth System Monitor*, 4(1), 4–8
- Yelland, M.J., B.I. Moat and P.K. Taylor, 1998a: Air-flow distortion around voluntary observing ships. May 1998 progress report. *Contract rep. for AES and BIO*.
- Yelland, M.J., B.I. Moat, P.K. Taylor, R.W. Pascal, J. Hutchings and V.C. Cornell, 1998b: Measurements of the open ocean drag coefficient corrected for air flow disturbance by the ship. *J. Phys. Oceanogr.*, 28(7), 1511–1526.
- Zeng, L.X. and R.A. Brown, 1998: Scatterometer observations at high wind speeds. *J. Appl. Meteorol.*, 37(11), 1412–1420.

REPORT ON BEAUFORT EQUIVALENT SCALES

Ralf Lindau, Institut für Meereskunde, Kiel, Germany

ABSTRACT The Beaufort scale derived by Lindau (1995) is recommended for converting visual marine wind estimates especially for climate study purposes, where a consistent conversion of entire data sets is essential. Since the shortcomings of earlier Beaufort scales can be mainly explained by the statistical method of derivation, a major part of this report is dedicated to basic statistical considerations.

1. INTRODUCTION

For over a century marine meteorologists have been searching for the definitive conversion of Beaufort estimates into metric wind speed. In principle, the derivation procedure is rather clear. Using a suitable technique, Beaufort estimates have to be compared to reliable wind measurements in their spatial and temporal vicinity. Finding a data set of high quality marine wind measurements is, at first glance, the most crucial prerequisite for an equivalent scale. Actually, the quality of the derived scale is indeed limited by the reliability of the calibration data set. Kaufeld (1981) used wind measurements from Ocean Weather Stations (OWSs) in the North Atlantic. For more than a decade, three hourly (at some stations even hourly) observations were made on a continuous basis by professional crews. In addition, the stations were situated in the open ocean. Therefore, coastal influences on the Beaufort estimates which are intended to be calibrated can be discounted. Another advantage is that in general the ships stayed at fixed positions so that measurement errors caused by the ship's speed did not occur. The large number of observations, together with their relatively high accuracy, qualify the wind measurements from OWSs as an excellent calibration data set.

After the principal decision of which data set should be used as reference, the data analysis follows. How to perform this final technical step has been debated for over one hundred years. This report reviews the various scales that have evolved and presents a statistical procedure for the correct derivation of a Beaufort equivalent scale. In the conclusion a definitive scale is recommended. Since questions about the appropriate statistical analysis are the most controversial part of the discussion, a detailed consideration of regression techniques is also necessary.

2. REGRESSIONS

As a basis for analysis, we first consider pure linear regressions. If data pairs from two samples X and Y are available, the correlation coefficient is defined as:

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (1)$$

which is equal to the covariance divided by the standard deviation of both samples. The regression of Y on X is defined as:

$$\hat{y} = r \frac{\sigma_y}{\sigma_x} (x - \bar{x}) + \bar{y} \quad (2)$$

where \bar{x} and \bar{y} denote the means of the two calibration data sets, with σ_x and σ_y being their respective standard deviations. The above regression line enables us to predict individual values for a given value of x; and predicting a wind speed

value for a given Beaufort estimate in this way is just what we might expect from an equivalent scale.

In order to gain a better insight into the problem, it is helpful to consider the historically used regression method as well (Figure 2). The regression line (2) is easy to calculate with modern computers, but in the past it was an arduous task. Therefore, the commonly applied technique was to sort the observation pairs into classes of constant Beaufort force and to compute the mean wind speed for each of these classes. Then, the regression line of the wind speed on the Beaufort force could be obtained by connecting these class averages. For the linear case, such a procedure is equivalent to the modern method. Actually, it can be even more powerful since non-linear relationships are detectable as well.

As a very simple example, let us consider two identical thermometers T_1 and T_2 , both of which provide time series of the temperature at two neighbouring sites. Because of their same construction and spatial proximity we suppose that there is no bias between them, and expect the same variance for both time series. Let us further assume a correlation coefficient of 0.6 between both instruments which is caused by the small but noticeable distance between them.

Since we defined a priori the universal relationship between both thermometers, a kind of equivalent scale is easy to determine. If we predict the measurements of T_2 from T_1 , it is obvious that:

$$T_2 = T_1 \tag{3}$$

would give the optimal estimate. But surprisingly, this holds true only if the characteristics of entire samples are considered. For the prediction of individual values, Equation (2) gives the best estimate. To make the example as vivid as possible, if we assume a mean temperature of 10°C, the one-sided regression of T_2 on T_1 tells us that $T_2 = 16^\circ\text{C}$ (Figure 1) would be the best prediction for the second thermometer, if the first shows a temperature $T_1 = 20^\circ\text{C}$ (and $T_2 = 4^\circ\text{C}$, when $T_1 = 0^\circ\text{C}$).

At this stage, two questions arise: Since Equation (2) seems to be in clear contradiction to our common sense, how can it be the optimal prediction for individual values? And, if we were convinced that this really is the case, why is Equation (2) not the appropriate basis for an equivalent scale?

2.1
PREDICTION OF INDIVIDUAL
VALUES - ONE-SIDED
REGRESSIONS

Let us address the first question. In our example, individual values can be regarded as comprising two components. The first is the mean temperature of the spatially extended surrounding of both thermometers, since the values can be regarded as individual realizations representative of the entire area. For this reason a prediction of one thermometer from the other is actually possible. The second is a modification by a stochastic spatial temperature gradient leading to

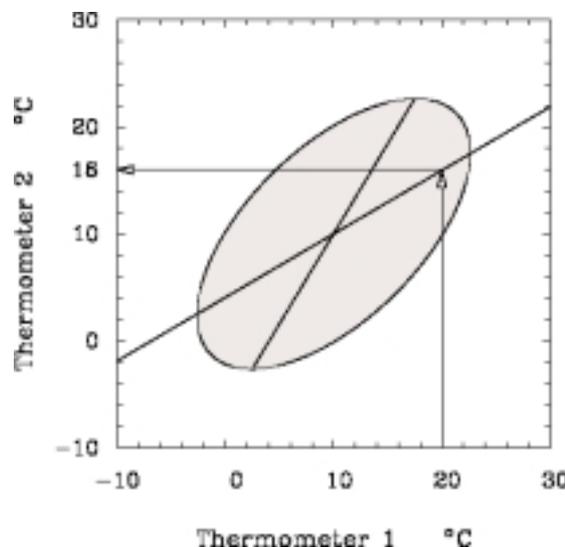
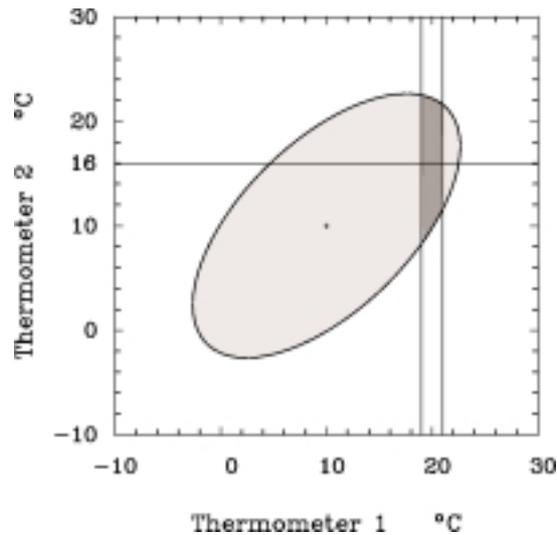


Figure 1—Thermometer 1 shows 20°C, the best estimate for thermometer 2 is 16°C, although both instruments are neighbouring and identical.

Figure 2—The historical method to calculate regressions. First, choose a value for the predictor, e.g. 20°C. Secondly, sort out all temperature pairs with $T_1 = 20^\circ\text{C}$ (dark grey area). Thirdly, calculate the mean temperature at T_2 for these cases. Finally, repeat the procedure for several predictor values and connect the results graphically.



slightly different values at both thermometers. Because of this variability a perfect prediction is not completely possible.

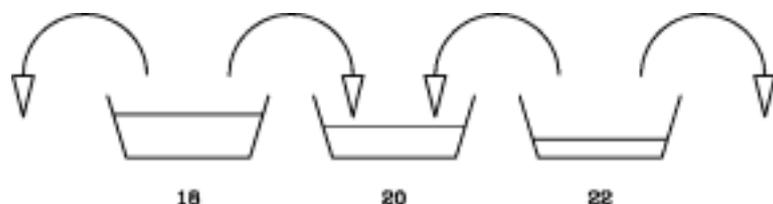
According to the historical method describe above (Figure 2), we can obtain the regression point-by-point using the following steps. First choose a fixed value for the predictor, e.g. $T_1 = 20^\circ\text{C}$, sort out all temperature pairs ($T_1; T_2$) with $T_1 = 20^\circ\text{C}$, and then calculate the mean temperature at T_2 for these cases. As we know already from Equation (2), the result will be 16°C .

Considering now the members of the 20°C -class of T_1 (Figure 3), we have to be aware that these values are already modified by a random deviation from their respective spatial mean. For example, it is possible that a modified value of 20°C results from a momentary spatial mean of 18°C combined with a local anomaly of $+2^\circ\text{C}$. On the other hand, 20°C may occur when the spatial mean for that time is 22°C together with an anomaly of -2°C . Since we assume that the local deviations are random, such positive and negative anomalies of the same amount have indeed the same probability. The deviations do not have a different probability, but the situations themselves do. Extreme situations are of course less frequent than situations closer to the overall mean. Applied to our example: situations with spatial means of 18°C are more frequent than those with 22°C , when the overall average is 10°C . Thus, considering the origins of the measurements of $T_1 = 20^\circ\text{C}$, colder spatial means are more likely than warmer ones, so that 16°C is the average of these situations.

The measurement at T_2 is just another realization of the instantaneous temperature in the considered area. But we are taking an average over several of these values, so that T_2 reflects finally the mean temperature of the selected sample, which is 16°C , as we have seen above, and not 20°C . Thus, for extreme values there is an increased probability that they are based solely on local events and cannot be found at a neighbouring station. It is therefore wise to predict a value closer to the overall mean.

It is obvious that the example can be generalized. If we substitute the expression 'spatial mean' by 'true value' and the expression 'local deviations' by 'observation errors', it will become clear that it does not matter whether real spatial differences or random observation errors are responsible for the reduced correlation coefficient.

Figure 3—Each measurement can be regarded as comprising two components: the spatially mean value which is representative of a broader area, plus a random deviation for the particular site. The buckets include situations with a mean temperature of 18°C , 20°C and 22°C , respectively. When the mean temperature is 10°C , 18°C is more frequent than 22°C . The actual temperature is a random deviation from the mean, that means in a figurative sense, splashing randomly in all directions. After this splashing procedure we examine the 20°C bucket, asking: where do these measurements come from? The probability of leaving a bucket is the same for all buckets and for both directions, but the 18°C bucket is fuller so that more 'splashes' come from lower temperatures. This means that if a thermometer shows 20°C , it is more likely that the surrounding is colder than 20°C .



Nevertheless, regression results similar to those discussed above can lead to the erroneous conclusion that T_2 underestimates the temperature for warm situations, and overestimates it for cold situations. Obviously this is not true, since a selection according to T_2 instead of T_1 would of course lead to the reversed result: considering only observation pairs with $T_2 = 20^\circ\text{C}$, it will now be T_1 which shows a mean temperature of only 16°C .

We have seen so far that Equation (2) is indeed the best prediction for a given individual value, so we can turn to the second question of why it should not be used as an equivalent scale. I will expound in the following that such one-sided regressions do not meet the requirements of an equivalent scale, but that an improved version of Equation (3) is better suited. Both equations have their own advantages, and we have to face the possibility that an optimum equation for all possible applications is not attainable. It is necessary to decide which of the scale characteristics are essential and which have a lower priority.

2.2
REQUIREMENTS FOR
EQUIVALENT SCALES - THE
ORTHOGONAL REGRESSION

When we consider converting by Equation (2) not individual values but an entire data set, the disadvantages of the one-sided regression are revealed. Such a theoretical data set, generated by the application of Equation (2), will contain only that part of the variance which is explained by the predictor.

The variance of the derived data set is:

$$\text{var}(\hat{y}) = \frac{1}{N-1} \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 \tag{4}$$

From Equation (2), it follows that:

$$\text{var}(\hat{y}) = \frac{1}{N-1} \sum_{i=1}^N r^2 \frac{\sigma_y^2}{\sigma_x^2} (x_i - \bar{x})^2 \tag{5}$$

which is equivalent to:

$$\text{var}(\hat{y}) = r^2 \sigma_y^2 \tag{6}$$

The diminution of variance by the factor r^2 has serious consequences. It leads to a substantial underestimation of the annual cycle since the correlation coefficient between wind speed and Beaufort force is significantly less than 1 (Figure 4). Therefore, monthly means would be systematically underestimated for one half of the year (with anomalously strong winds) and overestimated for the other half. Such performance is of course unacceptable for an equivalent scale.

To illustrate this point we calculated the two one-sided regressions and the orthogonal regression between the wind speed measurements at OWS K and the

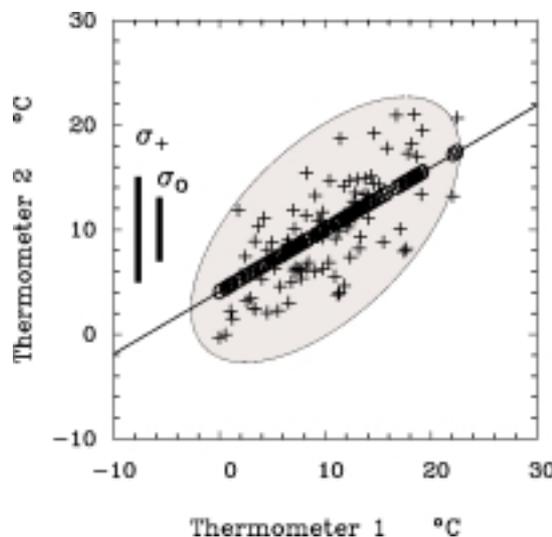


Figure 4—Schematic figure to illustrate the reduced variance of the predicted parameter. Crosses depict real values, circles are the prediction using the one-sided regression of thermometer 2 on thermometer 1. Owing to the prediction, all crosses are shifted vertically, lying finally on the regression line. It is obvious that the variance is decreased by this procedure.

Beaufort estimates of nearby passing merchant ships. The question is this: is it possible to predict the monthly wind speed at OWS K by the Beaufort estimates of the merchant ships by using the calculated regression lines for conversion? Figure 5 shows that a one-sided regression of wind speed on Beaufort seriously underestimates the annual cycle, while the orthogonal regression is in better agreement with the actual measurements at OWS K.

Another consequence is that one-sided regressions are not necessarily valid in other climates. Applying an equivalent scale in climatic zones where it has not been derived has attendant risks. But when using one-sided regressions, it is certain that even the longtime mean is not reproduced. If \bar{x} and $\bar{\xi}$ denote the mean derived and the mean applied Beaufort force, it follows directly from Equation (2) that the change in the obtained mean wind speed $\bar{\eta} - \bar{y}$ will be underestimated by the factor r .

$$\frac{\bar{\eta} - \bar{y}}{\sigma_y} = r \frac{\bar{\xi} - \bar{x}}{\sigma_x} \quad (7)$$

Considering again the thermometer example, another disadvantage of one-sided regressions becomes obvious. Let us assume that one calibration attempt is carried out in winter, with a mean temperature of 0°C, and a second experiment is performed in summer with 20°C as average. Leaving other circumstances unchanged, the winter regression will provide $T_2 = 6^\circ\text{C}$ as the best estimate for a given value of $T_1 = 10^\circ\text{C}$, because the correlation is assumed to be 0.6. However, the summer regression will give for the same value ($T_1 = 10^\circ\text{C}$) a best estimate of $T_2 = 14^\circ\text{C}$. For individual predictions this is reasonable; for the wintertime, a temperature of 10°C is a warm extreme and will have an opposite effect on the probability representative of its surrounding than is the case in summertime, when 10°C is a cold extreme. Nevertheless, it is hardly acceptable that the derivation of equivalent scales leads to different results depending solely on the average temperature regime. However, different wind climates might indeed justify different equivalent scales due to changed physical conditions. But in our example absolutely identical instruments were assumed; obtaining two different scales was absolutely unavoidable, for purely statistical reasons. Differences in physical conditions would only modify this outcome.

To assess the practical consequences, wind measurements at OWS K and Beaufort estimates of nearby passing merchant ships are investigated. The one-sided regression of wind speed on Beaufort and the reversed regression are given in Figure 7. For the conversion from Beaufort force into metric wind speed, the former ones are (if one-sided regressions are used at all) appropriate. However, in summer, the equivalent value for Beaufort 4, for instance, would be 14.5 kt, considerably lower than in winter with 18.6 kt. Figure 6 shows the orthogonal

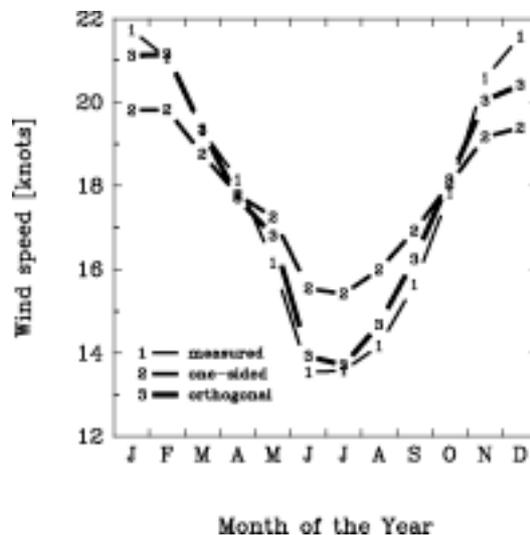


Figure 5—The mean annual cycle of the wind speed as measured by OWS K (1). Using the one-sided regression of wind speed on Beaufort as conversion (2), the annual cycle is considerably underestimated. The orthogonal regression (3) fits much better.

Figure 6—The orthogonal regressions between wind measurements at OWS K and Beaufort estimates of merchant ships in the vicinity, separately calculated for each month of the year.

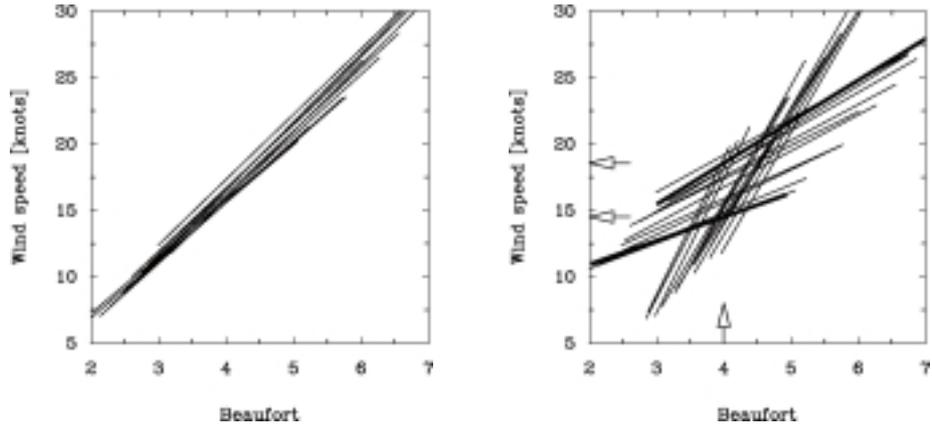


Figure 7—As Figure 6, but for the two one-sided regressions. As an equivalent value for Beaufort 4, the July-regression (lower thick line) would give 14.5 kt, but the January-regression (upper thick line) 18.6 kt.

regressions separately for each month of the year. The twelve regression lines coincide rather well, confirming that the orthogonal regressions closely reflect the common relationship between wind speed and Beaufort force.

In summary, we can state that although one-sided regressions are well suited to predicting individual values, such a conversion cannot be recommended for entire data sets. By using one-sided regressions as an equivalent scale, the statistical characteristics of the obtained data set will be changed substantially. The total variance will be underestimated, which causes, for example, a weakened annual cycle of the converted wind speed. For these reasons, one-sided regressions are not applicable in other wind climates, where even the obtained overall mean would be incorrect. If different scales are derived for different climates (say, twelve scales, one for each month of the year) the scales will not coincide even if Beaufort force and wind speed are actually connected by a commonly valid relationship (for which we are searching).

Hence, our first impulse to use Equation (3) as a conversion scale is reasonable, because we are not focussed on the optimal prediction of a single measurement but on the conservation of the overall statistical characteristics. Equation (3) is of course only valid for the above very simple and specific case, where equal variances and no bias were assumed. It is condensed from the following more general expression which is known as the orthogonal regression:

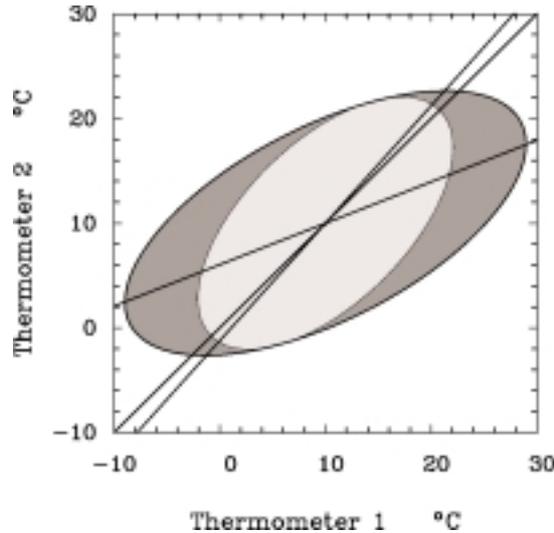
$$\hat{y} = \frac{\sigma_y}{\sigma_x} (x - \bar{x}) + \bar{y} \tag{8}$$

It is easy to show that this regression actually conserves the statistical properties discussed above. The variance of the converted data set remains unchanged, and application in other wind climates is principally possible. Calibration data sets with different total means will lead to the same results, unless there are significant contradictory physical reasons. Hence, the orthogonal regression is well suited to serve as an equivalent scale.

Nevertheless, a careful assessment of the calibration data sets used, i.e. wind measurements and Beaufort estimates, is necessary. Both the temporal resolution and the relative error variance play an important role (Figure 8). Consider again the example of two thermometers without any systematic difference between their measurements. An equivalent scale giving the correct universal conversion should obviously have a slope of 1. This remains true even if we suppose that one thermometer measures more accurately than the other. But the unequal error variances cause different total variances for both time series so that, according to Equation (8), the slope of the orthogonal regression will not be equal to 1. A comparable effect occurs when the standard deviations of both data sets differ due to the unequal resolutions of the time series considered. If one of the data sets contains temporally averaged values, its variance will be reduced compared to the other data set which consists of instantaneous measurements. As result, we obtain

Figure 8—Considering the thermometer example with the original data points lying in the light area, the 45° line is the best conversion for entire data sets.

This remains true, even if thermometer 1 is less accurate, which would cause an elongation of the scatter ellipse (dark grey area). Computing from this data the two one-sided regressions, it becomes obvious that the reversed regression line, where thermometer 2 is regarded as independent, is much better suited for a conversion than the regular one. The same effect would occur if thermometer 1 is not less accurate, but if the variance is increased by a higher temporal resolution of the measurements.



again a slope which differs from 1. In order to avoid such errors, we have to ensure that the data sets used for the calibration are of the same temporal and spatial resolution so that they actually contain a comparable amount of natural variability. A second requirement is that their relative error variance has to be equal.

Hence, the orthogonal regression is the most suitable statistical way to derive an equivalent scale. But beforehand, the possible effects of different resolutions and different error variances of both calibration data sets have to be eliminated. Now that the principal question of how to proceed has been clarified, let us review the numerous approaches of the last hundred years. This will show that progress has not always taken a straight course.

3. HISTORICAL SCALES

In the 19th century the first attempts were made to assign metric wind speeds to the 13 wind strength classes of Admiral Beaufort. The principal procedure for this purpose has not changed since then. The shipborne estimates are compared with reliable wind measurements in the temporal and spatial vicinity. The statistical analysis of these observation pairs then leads to an equivalent scale.

At the end of the 19th century, knowledge about regression techniques was just evolving, but access to fast data processing was not available. As mentioned in the previous chapter, the usual technique was to sort the data pairs into classes of constant Beaufort force and to compute the mean wind speed for each of these classes. Then the one-sided regression line of the wind speed on the Beaufort force can be obtained by connecting these class averages. Reversing the sorting and the averaging parameters gives the other one-sided regression of Beaufort force on wind speed. The second regression is suited to compute an individual Beaufort force from a given wind speed.

3.1. KÖPPEN AND SIMPSON

In 1888 a discussion began; which one-sided regression should be used as an equivalent scale? Based on a suggestion made by Köppen, Waldo (1888) proposed to take the regression of the Beaufort values on the measured wind velocity, i.e. to calculate the mean Beaufort estimate for a given wind speed class. This was in contrast to the common opinion at this time. Even nowadays Köppen's excellent argument is not always accepted since it is normal to predict wind speed from the Beaufort scale.

After 1888 Köppen explained his point of view in several publications (Köppen, 1916a, 1916b, 1926). In an article (Köppen, 1916a), he emphasized that both one-sided regressions are not optimal. However, treating the measurements of windspeed as the independent parameter would give much better results because they were averages over one hour, whereas the estimates were instantaneous values. Sorting the data pairs into classes of wind speed and averaging over the Beaufort estimates would further reduce the variance.

In London, Simpson (1906) published another Beaufort equivalent scale. Finally accepted by WMO in 1946 as Code 1100, this scale is commonly in use even today. It is remarkable that Simpson proceeded in the manner suggested by Köppen. He averaged estimates for fixed wind speed classes, thus obtaining the one-sided regression of Beaufort against wind speed. But in contrast to Köppen, Simpson considered the higher error variance of the estimates as the main reason for such treatment of data.

However, both authors were aware that the variance of the Beaufort estimates is increased, whether due to the higher temporal resolution or to the lower accuracy of the estimates; the regression of Beaufort against wind speed is preferable.

In 1916, Köppen conceded the plausibility of Simpson's point of view, namely that errors caused the higher variance of Beaufort estimates. Köppen was convinced by the fact that Curtis (1897) found no significant differences in his results when he calibrated the estimates against wind speed averages over 10 minutes instead of hourly means.

In the beginning of the last century the one-sided regression of Beaufort on wind speed was the commonly accepted equivalent scale. In an overview article, Köppen (1926) pointed out that he was well aware of the weaknesses inherent in one-sided regressions, and that improvements were still necessary. In those days, however, the available data sets were small, and experience with regression techniques was not sufficiently established to solve the problem fully.

3.2 THE 'METEOR' CRUISE

From 1925 to 1927, during the German Atlantic Expedition, the research vessel 'Meteor' cruised into the South Atlantic. During this voyage the diverse problems of wind observations at sea were investigated. In this context the actual Beaufort force was estimated hourly by eight different observers, while the wind speed was recorded by anemometers at several sites on the ship. From this data set Kuhlbrodt (1936) derived a new equivalent scale. Quoting Köppen's method of data analysis, he calculated the regression of Beaufort against average wind speed. Since the 'Meteor' touched nearly all climate zones, Kuhlbrodt computed scales for the tropics and extratropics. In order to evaluate the quality of such scales for different climates we have to keep in mind that Köppen's method is a better approximation than the reverse technique. Nevertheless, Köppen's method still leads to one-sided regression lines, which do not necessarily coincide in different climate zones, even though a universal scale may exist. Taking into account those considerations, Kuhlbrodt's attempt to derive equivalent scales for different climate zones by one-sided regressions is rather questionable.

3.3 VERPLOEGH AND RICHTER

In the years that followed, many other equivalent scales were derived. Verploegh (1956) used observations from two light ships on the Dutch coast. Three hourly Beaufort estimates together with anemometer measurements from a height of 7 metres were available for the years 1950 and 1951. After averaging the anemometer measurement over 10 minutes, Verploegh sorted the observation pairs according to the wind speed and averaged the estimates. Thus, he followed Köppen's method. However, the finally recommended scale is based not only on Verploegh's own calculations, but is an average of different scales. Among others, the results of the 'Meteor' cruise and those of Simpson (1906) were also taken into account. A scale derived by Richter (1956) was also included. This is interesting because Richter was one of the first who rejected Köppen's method; he returned to the antiquated procedure of calculating the mean measured wind speed for each Beaufort class. Before merging the various scales, Verploegh discussed their differences. Richter's scale and his own scale showed considerable differences, especially for low wind speeds, which he tried to explain purely by the different anemometer heights. From today's standpoint this is only half of the story. The antiquated deriving procedure of Richter inevitably leads to higher equivalent values for weak Beaufort classes and to lower values for the strong Beaufort forces. The last effect is compensated by the larger anemometer height so that only the first remains evident.

3.4 THE SCIENTIFIC “CMM-IV SCALE”

In the course of subsequent years, the credibility of the old WMO Code 1100 declined. In 1970 the Commission for Maritime Meteorology recommended a new scale, the “CMM-IV scale”, intended especially for scientific applications. Based on observations from the period 1874 to 1963, a regression analysis between Beaufort force and metric wind speed was made. Unfortunately, it was again the antiquated method which was used for derivation. Well aware that it is important to choose which of the parameters is regarded as independent, the authors cited Köppen (1898) and Curtis (1897). But accidentally, the originally correct statement was reversed. Consequently, the low Beaufort equivalents were overestimated, while the strong ones were underestimated. Köppen propagated just what he had intended to prevent by his unorthodox deriving method.

3.5 THE KAUFELD SCALE

Kaufeld (1981) published a new scale based on a comparison of wind speed measurements from OWSs with the Beaufort estimates of nearby passing merchant ships. The large and extraordinarily well-suited raw data material, and even more so the regression method used, gave the Kaufeld scale outstanding importance. Kaufeld pointed out that none of the one-sided regressions was able to provide optimum results; and consequently he derived a scale based principally on the linear orthogonal regression. Since in reality non-linear relationships are expected, special procedures are necessary. Kaufeld applied two techniques, the construction of the angle bisecting the two one-sided regressions and the method of cumulative frequencies, both leading to similar results.

Kaufeld's method is generally in accordance with our recommendations to use the orthogonal regression. But for practical application, the relative error variances of both measurements and estimates have to be equal. Simpson and Köppen had supposed that this was not case, assessing the observation errors of Beaufort estimates to be larger than those of the measurements. Actually, this was their reason for preferring the one-sided regression of Beaufort on wind speed to the reversed regression, although they knew that both are not completely correct. Lindau (1995) showed that estimation errors are indeed larger than measurement errors, at least for measurements from OWSs. After compensating for these error differences, Lindau derived a new Beaufort equivalent scale. His procedure will be reviewed in the next chapter.

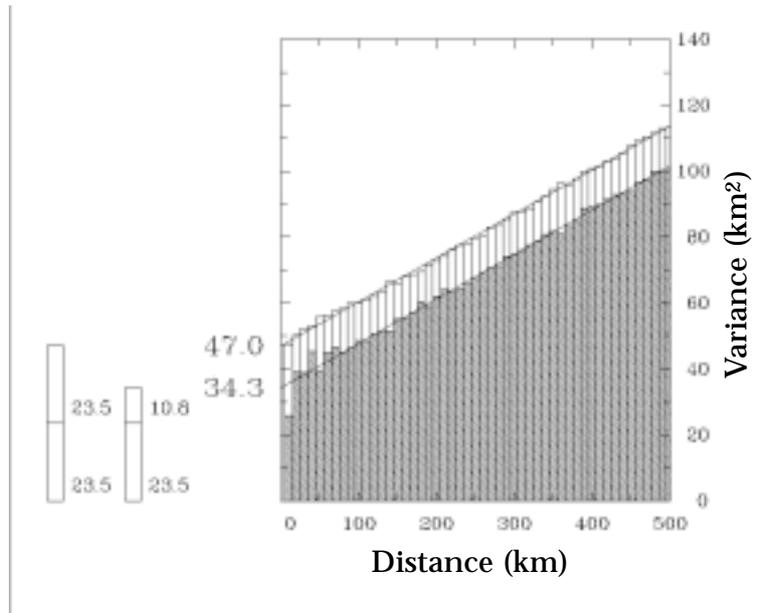
4. DESCRIPTION OF THE RECOMMENDED SCALE

4. Like Kaufeld, Lindau (1995) used the measurements of OWSs in the North Atlantic to calibrate the Beaufort estimates from merchant ships in their vicinity. Intending to apply the orthogonal regression method, the observation errors of both data sets were previously calculated.

In order to calculate the error variance of Beaufort estimates, pairs of simultaneous ship observations are considered as a function of the distance between both ships (Figure 9). As distance increases, the mean square difference between the two wind observations increases owing to growing natural variability. The error variance also contributes to the total variance, but it is independent of the distance and can be regarded as a constant for each distance class. For the potential distance of zero, the natural variability vanishes and only error variance remains. As pairs of ships are considered, the two error variances appear. Repeating the procedure with pairs of merchant ships and OWSs, the random observation errors of the OWSs measurements were computed. It turns out that their error variance is less than half that from merchant ship observations.

To compute a data pair, about six instantaneous merchant ship observations in the vicinity of the OWS are averaged and paired with daily means of the OWS measurements, consisting of only four individual observations. In this way, the larger errors of Beaufort estimates are compensated. However, since co-located and instantaneous measurements are not available, it is unavoidable that natural variability is also included in both averaging processes. For the OWS measurements, the temporal variability of one day is included. For the Beaufort measurements, an amount of spatial variability is included, depending on the extent of the radius of observations around the OWS. In order to attain comparable data sets, the spatial radius is computed to correspond to the

Figure 9—Determination of the mean observation error. Mean squared wind speed differences from VOS-OVS pairs (shaded) are compared to VOS-VOS pairs (total columns) as a function of distance between the ships. For the potential distance zero, only measurement errors contribute to the variance.



temporal variability of one day. Depending on station and season, radii of about 300 to 400 km were computed. After this procedure, Beaufort estimates and wind measurements should have the same accuracy and the same resolution (in a spatial respect for the Beaufort estimates and in a temporal respect for the OVS). By then applying the method of cumulative frequencies, the new scale was obtained (Table 1). Since the OVS measurements were previously reduced from 25 to 10 m, the scale is valid for a height of 10 m above sea level.

Beaufort	0	1	2	3	4	5	6	7	8	9	10	11	12
WMO	0.0	1.7	4.7	8.4	13.0	18.3	23.9	30.2	36.8	44.0	51.4	59.4	67.7
New	0.0	2.3	5.2	8.9	13.9	18.9	23.5	28.3	33.5	39.2	45.5	52.7	61.1
N	6	378	2287	8441	17197	11598	8870	4655	2068	597	122	15	1

Table 1—New 10 m-equivalent values (in knots) compared to WMO Code 1100. N gives the number of data pairs, which consist of daily means for OVS measurements and spatial means for Voluntary Observing Ships (VOSs).

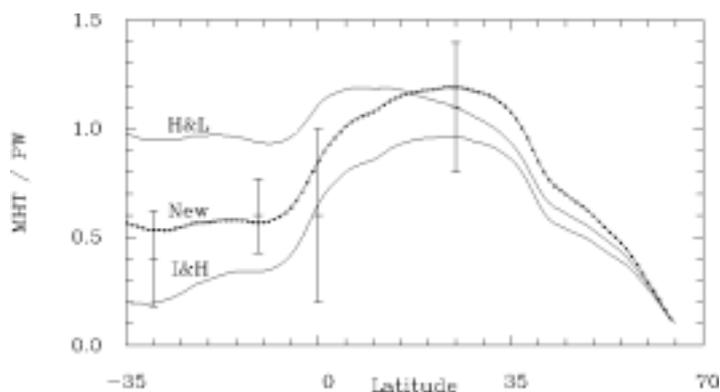
Kent and Taylor (1997) tested the performance of different Beaufort equivalent scales by comparing anemometer measured wind speeds with visual estimates, both from the Comprehensive Ocean-Atmosphere Data Set (COADS). An extraordinarily meticulous height correction of the measurements was performed using the individual anemometer heights for each ship. The agreement of converted Beaufort estimates with the corresponding measurements was checked for monthly averages within 1 degree squares. In conclusion, the Beaufort scale of Lindau (1995) was found to provide the most suitable conversion for the creation of a homogeneous monthly mean wind data set from anemometer and visual winds in COADS.

Lindau (2000) applied this scale to the marine meteorological reports of COADS. For the Atlantic Ocean, the wind-dependent latent and sensible heat fluxes, together with shortwave and longwave radiation, were calculated for a 40-year period. The reliability of the resulting total net heat flux field was estimated by comparing the induced meridional heat transport with independent oceanic measurements (Figure 10). A good agreement was achieved without any additional corrections which enhances the confidence in the used Beaufort scale.

5. CONCLUSIONS

The Beaufort scale derived by Lindau (1995) is recommended for converting visual estimates and metric wind speed. Especially for climate study purposes, it is essential that the characteristics of entire data sets are conserved when Beaufort estimates are converted into metric wind speed. A consistent conversion is only possible with orthogonal regression, whereas it is the domain of one-sided regressions to give the most probable wind speed for an individual Beaufort estimate,

Figure 10—When applying the Lindau scale on COADS, the turbulent fluxes at sea surface are calculated. Together with the radiative fluxes, the net energy exchange between ocean and atmosphere is concluded. The figure shows the northward meridional heat transport induced by the imbalance of the obtained net energy exchange, compared to the results of I & H: Isemer and Hasse (1987) and H & L: Hastenrath and Lamb (1978). Results of independent oceanographic sections are indicated at the relevant latitude together with their error bars: 30°S: Holfort (1994), 11°S: Speer et al. (1996), 0°: Wunsch (1984), 25°N: Bryden and Hall (1980).



and vice versa. However, if one-sided regressions are used at all, Köppen (1898) and Simpson (1906) realized independently that the regression of Beaufort on wind speed, i.e. considering the wind speed as the independent parameter, is at least more suitable to serve as an equivalent scale than the reversed regression. Larger errors in the estimates are the main reason for this, but different temporal resolutions of estimates and measurements, respectively, may also contribute. For a correct derivation of an equivalent scale, both effects - those of different errors and those of different resolutions - must be taken into account. Lindau (1995) equalized the different errors by averaging only a small number of measurements, but a somewhat larger number of estimates. At the same time, it was ensured that the included temporal and spatial variability was equal. This procedure resulted in a correct computation of the common relationship between Beaufort force and wind speed.

REFERENCES

- Bryden H.L. and M.M. Hall, 1980: Heat transport by currents across 25°N in the Atlantic Ocean. *Science*, 207, 884-886.
- Curtis, R.H., 1897: An attempt to determine the velocity equivalents of wind forces estimated by Beaufort's scale. *Quarterly J. Royal Meteorol. Soc.*, 23, 24 - 61.
- Hastenrath and Lamb, 1978: *Heat Budget Atlas of the Tropical Atlantic and Eastern Pacific Ocean*. University of Wisconsin Press, Madison, 90 pp.
- Holfort J., 1994: Großräumige Zirkulation und meridionale Transporte im Südatlantik. *Berichte aus dem Institut für Meereskunde an der Christian-Albrechts-Universität, Kiel*, 260, 96 pp.
- Isemer, H.J. and L. Hasse, 1987: *The Bunker Climate Atlas of the North Atlantic Ocean*. Volume 2: Air-sea interactions. Springer Verlag Berlin, 252 pp.
- Kaufeld, L., 1981: The development of a new Beaufort equivalent scale. *Meteorol. Rundschau*, 34, 17-23.
- Kent, E.C. and P.K. Taylor, 1997: Choice of a Beaufort equivalent scale. *J. Atmos. Ocean. Tech.*, 14, 228 - 242.
- Köppen, W., 1898: Neuere Bestimmungen über das Verhältnis zwischen Windgeschwindigkeit und Beaufortskala. *Archiv Deutsche Seewarte*, 5, 1 - 21.
- Köppen, W., 1916a: Beaufortskala und Windgeschwindigkeit. *Meteorol. Zeitschr.*, 33, 88 - 91.
- Köppen, W., 1916b: Die dreizehnteilige Skala der Windstärken. *Annalen d. Hydrographie u. Marit. Met.*, 44, 57 - 63.
- Köppen, W., 1926: Über geschätzte Windstärken und gemessene Windgeschwindigkeiten. *Annalen d. Hydrographie u. Marit. Met.*, 54, 362 - 366.
- Kuhlbrodt, E., 1936: Vergleich geschätzter Windstärken mit gemessenen Windgeschwindigkeiten auf See. *Annalen d. Hydrographie u. Marit. Met.*, 64, 14 - 23.
- Lindau, R., 1995: A new Beaufort scale. *Proc. Int. COADS Workshop* (Kiel, Germany). *Berichte aus dem Institut für Meereskunde*, 265, 232 - 252.
- Lindau, R., 2000: *Climate atlas of the Atlantic Ocean derived from the Comprehensive Ocean-Atmosphere Data Set (COADS)*, Springer Verlag, 514 pp.

- Richter, J., 1956: Geschwindigkeitsäquivalente der Windstärkeschätzungen nach Beobachtungen auf deutschen Feuerschiffen. *Annalen der Meteorologie*, 7, 267 - 287.
- Simpson, C.G., 1906: The Beaufort scale of wind-force. *Report of the Director of the Meteorological Office*, Official No. 180, London
- Speer, K.G., J.Holfort, T.Reynaud, G.Siedler, 1996: South Atlantic heat transport at 11°S. In: *The South Atlantic present and past circulation*, G. Welen & G. Siedler eds., Springer Verlag, Berlin (in press).
- Verploegh, G., 1956: The equivalent velocities for the Beaufort estimates of the wind force at sea. *Koninklijk Nederlands Meteorologisch Instituut, Mededelingen en Verhandelingen*, 66, 1- 38.
- Waldo, F., 1888: Vergleich von Beauforts Skala und Windgeschwindigkeit. *Meteorol. Zeitschr.*, 5, 239 - 241.
- World Meteorological Organization, 1970: The Beaufort scale of wind force (Technical and operational aspects), Commission for Marine Meteorology. *Reports to marine science affairs*. Rep. No. 3, WMO, Geneva, Switzerland, 22 pp.
- Wunsch, C., 1984: An eclectic Atlantic Ocean circulation model. Part 1: The meridional flux of heat. *J. Phys. Ocean.*, 14, 1712-1733.

EVALUATION OF OCEAN WINDS AND WAVES FROM VOLUNTARY OBSERVING SHIP DATA

Sergey K. Gulev*, Vika Grigorieva, Konstantin Selemenov and Olga Zolina, P.P. Shirshov Institute of Oceanology, RAS, Moscow

ABSTRACT

This paper considers the problem of the accuracy of Voluntary Observing Ship (VOS) wind and wave data, using individual wind and wave reports from the COADS (Comprehensive Ocean-Atmosphere Data Set). Additional information on the accuracy of marine wind and wave observations was available from a pilot questionnaire, SHIPMET, which was distributed among 400 marine officers with the aim of discovering the actual practice of marine meteorological observations onboard merchant vessels. The evaluation of true wind is one of the most important sources of error in wind observations. Estimates of the possible effects of inaccurate evaluation of true wind are presented. An estimation of random observational errors in wave parameters shows that wave fields can be successfully evaluated from the VOS data. Some approaches are recommended to remove systematic biases in visual wave estimates.

1. INTRODUCTION: CURRENT STATUS OF OUR KNOWLEDGE ABOUT VOS WIND AND WAVE OBSERVATIONS

Despite considerable progress in the development of satellite instruments and modelling, Voluntary Observing Ship (VOS) data are still the main source of our knowledge about ocean winds and waves, especially for the decades before the 1980s. During the last two decades these data have been assimilated in the Comprehensive Ocean-Atmosphere Data Set (COADS), which is currently the most complete collection of marine surface observations, assembled from the Global Telecommunication System (GTS) and log books and archived as Compressed Marine Reports (CMR) and Long Marine Reports (LMR) (Woodruff *et al.*, 1998). However, marine meteorological variables derived from COADS contain a number of biases and uncertainties connected with the observational accuracy and should be carefully validated before they are used for the flux fields production. In this context, wind and wave fields are the most 'questionable' and problematic VOS observations. Although the other surface variables (SST, SLP, air temperature and humidity) are also influenced by random and systematic observational errors, it is easier to assess their accuracy since they are exclusively instrumental observations. Taking into account the fact that winds and waves can be derived from satellite observations and model hindcasts with a better accuracy than other meteorological variables, we expect that the alternative products of these sea-air interface parameters will appear quite soon for the period covering the last several decades. In this context, it is very important to quantify the accuracy of the VOS winds and waves to provide the best possible VOS data possible for the intercomparison with remotely-sensed and model products. This paper addresses some issues on the accuracy of VOS wind and wave data on the basis of statistical analysis and new information about the actual measurement techniques used onboard merchant vessels which is provided by the questionnaire distributed among a representative population of officers.

(a) Wind observations

There are many uncertainties in VOS wind observations. First of all, a considerable part of wind observations are the visual estimates made by officers of

* corresponding author's address: P.P. Shirshov Institute of Oceanology, RAS, 36 Nakhimovsky ave., 117851, Moscow, Russia, Email: gul@gulev.sio.rssi.ru

merchant ships. The observational accuracy of these observations is reasonably low in comparison to in situ measurements. There are additional uncertainties connected with the systematic biases of different equivalent Beaufort scales in the low and high wind speed ranges. Instrumental wind observations show the mixture of measurements made by hand-held and fixed anemometers. Hand-held anemometer data are crucially affected by the ship's superstructure and sample procedure. Winds recorded by fixed anemometers are also influenced by the ship's superstructure; they are additionally affected by differences in anemometer heights onboard different ships and by the uncertainty of the procedure for evaluating true wind (or drop of it). Altogether, these uncertainties result in a coupled error which is comparable or overestimates the uncertainty of visual observations (Kent *et al.*, 1993). Finally, for the creation of climatologies, data from the fixed anemometers (hand-held anemometers are usually excluded from the analysis) are merged with visual observations. This leads to additional time- and space-dependent uncertainty of monthly averaging of inhomogeneous data.

During the last several decades the issue of the accuracy of wind observations at sea has been addressed in many works. To minimize biases in visual wind observations several alternative Beaufort equivalent scales were developed in addition to the traditionally used WMO Code 1100 scale (Cardone, 1969; WMO, 1973; Kaufeld, 1981; Isemer and Hasse, 1991; da Silva *et al.*, 1995; Lindau, 1995). Kent and Taylor (1997) comprehensively reviewed all these equivalent scales and found the Lindau (1995) scale to be the most unbiased. Considerable progress has been achieved during recent years as regards the problem of adjusting wind observations to a standard height, primarily by merging the WMO *International List of Selected, Supplementary and Auxiliary Ships* (WMO-No. 47) with the LMR available from the COADS collection. By matching the call signs from WMO-No. 47 and LMR, it is possible to get the actual observational heights of fixed anemometers at 30 to 60 per cent of marine carriers (Kent and Taylor 1997) and to adjust the wind to a standard level and neutral stability. During the VOS Special Observing Project in the North Atlantic (VSOP-NA) (Kent *et al.*, 1993), a large set of well documented surface meteorological data was collected in the North Atlantic mid-latitudes for the period from 1989 to 1991. Analysis of this data set makes it possible to quantify the most important biases in ocean wind observations and to implement the corrections. Laboratory and numerical modelling using typical ship superstructures helped to abate the impact of the ship on anemometer measurements (Yelland *et al.*, 1998). However, some biases remain unexplained. In particular, Gulev (1999) compared the high quality instrumental data to COADS winds for the 1980s and early 1990s in the north-west Atlantic, and found an overestimation of the COADS winds in low ranges and underestimation for the strong and moderate winds, i.e. the opposite tendency to that usually expected for such inter-comparisons. Since the application of the alternative equivalent Beaufort scales did not remove the bias and made it even more pronounced, it was concluded that such a disagreement results from the incorrect evaluation of true wind. Quantitative inspection of the procedure for evaluating true winds onboard merchant ships and correction of corresponding biases is difficult in contrast to research vessels, for example, whose data are much better documented (Smith *et al.*, 1999).

(b) Visual wave observations

VOS wave observations are exclusively visual estimates of heights, periods and directions of wind sea and swell. For a long time visual wave observations from limited collections were used to produce ocean wave statistics (Hogben and Lumb, 1967) and global wave statistics (Hogben *et al.*, 1986) widely used by sailors and naval engineers. Direct comparisons of wind waves to in situ observations from buoys and the other platforms (e.g. Wilkerson and Earle, 1990) reported about the large random and systematic errors in visual observations. Gulev and Hasse (1998, 1999) updated all visual observations in the North Atlantic from the COADS collection for the last 30 years and quantified many of the errors and

uncertainties. Intercomparison of Gulev and Hasse (1998) climatology with the European Centre for Medium Range Weather Forecasts (ECMWF) wave model (WAM) wave hindcast and altimeter measurements in the North Atlantic (Gulev *et al.*, 1998, Cotton *et al.*, 1999) shows a general similarity of spatial patterns and the co-location of local maxima. Although the mid-litudinal estimates of VOS waves were consistent with WAM hindcast and altimeter measurements, it has been found that overestimation of VOS waves in the tropics and subtropics is systematic. Small waves are relatively poorly resolved in WAM, leading to difficulties in their validation. However, small wave heights in the VOS data are influenced by systematic bias, which will be analysed below.

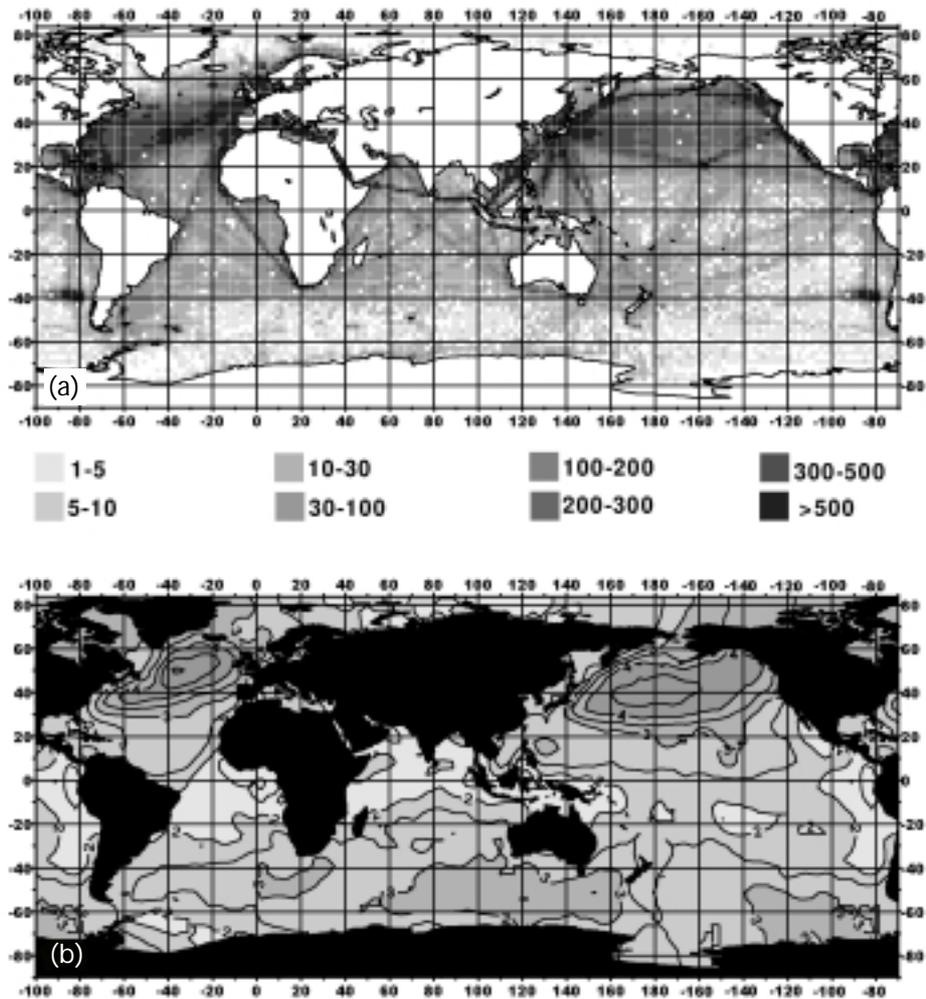
Considering the validation of the VOS waves over the global ocean, one of the main weaknesses of the VOS data is the inhomogeneity of the data coverage. In general, there is concern that the observational density of wave observations is considerably smaller in comparison to the other variables. Gulev and Hasse (1998) reported that 40 to 60 per cent of the total number of reports for the North Atlantic include wave information. Note that for the 1963-1979 period they used CMR and LMR that were available only from 1980. The inspection of newly updated LMR (Woodruff *et al.*, 1998) shows that the percentage of reports with wave observations closely matches 70 to 80 per cent of reports which contain wind information. Figure 1(a) shows the total number of reports with wave parameters for 2-degree boxes over the Global Ocean during January for a 15-year period from 1979 to 1993. The distribution of the number of wave reports is qualitatively similar to that of the other meteorological variables. Mid-litudinal and subtropical regions in the northern hemisphere are much better sampled than the southern hemisphere, where high observational density is observed primarily along the major ship routes. The quantitative comparison of the number of wave observations with the observational density of the basic variables (widely used for the creation of global scale climatologies) shows that the estimated 70 to 80 per cent of wind reports derived for the North Atlantic remains valid also for the other oceans. Even in the Southern Ocean, considered to be very poorly sampled, the density of wave observations considerably overestimates the density of humidity observations. In general, we can conclude that the number of observations is large enough for the creation of global scale climatology at least north of 40S.

The main problem with the validation of visual wave observations against instrumental measurements is the evaluation of significant wave height (SWH), which is usually reported in instrumental records. The traditional approach to deriving SWH from separate visual estimates of wind sea and swell is to apply the formula (Hogben *et al.*, 1988):

$$SWH = (h_w^2 + h_s^2)^{1/2} \quad (1)$$

where h_w and h_s are wind sea and swell heights, respectively. The results of intercomparison with instrumental measurements (Gulev and Hasse, 1998) show that Formula (1) overestimates the observed SWH in the majority of cases by several tens of centimeters with a mean deviation of -0.27 m. An alternative estimate of SWH was established by Wilkerson and Earle (1990), who analysed buoys, a majority of which had been deployed in the subtropics, and found that the highest of the two estimates was less biased. However, intercomparison of the measurements in both subtropics and mid-latitudes (Gulev and Hasse, 1998) showed a tendency of frequent underestimation of this SWH estimate. The best estimate of SWH was found to be a combined estimate, computed as recommended by Barratt (1991) (i.e. applying (1) when sea and swell are within the same 45° directional sector, and taking the higher of the two components in all other cases), but the optimal directional sector was found to be 30°. This combined estimate gives the mean 'buoy minus VOS' difference of -0.07 m in the Atlantic (Gulev and Hasse, 1998). The combined approach was chosen for the production of new global wave climatology recently developed at IORAS (Gulev *et al.*, 2001). Figure 1(b) shows an example of a climatological chart of January SWH computed using the combined estimate for the 1979-1993 period. It shows reasonable heights in the North Atlantic and North Pacific mid-latitudes. In the

Figure 1—Number of COADS reports with visual wave observations in climatological January for the period from 1979 to 1993 (a), and climatological January significant wave height over the Global Ocean (b).



South Atlantic, where the number of wave observations is considerably smaller, our climatology does not indicate ‘the belt’ of large wave heights as in model hindcasts. This difference results mostly from undersampling and considerable efforts are required to develop new procedures for the optimal interpolation of wave characteristics in poorly sampled areas.

Visual estimates of wave periods were found to be systematically underestimated in the VOS observations. Wilkerson and Earle (1990) reported about 0.2 sec ‘buoy minus VOS’ differences. Gulev and Hasse (1998) found that mean departure is about 0.26 sec with a std. dev. of 0.1 to 0.6 sec. Dacunha *et al.* (1984) and Hogben (1988) reported even larger systematic biases in periods for the Cobb seamount in the North Pacific. To correct biases several methods were developed. Ochi (1978) and Dacunha *et al.* (1984) recommended correcting joint probability distributions of wave heights and periods, making it possible to obtain the corrected mean periods. Gulev and Hasse (1998) developed a method for the correction of individual observations for periods with an accuracy of 0.12 sec. This method is based on the consideration of joint probability distributions of wave height and period in 17 locations of the North Atlantic. The application of this correction to the North Atlantic wave climatology shows that the largest corrections of 0.4 sec for sea periods and 0.8 sec for swell periods were applied in the north-east Atlantic.

Estimation of the observational accuracy of visual VOS wave data (Gulev and Hasse, 1999) shows that the day minus night-time difference in the visual wave estimates is not as large as in wind observations. In the North Atlantic it ranges from several centimetres to 0.2 m and does not have any pronounced spatial pattern. Another possible source of error in visual estimates of ocean waves is a poor separation of seas and swells by the observers. Gulev and Hasse (1999) tested

the success of this separation using joint probability distributions of the wave height and wind speed for the wind sea and swell, which were overplotted by the JONSWAP curves, representing wave height as a function of wind speed and duration in the formulation of Carter (1982). Most of the wind sea observations were bracketed by the JONSWAP curves corresponding to the 6- and 18-hour durations. Alternatively, only less than 20 per cent of swell observations were bracketed by the JONSWAP curves. Thus, there is evidence of quite good separation of seas and swells in the VOS wave data. However, as is the case with wind observations, further improvements to the accuracy of VOS wave data requires information on how the observations are actually made by the officers onboard merchant ships.

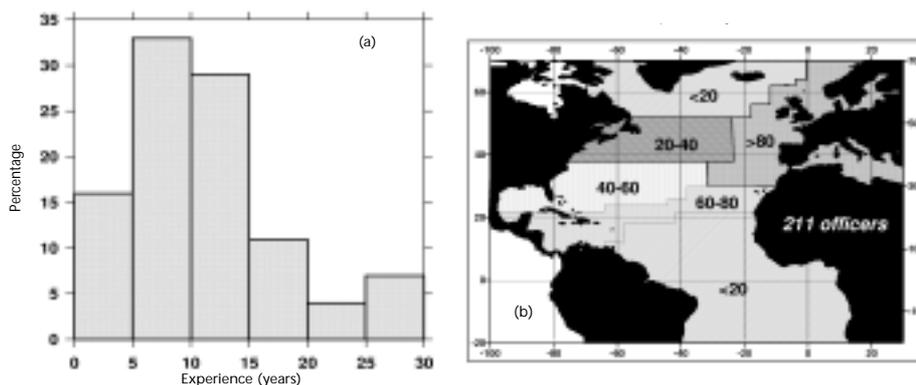
2. THE SHIPMET QUESTIONNAIRE

Houmb *et al.* (1978) were probably the first to report on interviewing observers as a method of estimating the observational accuracy of meteorological data. They investigated the effects of changing from meteorological assistants to mates on the visual wave estimates on some Norwegian ships. They found that mates tended to underestimate wave height in comparison to meteorological assistants. Unfortunately, this practice was not widely distributed for assessments of the observational accuracy of winds and waves at sea. To assess the impact of the 'personal factor' on the accuracy of VOS wind and wave observations, we designed the SHIPMET questionnaire (Gulev, 1996) and distributed it for a pilot pool among nearly 400 Russian ship officers. Such a population is considered to be very representative for narrow professional questionnaires. The questionnaire contained more than 60 groups of questions about the technical details of different meteorological observations (not only waves and winds) onboard different marine carriers. These questions were assembled according to the requirements of sociological pools. When one question is repeated at least several times in different contexts, it provides the possibility of testing the reliability of the given answers using the answers to the so-called 'sister questions'. Before the final list of questions was established, 11 officers were interviewed in a free manner. These interviews helped to provide details on the techniques used. Important questions were asked on which operational guide to use as the reference for the survey. In different Russian fleets (merchant, military and fishing) slightly different guides were used. Finally, the guide for the merchant fleet was taken as the reference and it was assumed that all officers were familiar with it. This guide elaborates on most of the details of meteorological observations for different types of meteorological onboard equipment. Some sample questions from the SHIPMET questionnaire which are analysed in this study are given in the Appendix below.

The 'response function' of the officers was quite good and more than 2/3 of the questionnaires were answered. After the expertise of the answered questionnaires, aimed at excluding unreliable samples, 211 of them were selected for the statistical analysis. Most sailors who participated in the pool were mid- and high-level officers and had the rank of mates, although approximately 15 per cent of them were sailors, appointed as low-rank officers after a couple of years of sailing experience. Figure 2(a) shows the distribution of officers' sailing experience. Most officers (63 per cent) have 5 to 15 years experience and this reflects the typical distribution of the experience of officers in most Russian ship companies. Figure 2(b) shows the regions in the North Atlantic where these officers operated. We asked them to mark roughly the most frequent ship routes along which they travelled. Thus, this picture contains some uncertainty. Nevertheless, it correlates well with the typical observational density over the North Atlantic (if the American and Canadian carriers are excluded), and we believe that we achieved an adequate representation of geographical regions. Most officers travelled along the ship routes that cross the North Atlantic mid-latitudes and subtropics and also the European basin.

Figure 3 displays some pilot results of the statistical analysis of the officers' answers to the questions concerning the determination of wind speed onboard the vessels. Figure 3(a) shows that most of the officers were quite familiar with the Beaufort scale details. Seventy five per cent of respondents either use a table with the description of the Beaufort scale or know it with varying degrees of accuracy.

Figure 2—Distribution of experience of respondents of the SHIPMET questionnaire (a) and North Atlantic areas where officers travelled most frequently (b). Numbers indicate the number of officers marked this region.



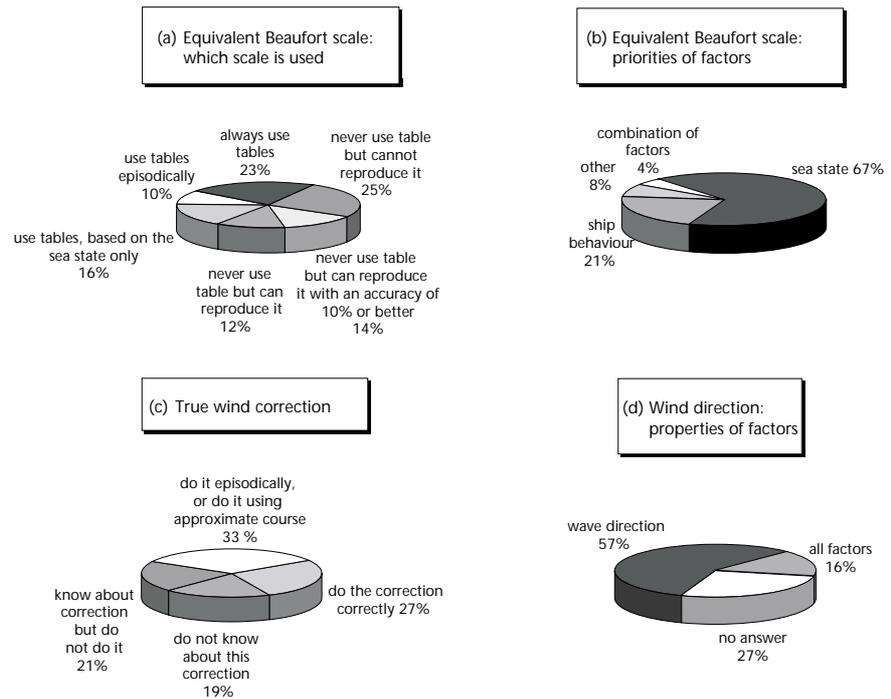
About 16 per cent of respondents use the reduced tables based on sea state only. However, Figure 3(b) shows that less than 1/3 of all officers account for ship behavior and other factors and estimate the Beaufort number, i.e. sea state remains the highest priority for most officers when determining the Beaufort number. If we assume arbitrarily that poor familiarity with the Beaufort table may result in the error of ± 1 Beaufort number for low and moderate winds and of ± 2 Beaufort numbers for strong winds, and apply these error estimates to 25 per cent of officers, who were completely or partly unaware of Beaufort scale details, the resulting absolute error of the reported Beaufort numbers will be ± 0.25 and ± 0.5 respectively.

Figure 3(c) shows results of the analysis of the evaluation procedure of true wind onboard merchant vessels equipped with anemometers. According to the answers of respondents, 19 per cent of officers do not know about the technique for evaluating true wind; 21 per cent know, but do not usually use it; 33 per cent use it either episodically or using the “approximate course and ship velocity”; and only 27 per cent use it correctly. Thus, according to our pool, about 40 per cent of officers omit true wind correction. Assuming roughly that anemometer measurements contribute 30 to 50 per cent of the total number of wind observations, the actual contribution of uncorrected winds is about 12 to 20 per cent of all wind reports. Additionally, considerable uncertainty stems from the 33 per cent of officers who do this correction episodically or using an approximate (i.e. expected, and not reported by the navigation system) ship course. Assuming very tentatively that half of the reports by this 33 per cent of officers can be considered as uncorrected, and using the same estimate of 30 to 50 per cent for the contribution of anemometer winds, we can increase our estimates of the total percentage of uncorrected winds by at least 5 to 8 per cent.

If we consider the determination of wind direction in the absence of an onboard anemometer (Figure 3(d)), nearly 60 per cent of respondents report it from the wave direction, and more than a quarter of sailors do not explain how the direction is derived. In this sense, the recently introduced simplification of the LMRF format (use of wind direction when the wind sea direction is not reported or deviates considerably from the wind direction) seems to be a reasonable step. At the same time, the situation with the evaluation of true wind from the relative wave direction for visual wind estimates (not shown here) is somewhat better than with the correction of anemometer winds. More than 80 per cent of officers ensured that in this case the reported wave and wind directions were absolute and not relative, although different approaches for evaluating the absolute directions were reported. Smith *et al.* (1999) reported on frequent confusions concerning the definition of true wind, used by meteorologists, oceanographers and the merchant marine. However, in our pool, more than 90 per cent of officers among those who are familiar with the technique of true wind correction used the meteorological definition of true wind (i.e. speed referenced to the fixed earth, and direction referenced to true north).

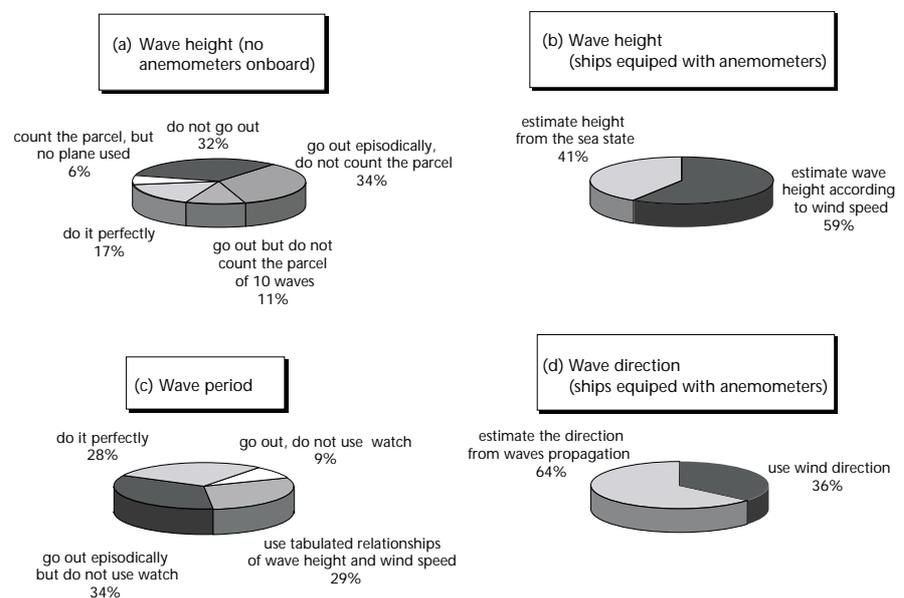
Figure 4 shows the results of the analysis of the observational techniques of visual wave estimates. Figure 4(a) summarizes how the officers of ships without onboard anemometers follow the recommendations on wave

Figure 3—Results of the SHIPMET questionnaire: (a) familiarity with the Beaufort table, (b) priority of factors for Beaufort estimates, (c) approaches used for true wind correction, (d) priorities of factors determining wind direction.



height measurements. These recommendations require the use of a special plane for the measurements of wave height, at least during the daytime. Ideally, the estimates of both wave height and period should be taken as an ensemble average within the parcel of 10 waves. In order to count the parcel, it is recommended to use a buoyant piece of red, yellow or white material to mark the reference point. The use of a watch is strongly recommended for the estimation of periods. However, in practice, many approaches are used. Remarkably, 32 per cent of observers do not even leave the bridge to make wave measurements. Although this has to affect the accuracy of visual wave estimates, note that the wind speed estimated by these sailors will be affected to the same degree. Only about 23 per cent of respondents reported that they count the parcel of 10 waves and about 17 per cent use a special plane during the daytime to estimate wave height. Forty-two per cent of observers report waves with intermediate accuracy, i.e. they leave the bridge, watch the sea surface but do not count the parcel of 10 waves, and do not use the plane to estimate wave height. Our questionnaire shows that high uncertainties of visual wave estimates can be

Figure 4—Results of the SHIPMET questionnaire: (a) approaches to determining wave height on ships without anemometers, (b) approaches to determining wave height on ships equipped with anemometers, (c) approaches to determining wave period, (d) approaches to determining wave direction.



associated with reports from vessels equipped with anemometers (Figure 4(b)). Officers onboard these ships consider ocean waves to be a low priority parameter with respect to wind, which is measured by fixed anemometers. Nearly 60 per cent of these sailors either directly use in situ measured wind speed to estimate wave height, or at least take wind measurements into account when they estimate waves. Assuming, as before, that 30 to 50 per cent of vessels are equipped with anemometers, in 15 to 30 per cent of cases we deal with some kind of simplified wave hindcast carried out by observers using wind information. Note that for ships equipped with anemometers 36 per cent of observers directly report wind direction as wind sea direction; and it is clear that many of these reports are affected by an inaccurate evaluation of true wind. Joint consideration of Figures 3(d) and 4(d) shows that the reported wind direction and wind sea directions in the majority of cases are not independent estimates in the VOS observations. Wave periods are perfectly estimated (i.e. a using watch) by just 28 per cent of respondents. In 45 per cent of cases we can expect that the low accuracy of reported wave periods are caused by simplifications of the observational technique. An important lesson is that nearly 30 per cent of observers use tabulated relationships between wind speed and wave height in order to estimate wave periods. Inspection of the tables used (the origin of these tables is always unknown) shows that their application usually results in the systematic underestimation of wave periods, which can partly explain the general underestimation of visual wave period estimates. The correction of these biases in wave periods requires the application of the procedures mentioned above in section 2.

Although the processing of the results of the SHIPMET questionnaire is still under way, the first pilot results show that the actual uncertainties inherent to the VOS collections of marine observations may be considerably larger than we expect from the traditional estimates of random and systematic errors in marine observations. Particularly, some of the reported approaches can result in systematic biases which should be taken into account. First of all, this is an inaccurate evaluation of true wind. There are reasonable questions about the reliability of the results of the SHIPMET questionnaire itself. According to sociological statistics, the random errors of narrow professional questionnaires are even higher than for the typical public pools, and ranged from 5 to 10 per cent. Thus, many of the conclusions based on these questionnaires should be considered more qualitatively rather than quantitatively. The most important question of all is whether we should believe that officers report reality rather than what the questionnaire expects of them (i.e. cite the instructions). The motivation of respondents is different for public relation pools and for professional pools, and this may result in additional uncertainty. An additional problem is connected with the question of whether Russian officers are representative of officers from other parts of the world. We estimated biases in winds and waves reported by the officers of different nations in the North Atlantic (Gulev and Hasse, 1998), using country code in COADS, and did not find any significant climatological biases. However, it is obvious that fleet-to-fleet differences in observational practices can be quite significant, especially if we consider the North Pacific where there is considerable contribution from the Japanese vessels.

3. ESTIMATION OF OBSERVATIONAL ERRORS IN MARINE WIND AND WAVE OBSERVATIONS

Kent *et al.* (1999) recently estimated random errors in basic meteorological variables reported by VOSs using the semivariogram technique. Random errors of the wind speed in the North Atlantic ranged from 2.0 to 2.3 m/s and did not indicate any significant spatial variability. This random observational error couples many particular uncertainties which affect wind observations at sea, and partly, of course, account for the random part of uncertainty associated with the evaluation of true wind. In general, there is concern that in many regions the problem of true wind evaluation does not seriously affect wind climatology. This is because on major ship routes the underestimation (overestimation) of true wind (if the correction is not done) when travelling in one direction will be compensated by the overestimation (underestimation) of true wind when travelling in the

opposite direction. However, this concern is based on several assumptions which may not necessarily apply to all regions. First, it is assumed that the directional steadiness of the dominant winds is quite high and is not affected by synoptic variability, or that the latter exhibits the random process. Secondly, it is assumed that ships always take the same routes travelling in both directions. The first assumption seems to be reasonable, at least for the mid-latitude regions, however weather regime changes may play an important role in wind speed and direction variations on weekly time scales. As regards the second assumption, it should be noted that many ships which contribute to the VOSs do not shuttle between two regions, but operate in different regimes. Moreover, marine carriers now use different routes travelling to the west and to the east. For instance, for the Newfoundland basin, the majority of ships, following the recommendations of meteoservices, use the southern routes when travelling from Europe to the USA and cross this region only on the way back. In this case, winds in this region will be slightly underestimated if the true wind evaluation is omitted (Gulev, 1999). Separate consideration of the zonal and meridional components of wind speed for this region shows that zonal wind speed (mostly affected by the ‘true wind effect’ under the dominant wind directions and ship routes for this region) indicates the larger disagreement between the COADS and high quality instrumental measurements than the meridional component.

We can demonstrate very roughly the possible bias in monthly climatological wind speed, which may result from an inaccurate evaluation of true wind from relative wind. Using the ship course and velocity reported in LMR, we recomputed winds for the North Atlantic assuming that the correction of relative winds was not applied at all. We also applied the reverse convergence to the winds, which were properly corrected. According to the SHIPMET pool, approximately 40 per cent of anemometer winds were not corrected. Thus, after the application of this procedure, approximately 60 per cent of wind observations were converted from true winds to relative winds and 40 per cent were corrected. Figure 5 shows the difference between climatological wind speed computed from the original VOS reports and from the reports after the ‘overall’ correction. All Beaufort estimates were used in the averaging of both arrays without any correction. The considerable positive difference of 0.5 to 1 m/s in the north-west mid-latitude Atlantic shows that the actual increase in the number of uncorrected reports (due to the application of the true wind correction to the already corrected winds, which were assumed to constitute a majority of reports) results in the underestimation of climatological wind in this region. Alternatively, overestimation is observed in the North Atlantic tropics and subtropics.

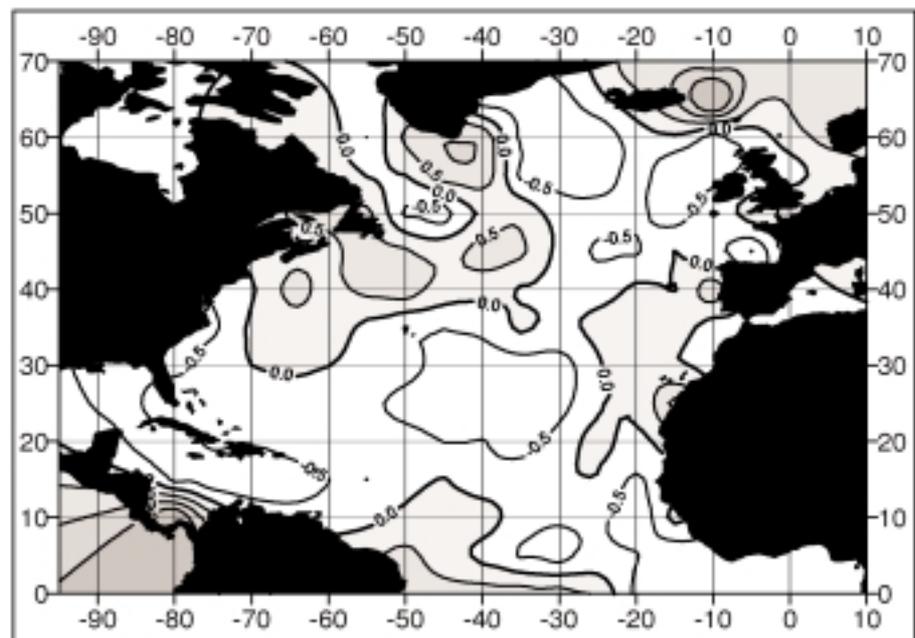
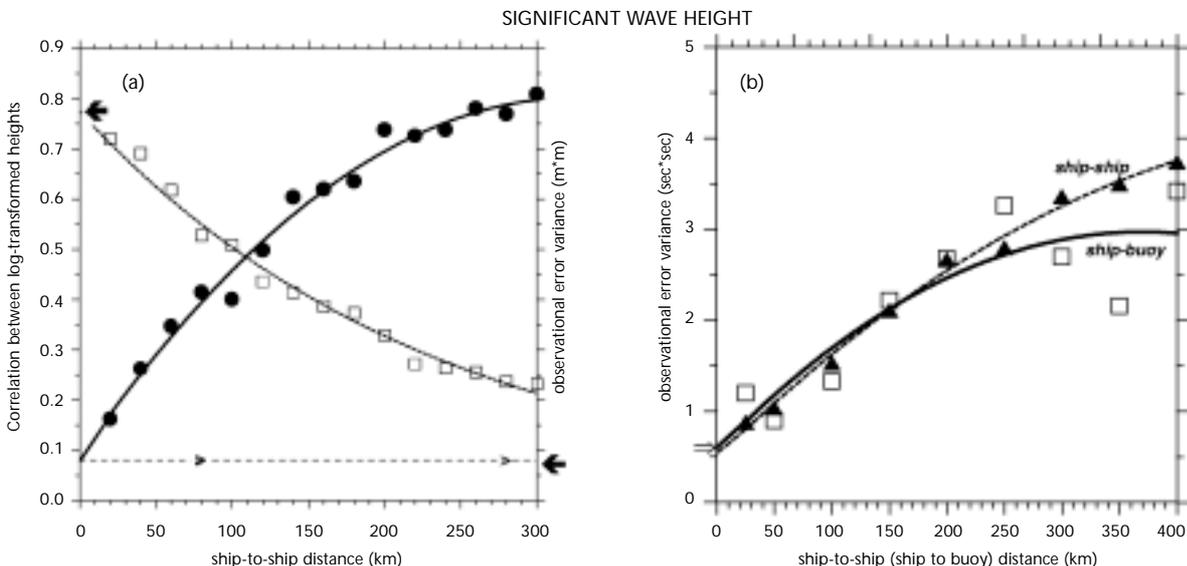


Figure 5—Difference between climatological wind speed in the North Atlantic ocean and wind climatology computed after the application of true wind correction to all anemometer measurements (see explanations in the text).

To estimate random observational errors in the visual wave observations, Gulev and Hasse (1999) used the approach of Lindau (1995) and Kent and Taylor (1997), who recommended computing the differences between simultaneous observations for certain classes of ship-to-ship distances. When the distance is equal to zero, natural variability does not contribute to the total variance, and the latter should represent only the error variance σ_o^2 , which has to be divided by two to get the squared measurement error $\epsilon_m^2 = \sigma_o^2/2$ (Lindau, 1995). To arrive at the σ_o^2 estimate, the polynomial extrapolation has to be used. An alternative approach was suggested earlier by Laing (1985), who introduced the dependence of the correlation between the log-transformed wave heights (r) on ship separation (x) as $r(x) = r_o \exp(-kx)$, where estimates r_o are reasonably not influenced by the spatial variability and should characterize the observational error.

Figure 6(a) shows the results of estimation of both ϵ_m^2 and r_o for significant wave height estimate for the North Atlantic Ocean after Gulev and Hasse (1999). The resulting estimate of r_o gives 0.76 for significant wave height, 0.69 for the wind sea, and 0.73 for the swell height. When we consider the regional correlations for 20-degree areas, the lowest correlation from 0.50 to 0.60 is found in the Western Atlantic subtropics and the highest (of about 0.83) in the North Atlantic mid-latitudes. The polynomial fit for ϵ_m^2 gives a standard deviation (std.) error of about 0.85 m² at $\Delta x=0$. However, if we use only classes of distances from 20 to 180 km, this estimate will be lower by about 0.07 m². We made an additional estimate for the class 0-10 km only and got ϵ_m^2 , which is a little bit less than 0.8 m². Figure 7 shows spatial distributions of the error estimates for wind sea and swell heights over the North Atlantic Ocean, computed by Gulev and Hasse (1999) for 20-degree boxes. The largest observational error of wind sea height of about 0.8 - 0.85 m² is obtained in the western subtropics, and the minimum (0.55 - 0.6 m²) is located in the eastern mid-latitudes. The spatial distribution of the observational error in swell height is quite different from that of the wind sea. The minimum error of around 0.8-0.85 m² is observed in the eastern mid-latitudes and the central subtropics and tropics. The largest errors up to 1 m² are observed in the western North Atlantic. A similar estimation of the random observational error in the resultant wave periods (after the correction of the wind sea and swell periods) has been carried out by analysing ship pairs and the data from the NDBC buoys in the subtropical Northwest Atlantic (Figure 6(b)). Although there is a disagreement between the two error estimates for large distances, the obtained ϵ_m^2 is quite comparable for both tests and closely matches 0.6 sec². However, this error grows by approximately 50 per cent in the mid-latitudinal North Atlantic. Thus, despite the fact that, according to the SHIPMET pool, less than 30 per cent of officers estimate wave periods perfectly, relative random errors in wave periods are not very high with respect to the observed magnitudes of seasonal and

Figure 6—Semivariogram (bold line, solid circles) and correlation (dashed line, boxes) estimates of random observational error in SWH over the North Atlantic (a) and estimates of random observational error in resultant wave period for ship-buoy (solid line, triangles) and ship-ship (dashed line, boxes) pairs (b).



interannual variability (Gulev and Hasse 1998, 1999). Furthermore, we can point out that the uncertainty of observational procedures results primarily in the systematic underestimation of wave periods.

Considering the possible systematic biases in wave height, we have to mention first of all the systematic overestimation of small seas and swells in VOS data. This overestimation results from the usage of the code figure '1' which is applied in COADS LMR to all waves smaller than 0.5 m. Therefore, all sea heights coded as '1' should represent a value that is somewhat lower than 0.5 m. Particularly, Gulev *et al.* (1998) found that the tropical VOS wave heights are slightly lower in comparison to the altimeter data and WAM hindcast. To resolve this problem we considered two-dimensional frequency distributions of wind speed and wave height for small waves, computed using instrumental data from NDBC buoys and from the VOS reports which give '1' as the measure of wave height and were sampled simultaneously with buoy measurements within a radius of 50 km. Buoy records report significant wave height and do not provide separate measurements of sea and swell. Thus, we selected the cases with the absence of swell in the VOS reports for this comparison. We required that the VOS wind speed estimate should not deviate from the wind speed measured at buoy by more than on 1 m. In total, more than 350 pairs of buoy and VOS

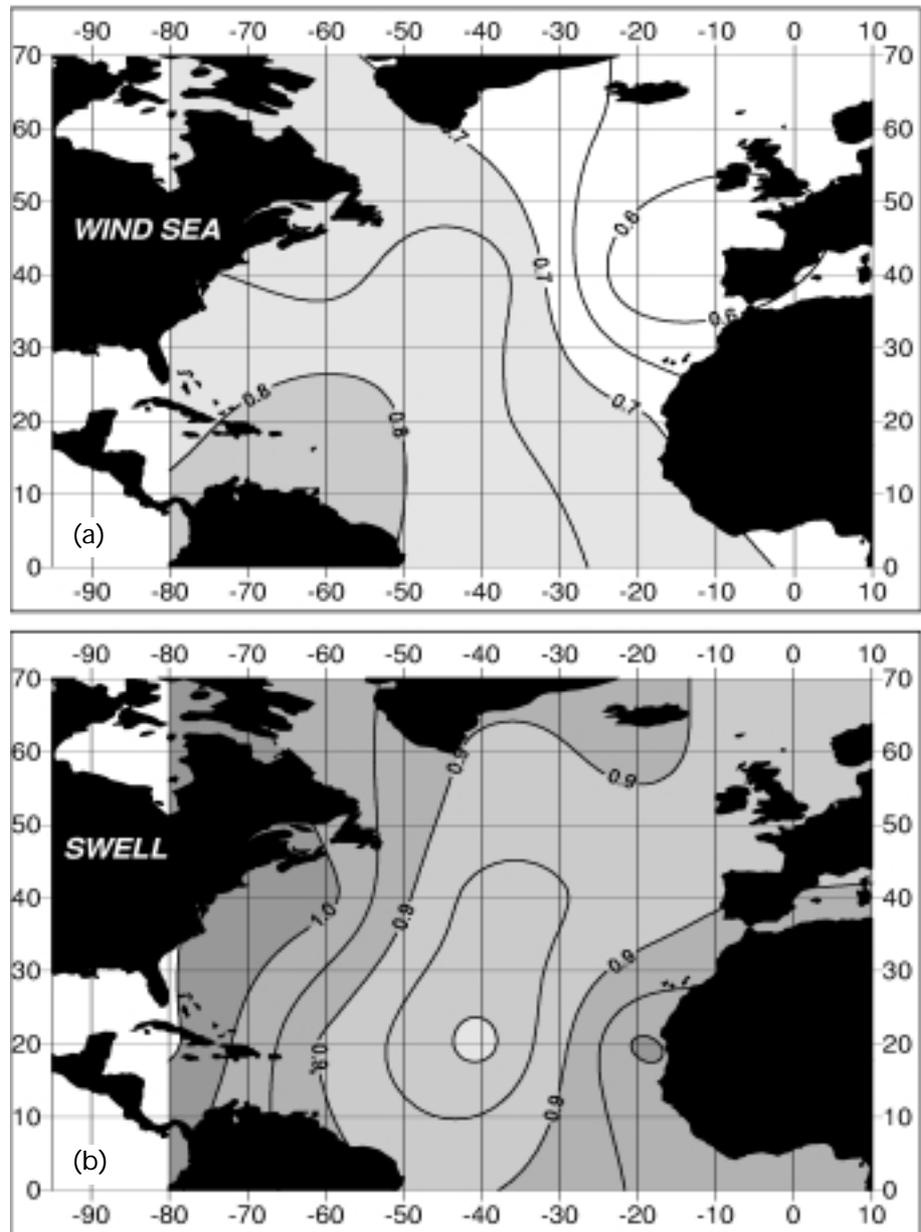


Figure 7—Spatial distribution of the random observational error (m^2) in the wind sea (a) and swell (b) visual estimates over the North Atlantic ocean. Contour interval $0.05 m^2$.

measurements of SWH and wind speed were chosen, primarily in the Gulf of Mexico and in the subtropical Atlantic. The analysis of probability functions for the wind speed, derived from the buoys and VOS reports, showed frequency distributions that were very close to each other. Subsequently, the two-dimensional probability density distribution of wind speed and wave height from the buoy measurements was considered for the wind speed range of 1.2 to 6 m/s and wave height of less than 0.5. For this range we derived a simple formula which can be used for correcting VOS wave height. The corrected sea height, reported with the code figure '1', has to be derived as $h_s = 0.5 - \exp(-0.658V)$, where $1.2 \leq V \leq 6$ is a wind speed. This formula makes it possible to correct small sea height with an accuracy of better than 20 per cent. Our attempts to derive the accurate correction of small swells were less successful. However, we can recommend with an accuracy of 30 per cent applying the correction of 0.15 m to all swells reported with the code figure '1'. Further details on the evaluation of small waves in COADS are given in Gulev *et al.* (2001).

4. SUMMARY AND CONCLUSIONS

We reviewed the accuracy of the VOS wind and wave observations using traditional statistical estimates together with the results of the questionnaire aimed at shedding light on the observational practices used by marine officers. Our initial experience with the questionnaire distributed among marine officers shows that this was quite a helpful tool for improving our knowledge of the actual uncertainties of winds and waves reported by VOSs. The statistical analysis of the pool results gives a reliable, although primarily qualitative, picture of the main sources of uncertainties inherent to the VOS observations.

It has been shown that the evaluation of true wind remains one of the most important sources of uncertainties in marine wind observations. Using ship course and velocity together with estimates of the percentages of uncorrected reports, which are available from the SHIPMET, it is possible to estimate roughly the possible error associated with the true wind correction, but it is still unclear how to correct the biases. As noted by Kent *et al.* (1993) and Gulev (1999), requirements to report both true and relative wind, or the relative wind only (even if satisfied), may result in additional uncertainty; this can affect the homogeneity of historical data. A remarkable example of this kind was the change of WMO swell period codes in 1968. This change was not simultaneously accepted by all nations and ship companies, and resulted in the biased swell periods for 1968-1969. However, the creation of some high quality regional and time limited subsets of the VOS data (like VSOP-NA) is very important for marine climatologists.

Wave parameters visually observed by marine officers can be successfully derived from the COADS collection of marine meteorological observations. Although the sampling frequency of wave observations is somewhat smaller in comparison to wind and temperature observations, it makes it possible to produce global and basin scale climatologies. However, south of 40S, VOS wave products should be considered with great care owing to the considerable undersampling of this region. It should be noted that VOS climatologies of the other parameters also show large sampling errors in these latitudes. We quantified the accuracy of visual wave observations. Random observational errors ranged from 0.5 to 0.85 m² for wind sea, from 0.8 to 1 m² for swell and from 0.4 to 0.7 sec² for wave periods. Beside the random observational errors, visual wave observations are influenced by some systematic biases. In particular, small waves are overestimated because of the coding system in the COADS, and periods are also underestimated by several tens of seconds. Simple corrections of these biases can be applied. At the same time, systematic biases associated with the differences in observational practices during the day and at night-time were not found.

Results of the SHIPMET questionnaire show that the approaches adopted by the officers affect the wave observations in the same degree as visual wind estimates. In this sense, the relative observational accuracy of wave observations should not be worse than for Beaufort wind estimates. Comparisons of visual wave estimates with instrumental measurements and alternative global data

(Gulev *et al.*, 1998) show a number of systematic biases between the VOS waves and alternative wave products. In particular, there is an evident overestimation of small waves, an underestimation of high waves, and an overall underestimation of wave periods. It should be noted that small and high waves in the model hindcasts and satellite products are also of a low accuracy. In this sense, buoys and the other in situ platforms provide in situ observations of a very high value, but further efforts are needed to obtain reliable estimates of sea and swell from wave recorders. Otherwise, direct comparisons with the VOS data will always be influenced by the uncertainty of evaluation of SWH from visual estimates of sea and swell. To quantify and correct the systematic biases in the VOS wave observations it could be desirable in the future to establish some kind of analog of equivalent Beaufort scale(s) for visual wave estimates.

Results of the SHIPMET questionnaire show that visual estimates of ocean waves and winds are, in fact, largely influenced by each other and are not fully independent. A considerable amount of wave observations are actually simplified local wave hindcasts carried out by sailors on the basis of wind information. The standardization of COADS formats also works in this direction. In this context, it is difficult to use jointly wave and wind information to cross-check the quality of wind and wave parameters. Nevertheless, these checks, if performed for the limited collections of truly independent observations, can help to considerably improve the accuracy of both wind and wave fields.

ACKNOWLEDGEMENTS

We would like to thank David Cotton of SATOBSYS (Godalming), Lutz Hasse of IFM (Kiel), Neil Hogben of BMT Ltd. (Teddington) and Andreas Sterl of KNMI (de Bilt) for the useful discussions. We appreciated the suggestions and criticism of two anonymous reviewers which helped to improve the paper. The COADS data for this study were made available courtesy of Steve Worley of NCAR (Boulder). Special thanks are also given to all marine officers who agreed to participate in the SHIPMET pool. This study was supported by the Russian Ministry of Science and Technology under the 'World Ocean' National Programme and EU INTAS grant 96-2089.

REFERENCES

- Barratt, M.J., 1991: *Waves in the North East Atlantic*. UK Department of Energy Report OTI 90545, HMSO, London, 16 pp. (with 40 figures).
- Cardone, V., 1969: *Specification of the Wind Distribution in the Marine Boundary Layer for Wave Forecasting*. Report TR69-1, New York University, N.Y., 131 pp.
- Cotton, P.D., P.G. Challenor, L. Redbourn-Marsh, S.K. Gulev, A. Sterl and R. Bortkovsii, 1999: An intercomparison of in situ, voluntary observing, satellite data, and modelling wind and wave climatologies. *Preprints of the CLIMAR99-WMO Workshop on Advances in Marine Climatology* (Vancouver, Canada, 1999), 110–120.
- Carter, D.J.T., 1982: Prediction of wave height and period for a constant wind velocity using the JONSWAP results, *Ocean Engng.*, 9, 17-33.
- Dacunha, N.M.C., N. Hogben and K.S. Andrews, 1984: Wave climate synthesis worldwide. *Proc. Int. Symp. on Wave and Wind Climate Worldwide*. Royal Inst. Naval Architects, London, 1–11.
- da Silva, A.M., C.C. Young, and S. Levitus, 1995: Toward a revised Beaufort equivalent scale. *Proceedings of the International COADS Winds Workshop* (Kiel, Germany, 1994). Ed. H. Diaz and H.-J. Isemer, NOAA-ERL, IFM (Kiel), 270–286.
- Gulev, S.K., 1996: *SHIPMET – pilot questionnaire for VOS officers*. Release 1. P.P. Shirshov Inst. of Oceanol., RAS, Internal note, Moscow, 43 pp.
- Gulev, S.K., 1999: Comparison of COADS Release 1a winds with instrumental measurements in the north-west Atlantic. *J. Atmos. Oceanic Technology*, 16, 133–145
- Gulev, S.K. and L. Hasse, 1998: North Atlantic wind waves and wind stress from voluntary observing data. *J. Phys. Oceanogr.*, 28, 1107–1130.
- Gulev, S.K., D. Cotton and A. Sterl, 1998: Intercomparison of the North Atlantic wave climatology from voluntary observing ships, satellite data and modelling. *Physics and Chemistry of the Earth*, 23, n.7/8, 587–592.

- Gulev, S.K. and L. Hasse, 1999: Changes of wind waves in the North Atlantic over the last 30 years. *Int. J. Climatol.*, 19, 1091–1018.
- Gulev, S.K., V. Grigorieva and A. Sterl, 2001: Global climatology of wind waves from voluntary observing ship data. *J. Geophys. Res.*, submitted.
- Hogben, N., 1988: Experience from compilation of global wave statistics. *Ocean Engng.*, 15, 1–31.
- Hogben, N. and Lumb, 1967: *Ocean Wave Statistics*. The Ministry of Technology, HMSO, London, 263 pp.
- Hogben, N., N.M.C. Dacunha and G.F. Oliver, 1986: *Global Wave Statistics*. Unwin Brothers, London, 661 pp.
- Houmb, O.G., K. Mo and T. Overvik, 1978: *Reliability Tests of Visual Wave Data and Estimation of Extreme Sea States*. Division of port and ocean engineering. Univ. Trondheim, Norwegian Institute of Technology, Report No. 5, 28 pp with figures and tables.
- Isemer, H.-J. and L. Hasse, 1991: The scientific Beaufort equivalent scale: Effects on wind statistics and climatological air-sea flux estimates in the North Atlantic Ocean. *J. Climate*, 4, 819–836.
- Kaufeld, L., 1981: The development of a new Beaufort equivalent scale. *Meteor Rundsch.*, 34, 17–23.
- Kent, E.C. and P.K. Taylor, 1997: Choice of a Beaufort equivalent scale. *J. Atmos. Oceanic Tech.*, 14, 228–242.
- Kent, E.C., P.K. Taylor, B.S. Truscott and J.S. Hopkins, 1993: The accuracy of voluntary observing ships' meteorological observations—results of the VSOP-NA. *J. Atmos. oceanic Tech.*, 14, 591–608.
- Kent, E.C., P.G. Challenor and P.K. Taylor, 1999: A statistical determination of the random observational errors present in voluntary observing ships meteorological reports. *J. Atmos. Oceanic Tech.*, 16, 905–914.
- Laing, A.K., 1985: An assessment of wave observations from ships in Southern Ocean. *J. Clim. Appl. Meteor.*, 24, 481–494.
- Lindau, R., 1995: A new Beaufort equivalent scale. *Proceedings of the International COADS Winds Workshop* (Kiel, Germany, 1994). Ed. H. Diaz and H.-J. Isemer, NOAA-ERL, IFM (Kiel), 232–252.
- Ochi, M.K., 1978: Wave statistics for the design of ships and offshore structures. Proc. SNAME, November 1978, New York, N.Y.
- Smith, S.R., M. Bourassa and R.J. Sharp, 1999: Establishing more truth in true winds. *J. Atmos. Oceanic Tech.*, 16, 939–952.
- Wilkerson, J.C. and M.D. Earle, 1990: A study of differences between environmental reports by ships in the voluntary observing program and measurements from NOAA Buoys. *J. Geophys. Res.*, 95, 3373–3385.
- WMO, 1973: Commission for Marine Meteorology. *Abridged Final Report of the Sixth Session*. Geneva, Switzerland, 117 pp.
- Woodruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley, 1998: COADS Release 2 data and Metadata enhancements for improvements of marine surface flux fields. *Physics and Chemistry of the Earth*, 23, 7/8, 517–526.
- Yelland, M., B.I. Moat, P.K. Taylor, R.W. Pascal, J. Hutchings and V.C. Cornell, 1998: Wind stress measurements from the open ocean corrected for airflow distortion by ship. *J. Phys. Oceanogr.*, 28, 1511–1526.

APPENDIX

SAMPLE QUESTIONS FROM THE SHIPMET QUESTIONNAIRE

1. Visual Beaufort estimates
 - 1.1 Do you use tables with the qualitative description of the Beaufort numbers? yes, always / yes, episodically / yes, but the “reduced table” based on the sea state only / no
 - 1.2 If NO, can you reproduce the table?
 - 1.3 What is the dominant factor for the determination of the Beaufort number? sea state / ship behaviour / combination of factors / other (specify)
 - 1.4 What is the dominant factor for the estimation of wind direction in the case of visual estimates of winds? wave direction / combination of factors

2. True wind correction
 - 2.1 Do you know about the necessity to apply true wind correction to the winds measured by anemometers? yes / no
 - 2.2 If YES, do you do this correction? yes/no
 - 2.3 If YES, describe briefly your actions needed to recompute relative wind to the true wind.

3. Wave height determination
 - 3.1 When estimating wave height, do you go out of the bridge? yes / no / episodically
 - 3.2 If YES, or EPISODICALLY, do you count the parcel of 10 waves to estimate wave height? yes / no
 - 3.3 If YES, do you use the plane to estimate wave height? yes / no
 - 3.4 When estimating wave height onboard the ship equipped with an anemometer, do you take into account the measured wind speed? yes, always / yes, during night time / no

4. Determination of wave periods
 - 4.1 When estimating wave period, do you go out of the bridge? yes / no / episodically
 - 4.2 If YES or EPISODICALLY, do you use a watch to estimate the period? yes / no
 - 4.3 If you do not go out, do you use any tabulated relationships between wave height, wind speed and wave periods? yes / no

5. Determination of wave direction

When estimating wave direction onboard a ship equipped with anemometer, do you use wind direction measured wind speed, for your estimate? yes / no

EVALUATION OF NCEP REANALYSIS SURFACE MARINE WIND FIELDS FOR OCEAN WAVE HINDCASTS

Vincent J. Cardone and Andrew T. Cox¹, Val R. Swail²

ABSTRACT The National Center for Environmental Prediction (NCEP) Reanalysis (NRA) surface marine wind fields are evaluated as the forcing of a third-generation ocean wave model adapted to the North Atlantic (NA) Ocean on a high resolution grid. This evaluation is part of a larger study to produce a high-quality, homogeneous, long-term wind and wave database for assessment of trend and variability in the wave climate of the NA.

It is found that while NRA wind fields appear to be a significant improvement over operational wind fields, if for no other reason than they are more homogeneous over time than real time products, they still suffer from poor resolution of areas of high winds in extratropical storms and lack of resolution of most tropical systems. It is shown that the NRA wind fields may be improved by re-assimilation of measured wind data in a kinematic analysis approach, but only after the limitations of each data source are considered to reduce bias associated with variable measurement height and averaging interval and to recognize limitations of dynamic range, especially for remotely sensed wind speed.

1. HISTORICAL PERSPECTIVE One of the products of the NCEP/NCAR Reanalysis project (henceforth NRA, Kalnay *et al.*, 1996) is a description of the global marine surface wind field on synoptic time (6-hourly) and space (roughly 2-degrees) scales. The NRA is an appealing and convenient source of forcing for ocean response modelling but it is fair to ask whether it is sufficiently accurate and free of bias for such purposes. The principal purpose of this paper is to describe our evaluation of the NRA winds through analysis of the errors in a simulation of the wave climate (Swail and Cox, 1999) made when NRA winds are used to force a proven spectral ocean wave model adapted to the North Atlantic Ocean. Wave modelling has been shown to be particularly well suited to the evaluation of marine wind fields (Cardone *et al.*, 1995). Before presenting our evaluation, however, it is interesting to review the more traditional approaches to specification of marine surface wind fields, because we will find that some elements of those methods still have a role to play in the derivation of wind fields of maximum accuracy from NRA products.

At the time when the first author of this paper first became interested in specification of marine surface wind fields (circa 1965) and as recently as the late 1970s, basically only one data source and two approaches were available (the analyst's life was therefore quite simple though the results were not always rewarding!). The data source consisted of ships' synoptic weather reports, mainly from transient merchant vessels supplemented in the northern hemisphere (NH) by a few stationary ocean station vessels. The two approaches consisted of: (1) derivation of winds from fields of sea level pressure and other Marine Planetary Boundary Layer (MPBL) variables, themselves derived from ships' observations of sea level pressure, air temperature and sea temperature, using simple empirical rules or fairly complex MPBL models; (2) kinematic analysis of ship wind observations.

1 Oceanweather Inc. – Cos Cob, CT

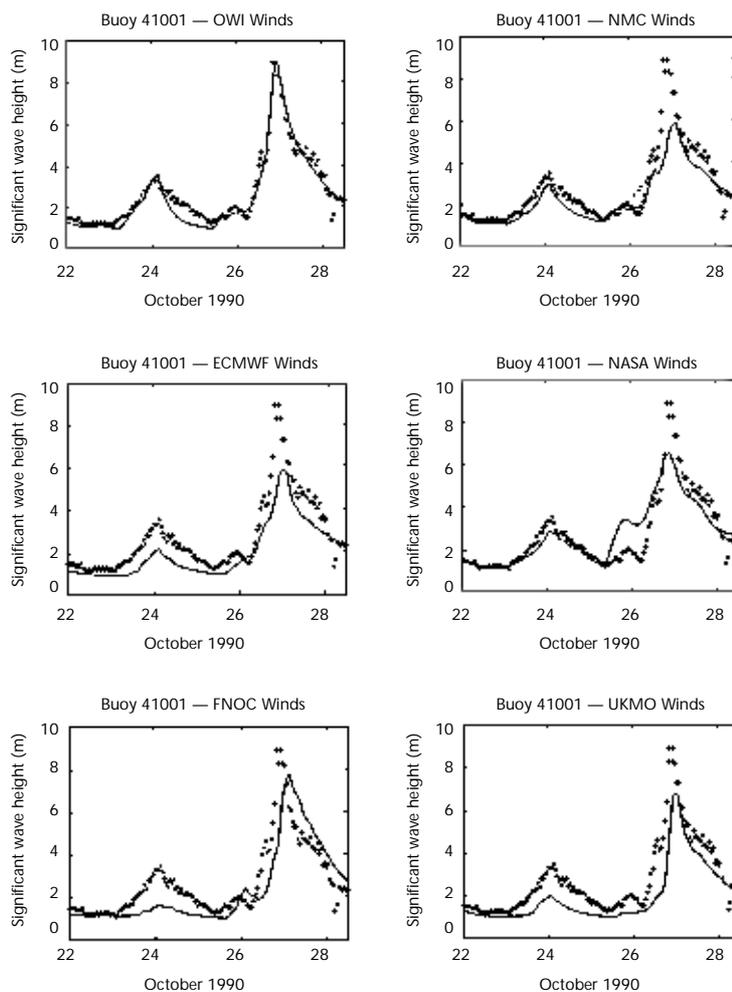
2 Environment Canada – Downsview, Ontario

The error characteristics of wind fields hindcast for a sample of NH extratropical storms by the alternative approaches were explored by Cardone *et al.* (1980). It was found that wind fields produced by application of an MPBL model to either hand analysed or objectively analysed pressure fields tended to be negatively biased (typical bias of -1.5 to -2.0 m/s) with the bias contributed to mainly by the higher wind speeds. It was suggested that the bias in wind speed was, therefore, better expressed as a percentage (10-15 per cent) reduction. Cardone (1991) summarized a number of similar evaluations of MPBL-derived winds conducted through the 1970s and 1980s (Overland and Gemmill, 1978; Gemmill *et al.*, 1988; Dobson and Chaykovsky, 1991) and concluded that random wind speed errors in MPBL-derived wind speeds derived from carefully reanalysed pressure fields are about 3 m/s (rms) about a mean negative bias of about -0.5 m/s when the MPBL winds were compared to NOAA buoy winds or GEOSAT altimeter winds and -1.6 m/s when evaluated against ship winds after the ship winds were adjusted for measurement height and type and stability (Cardone *et al.*, 1990).

With the widespread implementation of multi-level primitive-equation numerical weather prediction (NWP) models in the 1980s, new sources of marine surface wind data became available as a by-product of the NWP analysis-forecast cycle of the major centres such as the Canadian Meteorological Center (CMC), the US Navy's Fleet Numerical Meteorology and Oceanography Center (FNMOC, formerly FNOC) NCEP, the United Kingdom Met Office (UKMO), the European Centre for Medium Range Weather Forecasting (ECMWF) and the Goddard Space Flight Centre of the National Aeronautics and Space Administration (NASA). At some centres, the 10 m level was explicitly resolved in the NWP model and the initial analysis benefited from the assimilation of winds measured from ships, moored buoys and, by the early 1990's, from satellite sources of wind data. Unfortunately, the data assimilation eliminates from consideration marine data that could have otherwise served as independent data to evaluate the accuracy of the NWP wind fields.

The Surface Wave Dynamics Experiment (SWADE) conducted off the US East Coast in 1990 (Weller *et al.*, 1991) provided an opportunity to develop a surface wind fields database with much better coverage from measured data than had been previously possible. This is because the SWADE wind fields database itself incorporated a second database of storms which incorporated high quality surface wind measurements from buoys. These buoys were sufficiently well distributed to ensure, for the first time in such a database, the avoidance of gaps typically found in similar data sets for open ocean areas. Initially, it was thought that the availability of the SWADE enhanced database in real time to the NWP centres' objective analysis and data assimilation schemes would then necessarily lead to high-quality wind fields. Unfortunately, when those NWP centre wind fields for SWADE IOP-1 (an 11-day period centered on the development of an intense US East Coast cyclone of 23-31 October 1990) were used to drive the WAM-4 wave model adapted to the SWADE area at high resolution, errors in modelled sea states were found to be intolerably large (Graber *et al.*, 1991). However, when the same database was subjected to an intensive manual analysis using classical kinematic analysis and the resulting wind fields were used to drive the WAM-4 wave model, wave hindcasts of unprecedented skill were found (Cardone *et al.*, 1995). Figure 1 compares hindcast and buoy measurements of significant wave height (HS) at NOAA buoy 41001 moored east of Cape Hatteras, from WAM-4 hindcasts driven by the various NWP analysis wind fields and by the kinematically-derived winds (labelled OWI in the figure). The maximum wind speed and HSL observed in the SWADE array during IOP-1 were about 25 m/s and 9 m respectively. Therefore, at least for this regime of moderate wind forcing, the SWADE study demonstrated that wind field errors could be reduced to very low levels through an available, though tedious, analysis method, namely kinematic analysis, provided that accurate surface wind measurements are available at a data density roughly comparable to that achieved in the buoy array off the US East Coast during SWADE.

Figure 1—Comparison of WAM-4 hindcasts (solid line) of significant wave height and buoy measurement at buoy 41001 (East of Hatteras) in SWADE IOP-1 (Cardone *et al.*, 1995).



The most significant wind field features found in the storms modelled in SWADE and in other recent storm studies (Cardone *et al.*, 1996), in terms of generation of storm peak sea states, were relatively small scale, rapidly propagating surface wind maxima or 'jet streaks' (typical jet core widths of 200 km or less) which by virtue of their spatial and temporal coherency provide a dynamic fetch to couple very effectively to the surface wave field. The propagation speeds of these jet streaks, typically 15-20 m/s, do not necessarily match the speed of the parent cyclone centre. The most extreme sea states in storms containing jet streaks are normally observed at buoys directly in the path of the core of jet streaks. Validation of wave hindcasts, therefore, provides a sensitive measure of skill in wind fields.

Unfortunately, the SWADE hindcast study also shows that the objective analysis systems used at major NWP centres did not realize the full potential of the enhanced buoy array for surface wind analysis, and did not resolve accurately the small scale rapidly evolving features. The wind fields provided by objective analysis at such centres have been used to drive wave models to provide hindcast time series for climate assessment, such as the US Navy's 20-year Spectral Ocean Wave Model (SOWM) and Norwegian 35-year Waves in Norwegian Coast-Hindcasting (WINCH) data sets. It is not surprising, therefore, that such data sets, though useful, are subject to both bias and scatter.

It was found that the deficiencies of the operational NWP wind fields observed during SWADE could not be attributed to model grid spacing or the size of the time step. This was shown by Graber *et al.* (1995), who used the SWADE kinematic winds in IOP-1 to systematically investigate the effect of degrading the spatial and temporal resolution of the reference SWADE wind fields on the accuracy of the hindcasts. The effect of degrading the temporal and spatial resolution was investigated through the validation of alternative SWADE hindcasts with the

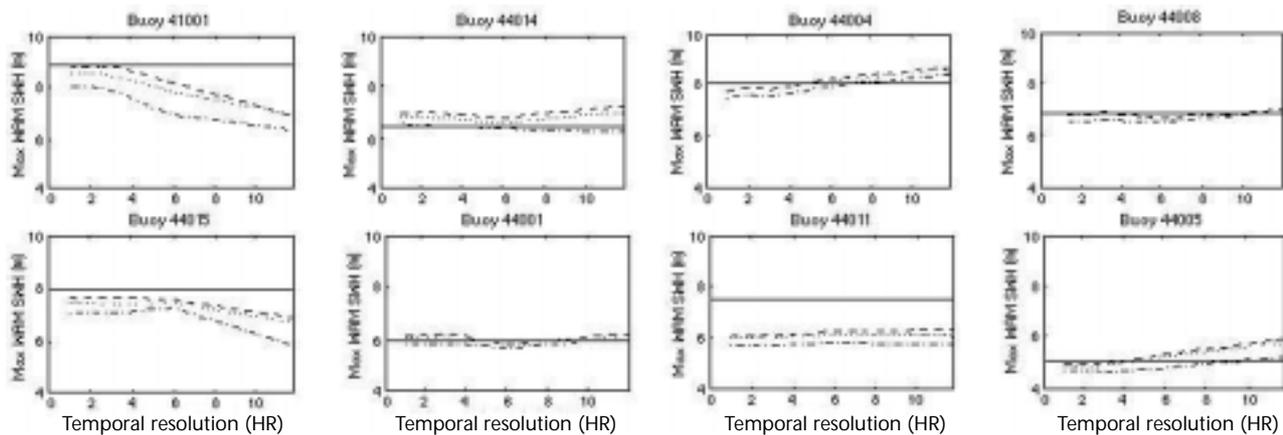


Figure 2—Hindcast of peak event HS relative to that measured (solid horizontal line) as a function of indicated temporal resolution for three indicated spatial resolutions: 0.5° (OWI) (dashed line); 1° (dotted line); 1.5° (dashed-dot line) (from Graber et al., 1995). Buoy locations: 41001 – 34°55.5'N, 72°57.1'W; 44014 – 36°35.0'N, 74°50.0'W; 44004 – 38°32.2'N, 70°42.3'W.

same wave model used for the reference SWADE hindcasts (Figure 2). The reference winds were specified on a 0.5 degree grid at hourly intervals. It was found that at the buoy directly in the path of the jet streak (41001), wind fields with a 0.5 degree spatial resolution and 3-hour time step were required for accurate specification of the peak HS. At buoys moored north of the storm track (e.g. 44014, 44004) in a nearly linear slowly evolving wind field, even 12-hour sampling and 1.5 degree spacing did not degrade specification of the local HS storm peaks. Well outside the SWADE array (e.g. 44011), where even the reference winds were not very accurate, the storm peak HS was uniformly underestimated for all resolutions simulated. Within the SWADE array, however, it was found that the errors in the hindcasts of storm peaks resulting from the actual operational wind fields (Figure 1) were always significantly greater than the errors for the particular cases simulated which matched the spatial and temporal resolution of the operational centre winds, thereby confirming the presence of additional error sources in the NWP centre wind fields.

The deficiencies exhibited in the NWP winds during SWADE (conducted in 1990) may not be indicative of the accuracy of NWP winds later in the 1990s and at the present time because analysis and data assimilation methods have undergone almost continuous refinement. Also, in some areas the volume of high-quality in situ measured data has increased, particularly off the east and west coasts of North America and offshore Western Europe. In addition, remotely-sensed marine wind data became available on a global basis in the early 1990s from passive and active microwave sensors. However, there remain questions of accuracy and bias, especially at wind speeds above about 15 m/s with all types of in situ and remotely-sensed marine wind observations, which have not been fully resolved and will be discussed further below.

Nevertheless, the NRA provides a new and convenient database and indications are that the NRA marine wind fields will be widely used for ocean response modelling. Section 2 of this paper gives our evaluation of the alternative files of marine winds available within the NRA database. In section 3, we describe the remaining deficiencies of even the best of the NRA wind fields evaluated, and we describe how they were resolved at least in part by applying kinematic analysis and manual intervention to the NRA database with a graphical user interface. Section 4 elaborates on the above-noted observational error issues and implications of same on the development of an optimum marine wind analysis system. Conclusions are given in section 5.

2. EVALUATION OF THREE ALTERNATIVE NRA WIND FIELDS

A. EVALUATION METHODOLOGY

In the evaluation phase of our study we compared three alternative NRA sources of marine boundary layer winds: (1) the 1000 mb wind fields on the 2.5° latitude-longitude grid; (2) the lowest sigma level (0.995) wind fields on the 2.5° latitude-longitude grid; and (3) the 10 m surface wind fields on the so-called Gaussian grid. A fourth method is available, namely the application of a diagnostic MPBL model applied to NRA pressure fields and other MPBL variables, but this was not utilized because it was expected a priori that the boundary layer

formulation within the NRA NWP model provided a physically more correct representation of the boundary layer than that provided by any steady state diagnostic MPBL. However, we have recently had cause to reconsider the validity of this assumption.

Eight months were selected from the available period (1979-1995) for the wind field evaluation. Months 8103 and 8301 were chosen for having the highest and lowest values, respectively, of the mean North Atlantic atmospheric zonal circulation index described by Kushnir (1994). The months 9110, 9303 and 9504 each contained extreme western North Atlantic storms hindcast in recent studies (Cardone *et al.*, 1996; Swail *et al.*, 1995), while 9509 was chosen as a hurricane-dominated month. The remaining months (7906, 8808) were added to provide a more even representation over time of the part of the NRA available (1979-1995).

Wind fields for each month were interpolated from the NRA source grids onto a 0.625° by 0.833° latitude-longitude wave model grid covering the North Atlantic Ocean using the IOKA (Interactive Objective Kinematic Analysis) algorithm (Cox *et al.*, 1995) and then time interpolated linearly from a six-hour time step to a one-hour time step. Oceanweather's third generation (OWI3G) wave model (Khandekar *et al.*, 1994) was used in deep water mode for all hindcasts. Wave and interpolated wind results were then compared (time series, scatter plots and statistics) to all available deep-water buoys (US, Canadian and European), offshore North Sea platforms, US C-MAN (Coastal Marine Automated Network) and ERS-1/2 altimeter and scatterometer measurements. All measured winds were adjusted for height and stratification to a 10 m reference height and neutral stability (Cardone *et al.*, 1990), while hourly wind and wave measurements were smoothed over ±1 hour using equal weights (1,1,1). ERS-1/2 altimeter and scatterometer measurements were extracted from Ifremer's CD-ROM set using the recommended quality controls, temporally binned within a 6-hour window, and then spatially binned onto the wave model grid every 6 hours.

B. RESULTS

The results of the statistical comparisons of the three sets of NCEP winds and the modelled waves with all buoys, platforms and C-MAN stations on the western and eastern Atlantic continental margins, and with ERS-1/2 satellite altimeter winds and waves, are summarized in Tables 1 and 2. Table 1 shows statistical comparisons for March 1993 – the other evaluation months showed generally comparable results. While the statistics for correlation coefficient and scatter index for winds were similar among all wind fields, there were clear advantages in bias, scatter index, and ratio for the waves produced by the surface wind fields. From these and other properties of the hindcast results studied it was concluded that there was no advantage in selecting the 1000 mb winds; therefore the 1000 mb winds were dropped from further consideration. Table 2 shows the bias and scatter index comparisons for all eight evaluation months versus the in situ measurements and for the three months for which ERS-1/2 altimeter data were available. Table 2 shows that the best wind field was the Gaussian grid 10 m surface wind field. The bias for these winds was generally lower for both winds and resulting waves; the scatter indices for winds were similar for both data sets, although the independent satellite comparisons always favoured the surface winds. The scatter index for waves hindcast from the surface winds was always superior.

Table 1—Comparison wave summary statistics (wind statistics in brackets) for March 1993 for NRA 10 m surface, sigma and 1000 mb input wind fields (Scatter index is the ratio of the standard deviation (SD) of the difference between hindcast (H) and measurement (M) and the mean of the measurements; ratio is percentage of points above/below the 1:1 line on a scatter plot (0.5 is ideal) of the paired hindcast-measured data).

Wind field	Bias (H-M) m (m/s)	RMS error m (m/s)	Scatter index	Ratio	Corr. coeff.
Surface	0.0 (0.0)	0.98 (2.74)	0.44 (0.35)	0.52 (0.51)	0.83 (0.82)
Sigma	1.0 (2.0)	1.65 (3.36)	0.60 (0.34)	0.85 (0.79)	0.81 (0.83)
1000 mb	0.6 (1.2)	1.36 (3.13)	0.54 (0.36)	0.76 (0.68)	0.78 (0.80)

Table 2—Comparison of wind and wave bias and scatter index values by month for NRA sigma and 10 m surface winds (bold italics show closer agreement with measurements).

	Wind speed				Significant wave height			
	Bias (H-M)		Scatter index		Bias (H-M)		Scatter index	
	Surface	Sigma	Surface	Sigma	Surface	Sigma	Surface	Sigma
<i>Vs. in situ</i>								
7906	-0.4	1.1	0.44	0.45	0.0	0.4	0.56	0.60
8103	-0.4	1.2	0.27	0.27	-0.4	0.4	0.27	0.33
8301	0.1	0.8	0.27	0.23	-0.3	0.1	0.27	0.29
8808	0.2	2.2	0.48	0.50	-0.2	0.4	0.51	0.61
9110	-0.5	1.4	0.39	0.37	-0.4	0.4	0.61	0.72
9303	0.0	2.0	0.35	0.34	0.0	1.0	0.44	0.60
9504	-1.2	0.3	0.38	0.35	-0.2	0.4	0.44	0.46
9509	-1.2	0.5	0.36	0.32	-0.4	0.2	0.36	0.43
<i>Vs. altimeter</i>								
9110	0.1	1.4	0.30	0.34	0.0	0.8	0.34	0.54
9303	0.6	2.2	0.33	0.37	0.1	1.2	0.45	0.63
9504	0.2	1.6	0.30	0.33	0.1	0.9	0.41	0.56

3. DEFICIENCIES AND CORRECTIONS OF NRA WINDS

A. DEFICIENCIES

While the NCEP surface wind fields produce the least biased and most skillful wave hindcasts overall, the scatter index values were much higher (hence less skill) than found in hindcast studies of continuous periods (Cardone *et al.*, 1995) or storms (Cardone *et al.*, 1996) where kinematically reanalysed wind fields were used to drive the wave model. The hindcasts were also found to systematically underestimate storm peaks. For example, Figure 3 (left-hand side) shows the effect on the hindcast of the poor NRA representation of the winds at a buoy off the US East Coast during SWADE IOP-1. It was also found that tropical storms were poorly resolved in the NRA wind fields as shown in Figure 4 for Hurricane Emily (September 1993).

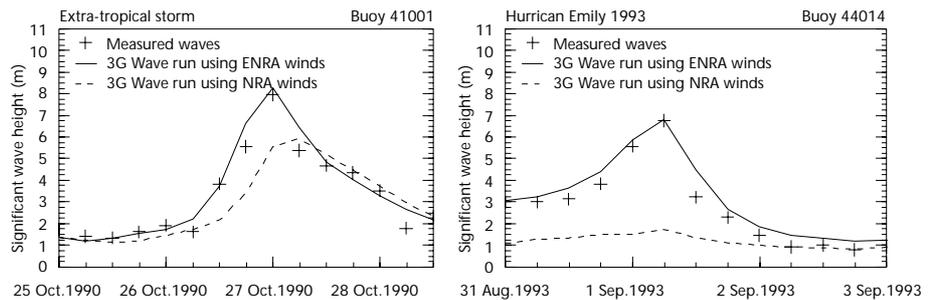
Table 3 shows the results of hindcasts using the NRA 10 m surface wind fields for four of the eight months selected (those months for which ERS 1/2 altimeter data were available); results for the other four months indicated similar results (not shown). The hindcasts were compared to measurements from buoys moored in deep water off the US and Canadian East Coasts and off Northwest Europe and to the satellite data over the whole of the model domain. With respect to the buoy comparisons overall, the HS SI of 26 per cent indicates less skill in these hindcasts than provided by kinematically reanalysed wind fields. On the other hand, this skill is equal to, or better than, the best of the SWADE hindcasts driven by the wind fields from the operational centres (Cardone *et al.*, 1995). The HS bias of 3 cm is satisfyingly small.

The altimeter comparisons in Table 3 provide evaluation of the hindcast over the whole of the NA. These comparisons exhibit a mean difference of 18 cm and HS SI of 23 per cent. Interestingly, these comparisons suggest that the skill

Table 3—Validation of North Atlantic Ocean continuous hindcasts of indicated months with OWI-3G driven by NRA 10 m surface winds, against buoy and ERS-1 altimeter wave measurements.

Year/ Month	Variable	Num	All Buoys			ERS-1 Altimeter			
			Bias	rms	S.I.	Num	Bias	rms	S.I.
9110	WS (m/s)	882	0.12	2.96	0.34	16,808	0.34	2.13	0.29
	HS (m)	758	0.01	0.77	0.24				
9303	WS (m/s)	868	-0.28	2.31	0.24	17,517	0.43	2.19	0.26
	HS (m)	871	-0.07	0.73	0.24				
9504	WS (m/s)	600	-0.15	2.30	0.33	17,693	0.37	1.97	0.27
	HS (m)	720	0.04	0.60	0.26				
9509	WS (m/s)	761	0.36	2.68	0.41	18,081	0.05	2.30	0.35
	HS (m)	834	-0.11	0.62	0.30				
All Months	WS (m/s)	3,111	0.01	2.59	0.33	70,099	0.30	2.15	0.29
	HS (m)	3,183	-0.03	0.68	0.26				

Figure 3— Effect of kinematic analysis on wave hindcast.



indicated by the buoy comparisons is indicative of skill over the whole of the model domain.

Another deficiency in the NRA reanalysis concerns the assimilation of surface marine wind data from the Comprehensive Ocean-Atmosphere Data Set (COADS). The assimilation scheme treated all observations at a 10 m reference level, whereas ship and drilling platform observations may actually range from about 15 m to more than 100 m, and buoy observations are typically taken at about 5 m. Over the 40-year duration of the NCEP reanalysis this may introduce biases similar to those found by Cardone *et al.* (1990) due to the increasing heights of shipboard anemometers and the higher fraction of wind measurements compared to wind estimates. To overcome any potential bias in this project, all surface wind data were reassimilated after first being adjusted to the 10 m reference level (Cardone *et al.*, 1990).

B. ENHANCEMENT OF NRA WINDS

While it has been shown that NRA surface wind fields produce wave hindcasts of good quality, they are evidently susceptible to further improvement to achieve skill comparable to hindcasts driven by kinematically reanalysed wind fields. Of particular concern was the finding that the hindcasts tended to systematically underestimate storm peaks.

Basically, three steps were taken to enhance the NRA winds. First, the NRA wind fields and the wind observations were processed to make them representative of the average effective neutral wind at 10 m height. This was done for the NRA surface winds by computing an equivalent neutral wind using the NRA 2 m surface temperature and sea-surface temperature fields and the algorithm described by Cardone *et al.* (1990). To remove potential biases in the data to be reassimilated into NRA, all wind observations including buoy observations, ship reports (from COADS) and C-MAN stations were also transformed to effective neutral 10 m wind speed taking into account the method of observation, anemometer height and stability. ERS 1/2 scatterometer winds were made available to the analysis only after a meteorologist had the opportunity to filter areas of suspected saturation of wind speed and incorrect wind directions due to obvious failure of the ambiguity removal algorithm.

Second, wind fields for all significant storms were kinematically reanalysed using the IOKA system with the aid of an interactive wind workstation (Cox *et al.*, 1995). The NRA surface wind fields were brought into the wind workstation every six hours in monthly segments for evaluation by a trained marine meteorologist. The interactive hindcast methodology used by the analysts follows similar previous hindcast studies (Cardone *et al.*, 1995, 1996). Particular attention was given to strong extratropical systems and the quality control of surface data. Kinematically analysed winds from previous hindcasts of severe extratropical storms in the north-west Atlantic (Swail *et al.*, 1995) were incorporated into the present analysis on the North Atlantic wave model grid.

Altimeter wave measurements were used in an inverse wave-modelling approach as follows. First of all, a global coarse wave run was made and hindcast wave heights over the North Atlantic Ocean were compared to altimeter wave measurements. The global wave fields were generated using Oceanweather's 1-G wave model (Khandekar *et al.*, 1994) adapted to a 1.25° by 2.5° latitude-longitude grid for the entire globe. NRA surface winds (adjusted to neutral stability) were

Figure 4(a)—NRA surface wind field (unmodified).

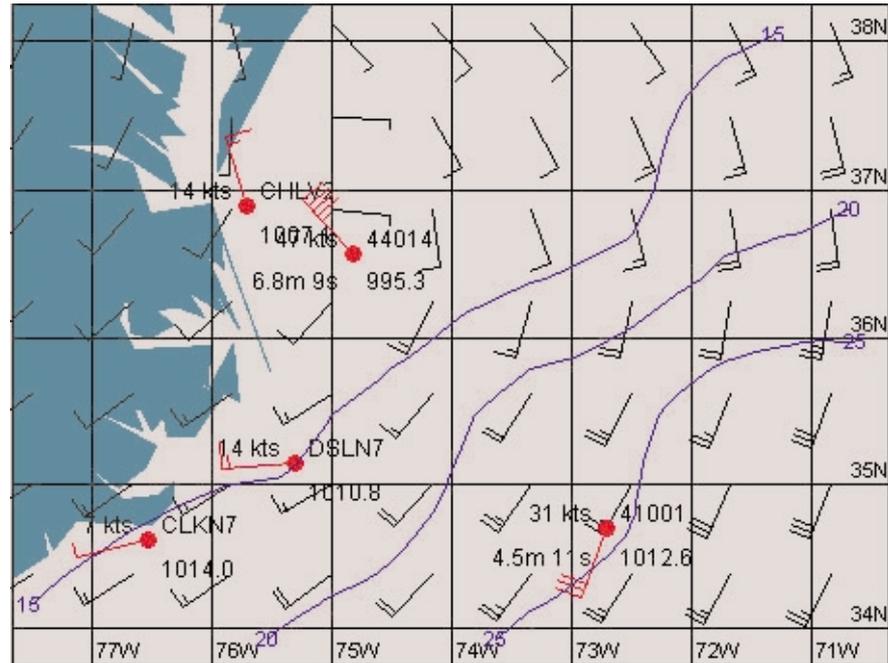
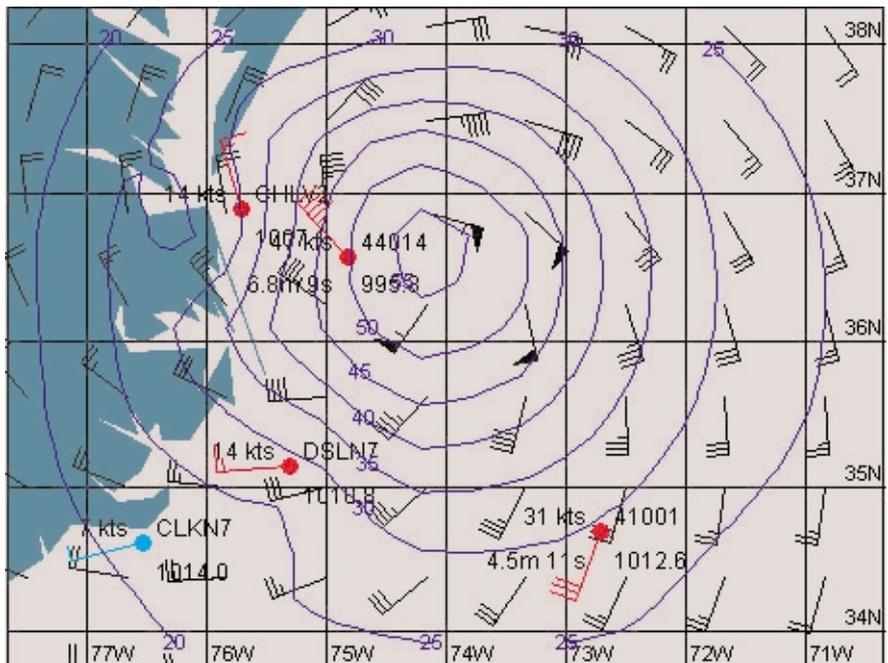


Figure 4(b)—ENRA final wind field with tropical vortex model winds incorporated.



used to drive the global wave model. Areas where the resulting wave fields were deficient, as indicated by the altimeter, were brought to the analyst's attention and the analyst subjectively altered the wind fields in the relevant space-time domains until the output from the 1G wave model agreed better with the altimeter measurements.

Third, high resolution surface wind fields for all tropical cyclones, as specified by a proven tropical cyclone boundary layer model (Cardone *et al.*, 1994; Thompson and Cardone, 1996), were assimilated into the wind fields to provide greater skill and resolution in the resulting wave hindcasts. Track and initial estimates of intensity were taken, with some modification, from the NOAA Tropical Prediction Center's (TPC) HURDAT database. The radius of maximum wind was determined using a pressure profile fit to available surface observations and aircraft reconnaissance data. Surface winds generated from the model were then evaluated against available surface data and aircraft reconnaissance wind observations adjusted to the surface as described by Powell and Black (1989). Model winds

within 240 nautical miles from the centre were then exported on a 0.5° latitude-longitude grid for inclusion and blending using the wind workstation. Approximately 400 tropical cyclones were added to the NRA in this way.

C COMPARISON OF HIGH-FREQUENCY WIND AND WAVE RESULTS

Figure 3 (left-hand side) shows the hindcast made with NRA surface winds at a buoy off the US East Coast during SWADE IOP-1 and the hindcast made after the NRA winds were kinematically enhanced (hereafter ENRA). This case is typical of the improvement in skill of the hindcast overall and the reduction in the underestimation of storm peaks when the NRA surface wind fields were reanalysed.

Figure 4 compares the NRA winds and ENRA winds during Hurricane Emily (September 1993). The improvement is achieved through a combination of interactive kinematic analysis of the wind fields in conjunction with winds generated by a proven tropical cyclone model as described above. The resulting wave comparison at buoy 44014 is shown in Figure 3 (right-hand side).

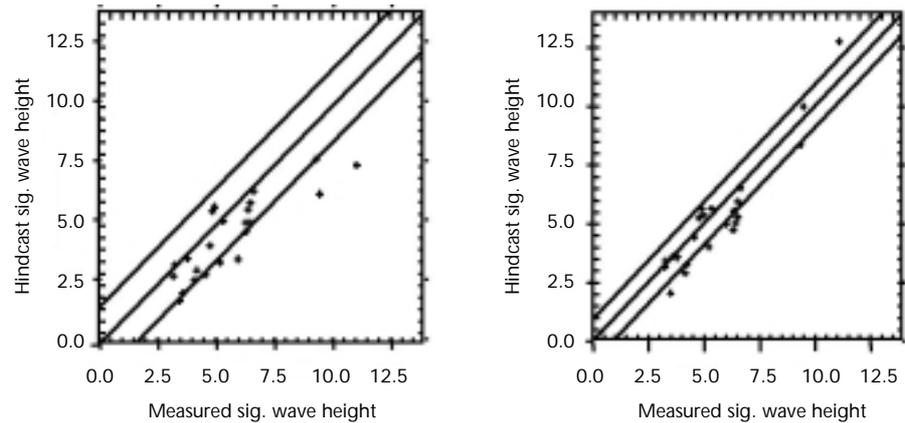
Table 4 shows the validation of the hindcasts against buoy and altimeter data for hindcasts made using the ENRA wind fields for the same four months shown in Table 3. At the buoys there is a significant reduction in the scatter index for wind speed, nearly a factor of two reduction over all buoys, which is to be expected because the buoy winds have been reassimilated at the correct height. The wave height SI is reduced as well but by only about 10 per cent overall. Altimeter wind speeds and wave heights were not assimilated so the altimeter statistics give an independent measure of skill in the hindcasts. By comparing Table 3 and Table 4 it is seen that there is no significant difference in the scatter statistics (i.e. rms and SI) between runs made with NRA and ENRA winds. This result is not surprising since the scatter statistics were dominated by lower sea states, which would not be changed substantially by the IOKA process. However, there is a reduction in the wave height bias overall from 18 to 4 cm. This reduction in bias is contributed to mainly by increased skill in specification of storm generated sea state. Figure 5 shows the comparison of storm peaks greater than 3 m (as measured by the buoy) at buoy 44138 for the four overlapping evaluation and production months. This figure shows a clear reduction in both the bias and scatter when using the ENRA wind fields.

Figure 6 shows the wave model grid-averaged altimeter wave measurements binned every 2 m compared with the matching hindcast waves (within ±3 hours), showing the mean bias for each bin over the four evaluation months. While the buoy comparisons indicate the skill in the hindcasts near the continental margins, the altimeter samples the entire North Atlantic basin more or less even in space and time. It is encouraging, therefore, that wave hindcasts show very good agreement with the altimeter throughout the range of wave heights. The mean in bias in wave height derived from the ENRA winds over the four months is within ± 30 cm, while the NRA analysis had biases of nearly twice that value. Hindcast wave heights of less than 1.5 m show a slight systematic overestimation, which may be attributed to an inherent tendency for the gridded wind and wave fields to fail to resolve small areas of calm winds and seas.

Table 4—Validation of North Atlantic Ocean continuous hindcasts with OWI-3G driven by ENRA winds against buoy and ERS-1 altimeter wind speed and wave height measurements.

Year/ Month	Variable	Num	All Buoys			ERS-1 Altimeter			
			Bias	rms	S.I.	Num	Bias	rms	S.I.
9110	WS (m/s)	882	0.69	2.41	0.26	16,808	0.39	2.19	0.30
	HS (m)	758	0.26	0.76	0.25	16,703	-0.06	0.64	0.25
9303	WS (m/s)	868	0.19	1.04	0.11	17,517	0.46	2.26	0.27
	HS (m)	871	0.09	0.68	0.22	16,972	0.05	0.63	0.21
9504	WS (m/s)	600	-0.05	1.85	0.09	17,693	0.38	1.94	0.27
	HS (m)	720	0.11	0.55	0.22	17,551	0.07	0.53	0.22
9509	WS (m/s)	761	0.42	1.28	0.19	18,081	0.13	2.20	0.34
	HS (m)	834	0.09	0.53	0.26	18,059	-0.23	0.60	0.24
All Months	WS (m/s)	3,111	0.40	1.73	0.17	70,099	0.39	2.15	0.29
	HS (m)	3,183	0.13	0.64	0.24	69,285	-0.04	0.60	0.23

Figure 5—Comparison of peak-to-peak wave height using NRA (left) and ENRA (right) wind fields to drive 3-G wave model for four months.



Total points: 23
 Mean X: 5.666
 Mean Y: 4.569
 Mean diff: -1.097
 Root mean square: 1.524
 Standard dev.: 1.058
 Scatter index: 0.187
 Ratio: 0.130
 Correlation coeff: 0.852

Total points: 23
 Mean X: 5.666
 Mean Y: 5.280
 Mean diff: -0.386
 Root mean square: 0.919
 Standard dev.: 0.834
 Scatter index: 0.147
 Ratio: 0.304
 Correlation coeff: 0.938

Given the emphasis in the ENRA on specification of storm wind fields, it is interesting to compare the production wave hindcasts with wave hindcasts made with the NRA surface winds during storm peaks. In Figure 7, TOPEX altimeter wave measurements along a swath are compared in an extratropical storm off the east coast of Canada. The improvements resulting from the ENRA winds are clearly evident along the TOPEX track; the figure shows that not only does the ENRA capture more accurately the peak of the storm but also the spatial characteristics of the wave field.

D. REPRESENTATION OF WIND AND WAVE CLIMATOLOGY

Comparisons of the ENRA wind and wave climatology at six buoys and platforms selected to give a comprehensive geographical coverage over the North Atlantic Ocean, well away from the coast, in deep water, were carried out for the period 1990–1995.

The hindcast and measured wind speed climatologies are not independent since all the wind data used contributed heavily to the data assimilation scheme in the NCEP reanalysis, and again in the kinematic reanalysis. Nevertheless, it is

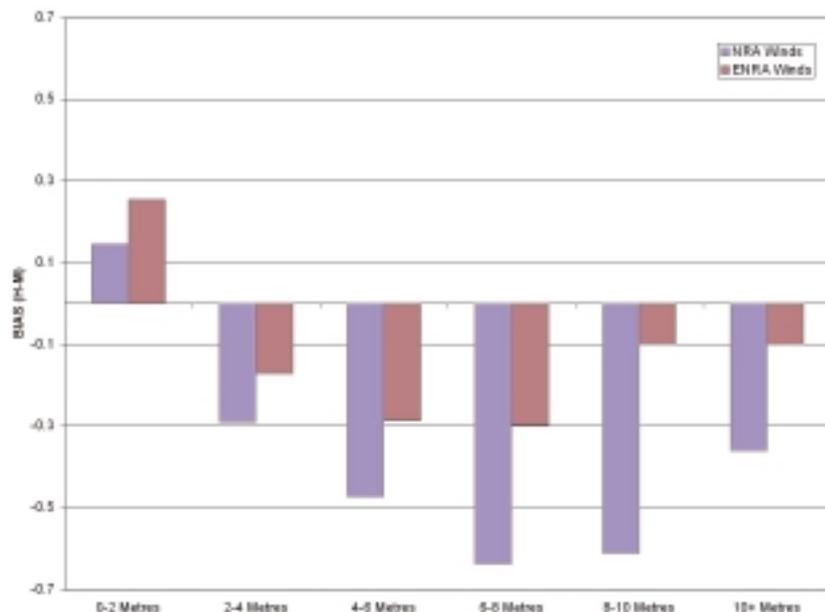


Figure 6—Comparison of bias statistics (H-M) vs. binned ERS altimeter measurements.

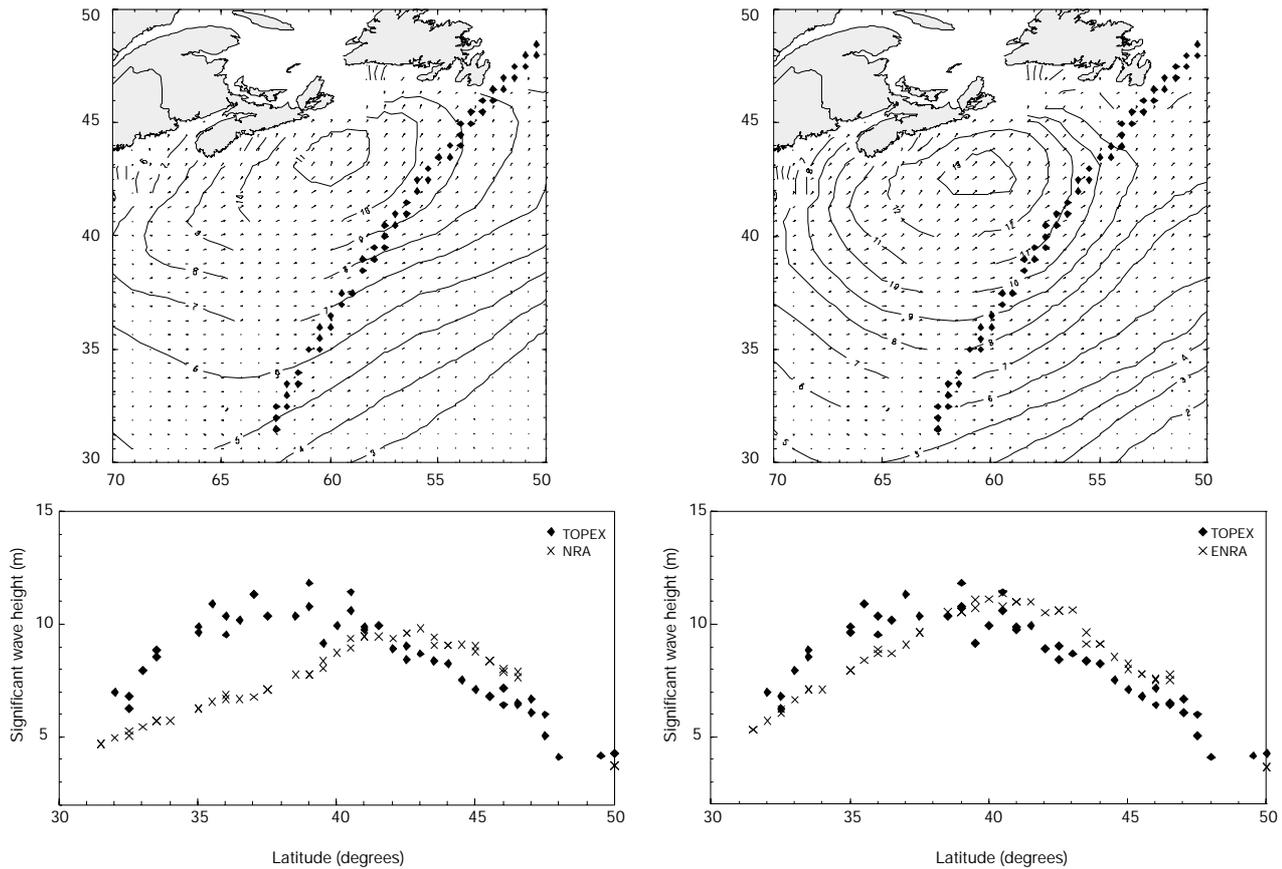


Figure 7—Hindcast and measured HS along indicated TOPEX track for NRA (lower left) and ENRA (lower right) winds. Upper panels show contours of hindcast HS (m) from NRA (left) and ENRA (right) winds.

useful to compare the two data sets to verify that the various adjustments for elevation and interpolation onto the wave model grid have not compromised the hindcast data set.

Figure 8 shows quantile-quantile (Q-Q) plots for ENRA hindcast wind speed versus measured wind speed for each of the six selected sites. Q-Q plots illustrate the comparison of the full frequency distributions, particularly in the extreme tails. These plots show very good agreement across the entire frequency distribution. There is a tendency for the ENRA winds to be slightly higher at Canadian buoys (44137), particularly for the highest wind speeds, possibly related to the vector averaging of the buoy wind samples as opposed to scalar averages elsewhere (see section 4 below). At the platform (LF3J) the model is noticeably higher than the measurements for the low end of the wind speed distribution.

Figure 9 shows Q-Q plots for model versus measured wave height for each of the six selected sites. These plots show very good agreement across the entire frequency distribution. There is a slight tendency for the model to overestimate the wave height compared to the measurements for low values of sea state. The model also is consistently higher at the platform, although the differences are negligible for the few highest observations. The effect of the Halloween storm (October 1991) is clearly seen at 44137 and 44138, where the peak measured waves clearly exceed the hindcast values. The Gullfaks platform in the North Sea (LF3J) does not strictly satisfy the conditions of deep-water open ocean; a model of much higher grid resolution would be required to properly describe the propagation of wave energy from the North Atlantic Ocean into the North Sea through the British Isles.

4. DISCUSSION OF VARIOUS MARINE WIND DATA SOURCES

As noted in section 1, there has been a tremendous increase in the volume of instrumentally measured winds within the last two decades as acquired from moored buoys, automatic coastal weather stations, fixed platforms and satellites. The USA alone maintains over 160 instrumented sites. The Canadian Government supports more than 40 buoys. In the North Sea, Norwegian Sea and western approaches to Europe there are over 50 sites. In this section, we discuss

Figure 8—Quantile-quantile plots of wind speed for selected measurement locations, based on ENRA-driven hindcasts.*

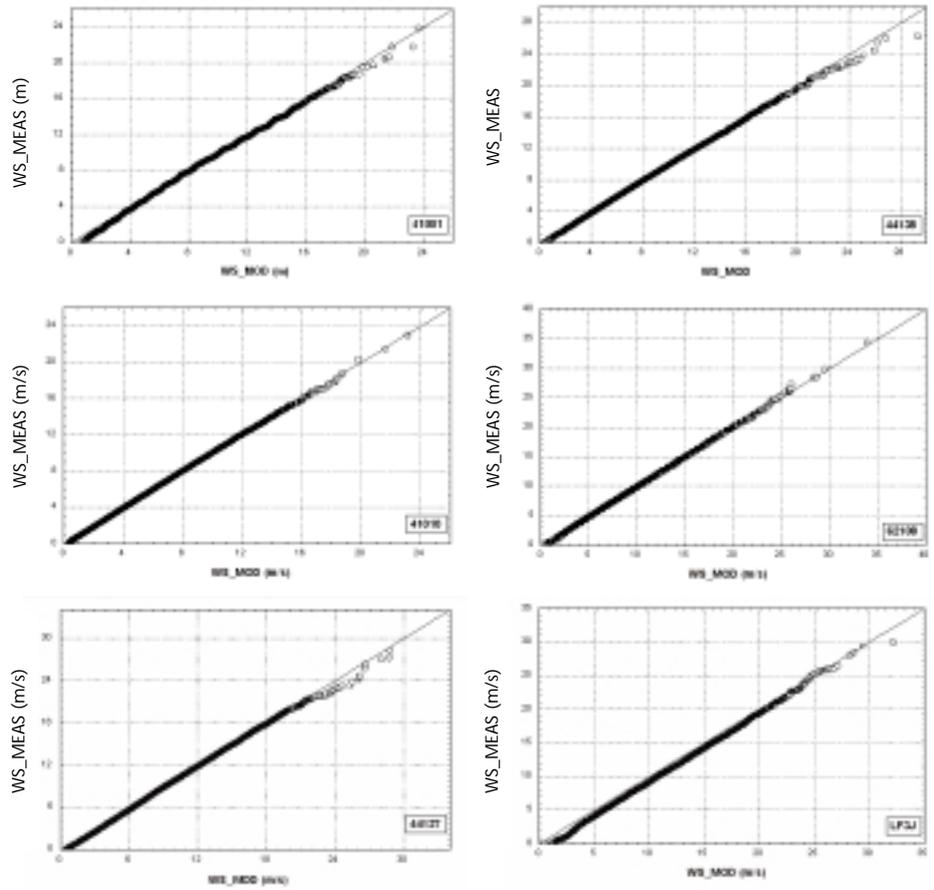
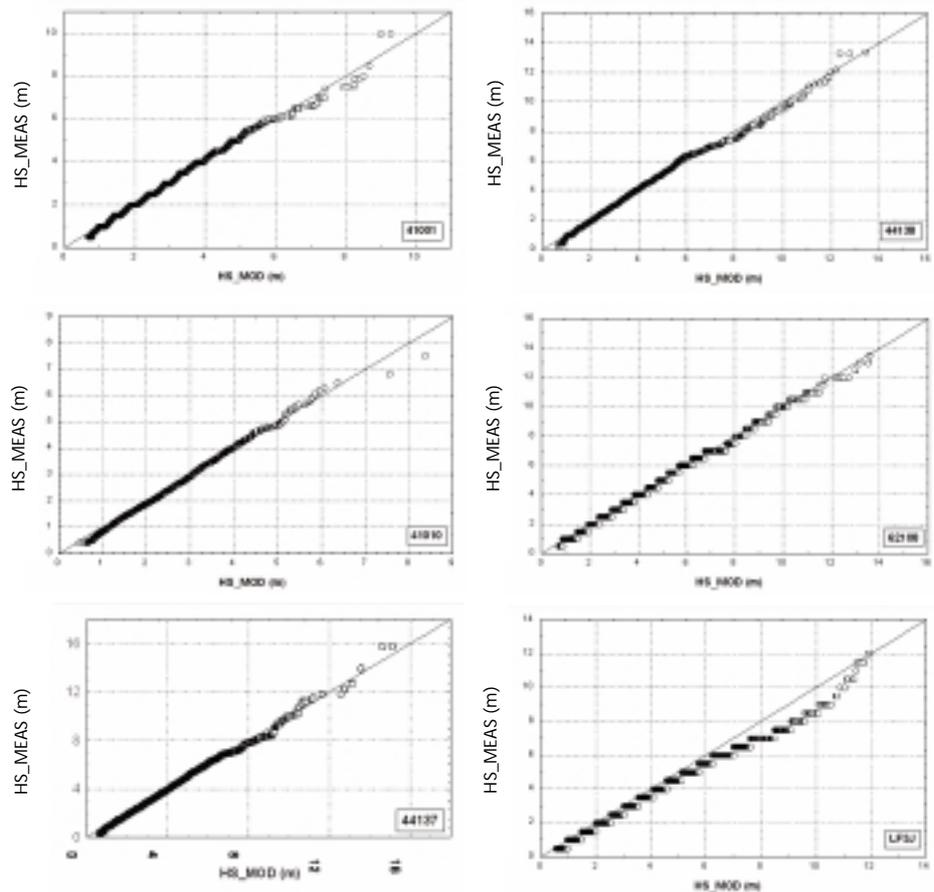


Figure 9—Quantile-quantile plots of significant wave height for selected measurement locations, based on ENRA-driven hindcasts.*



*Buoy locations for Figures 8 and 9:

41001	34°55.5'N	72°57.1'W
41010	28°52.8'N	78°32.0'W
44137	41°11.6'N	61°07.8'W
62108	53°12.0'N	15°00.0'W
LF3J	61°12.0'N	02°18.0'E

the uncertainty in ship, buoy and satellite measurements of surface wind and sea state, emphasizing extreme conditions from the perspective of some new insights gained from recent and ongoing research programmes.

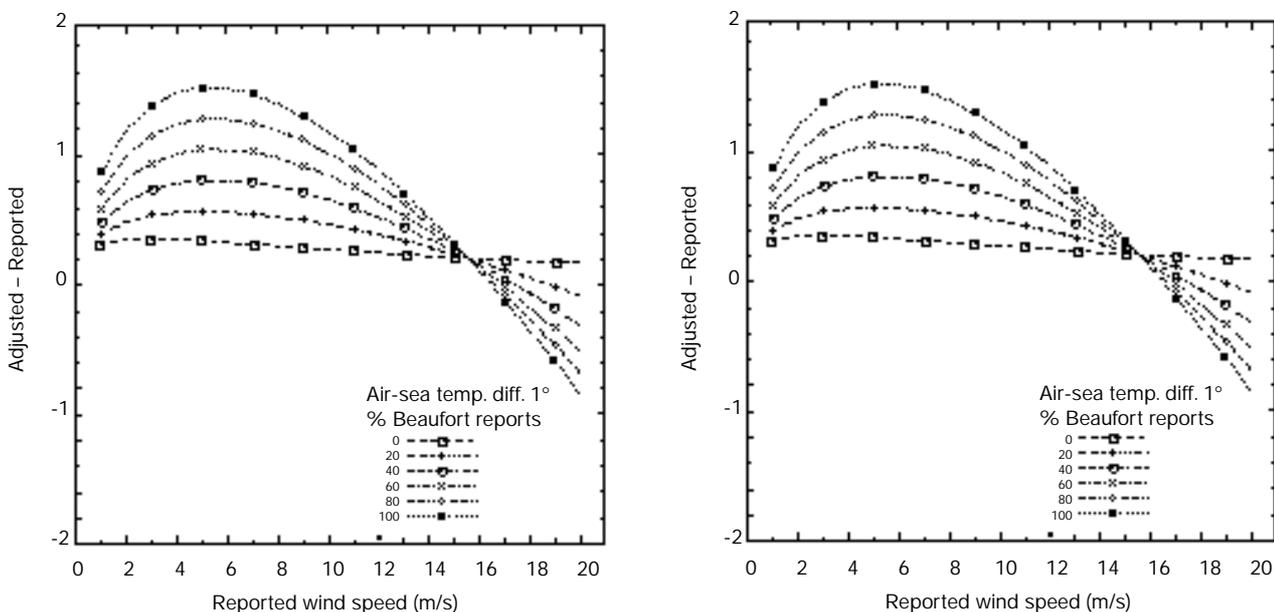
A. SHIP REPORTS

Ship reports of wind are either Beaufort estimates or anemometer measurements, and it is not always known which type a given report falls into. A great deal of new research has been reported to improve the conversion of Beaufort force or number into equivalent wind speed (e.g. Cardone *et al.*, 1990; see also Taylor *et al.*, 1995). While there are some differences between these proposed alternatives, the new scales imply that present and historical Beaufort wind speeds below about 15 m/s should be raised, and higher wind speeds lowered. Figure 10 (derived using the Cardone *et al.*, 1990 scale) shows the systematic differences between means of a population of ship wind speed reports and the ‘true’ mean, assuming that individual reports of Beaufort force and anemometer wind speed are themselves unbiased. The difference is a function of the proportion of anemometer measurements to Beaufort estimates and of the mean air-sea temperature difference. Until there is (if ever) a release of COADS in which ship winds have been adjusted to 10 m neutral winds in a systematic way, the adjustments must be made on a study-by-study basis.

For estimation of extremes in harsh climates, there is an additional limitation of ship reports. The upper limit of the Beaufort scale, namely Beaufort 12, is equivalent to wind speeds which vary according to which scale is adopted. This limit varies from 56 knots (29 m/s) according to the Cardone *et al.* (1990) scale to ‘>63 knots’ (32 m/s) for the official WMO scale. Thus, even if the estimation of Beaufort force by an observer was unequivocal and the perfect equivalency scale was known, this system simply lacks adequate dynamic range to extend to wind speeds associated with the generation of extreme sea states. Even in extratropical storms, the maximum average 10 m wind speed may range up to 40 m/s.

There are numerous sources of error or uncertainty associated with wind measurement from ships, including the height of the anemometer above sea level, corrections (or lack thereof) for ship motion, averaging interval of the measurement and distortion of the true marine wind field by the superstructure of the ship itself. A detailed review of the accuracy of ship measurements is given by Taylor *et al.* (1995). The flow distortion errors are almost always non-negligible and may be the dominant factor at high wind speeds, depending on the location of the anemometer and relative direction of the wind to the ship. The errors may also be of either sign. For this reason, Dobson (1983) recommended that corrections to measured winds from ships for anemometer height not be done unless corrections were also done for flow distortion. The latter is very difficult since

Figure 10—Difference between adjusted and reported mean wind speed for indicated percentages of estimated winds, and air minus sea virtual potential temperature (from Cardone *et al.*, 1990).



there are many different, usually unknown, effects which contribute to the flow distortion problem. However, Cardone *et al.* (1990) and Taylor *et al.* (1995) both find improvements in overall wind estimation by adjusting for anemometer height. Recently, Kent *et al.* (1999) show that after all adjustments (except flow distortion) are made, the mean random observational error of ship reports of wind speed appears to be about 2 m/s, which is about half the value previously derived from comparisons of ship and US buoy winds (Wilkerson and Earle, 1990).

B. BUOY WINDS

Meteorological buoys are widely considered to be the best source of data for marine winds. In addition to their direct use in climate analysis, buoy winds are widely used for a number of different applications: operational NWP analysis schemes; validation of hindcast and forecast wind fields; and as 'truth' for the validation and calibration of satellite and radar remote sensing systems. Buoy winds by no means form a homogeneous data type. For example, considering only the US NOAA and Environment Canada arrays we find the following differences: (1) winds from NOAA buoys are 8.5 minute scalar average speeds and directions are unit vector averages; winds from Canadian buoys are historically 10 minute vector (now scalar) average speeds and directions; (2) winds from NOAA buoys may be at either 5, 10 or 13.8 m level; wind observations from Canadian NOMAD buoys are at 4.6 to 5.4 m; (3) NOAA buoys report the highest 5 second window average obtained in the 8.5 minute sample; Canadian buoys report the highest 8 second (now being changed to 5 second) running scalar mean peak wind speed in the 10 minute sample.

The error characteristics of winds from buoys need to be better understood over a wide range of environmental conditions. Considerable work has been devoted to the demonstration of buoy capability in low to moderate sea states (e.g. Gilhousen, 1987). However, there has been little or no investigation of buoy winds in severe conditions. It is commonly believed by operational meteorologists in Canada and the USA that the buoy average wind speeds are significantly underestimated in these conditions and that the reported gust speed is a more reasonable measure of the true sustained wind speed.

A field programme supported by Environment Canada was undertaken during the winter of 1994/95 off the west coast of Canada (SWS-1) and near the Hibernia platform in the winter of 1997/98 (SWS-2) in which measured winds and waves from a NOMAD buoy were recorded twice per second when significant wave heights exceeded 8 m. Air temperature, heave, magnetometer, buoy heading and vertical wind speed were also recorded at 2 Hz; sea surface temperature was recorded every 10 minutes. Preliminary results (Skey *et al.*, 1998) show that wind speeds vary considerably over a very short time frame, e.g. a factor of 2 over less than 10 seconds even at moderate wind speeds. The wind direction may vary by more than 100 degrees over 10 minutes, with a standard deviation of 16 degrees. This variability will have a significant impact on the vector mean wind speed computed for the hourly wind report. Detailed analysis is presently being carried out to assess the magnitude of errors introduced by this vector averaging, as well as potential effects due to sheltering of the anemometers by the high waves and errors due to buoy motions. Preliminary estimates indicate that buoy average winds may be biased low by 20 per cent or more in extreme sea states (say HS > 10 m).

C. PLATFORM WINDS

Winds measured from offshore platforms are potentially the most accurate source of marine winds in extreme storms. Instrument error can be very low provided that the sensor is calibrated and checked periodically, that there is no appreciable sensor motion and that flow distortion is minimal for sensors mounted well above the platform superstructure. These conditions are increasingly being satisfied for the newer platforms in the Gulf of Mexico, North Sea, Norwegian Sea and in other frontier areas of offshore exploration and production. Typically, the anemometer is of a modern design, calibrated, and mounted at the top of the drilling derrick at heights of 40 m to as much as 140 m above the sea surface and electronically records average wind speed and direction. The only adjustments normally needed for such measurements are for sensor height and adjustment to neutral stratification. Interesting data sets have been acquired in recent North Sea

extreme storms which indicate that sustained winds, defined as maximum one-minute scalar averaged wind speeds, in the marine boundary layer reduced to equivalent 10 m neutral stratification, can range as high as 50 m/s with gusts to as high as 60 m/s. Curiously, even in the recent storms in which buoys moored in the western North Atlantic have measured HS greater than 17 m, recorded maximum sustained wind speeds from buoys have not exceeded about 30 m/s. A remaining issue of concern, however, for the higher platform-mounted anemometers (above say 50 m) is that the sensor may be above the constant stress or surface boundary layer. It may be in the part of the boundary layer where more complex wind profiles (than simple power law or logarithmic) are needed to derive the 10 m neutral wind.

D. SATELLITE WINDS

Remote sensing of the ocean is clearly an essential component of the future global observing system, due to the immense area to be covered and the difficulties and expense of using conventional in situ systems. Several types of satellite sensors capable of producing information on ocean waves and marine winds have been developed in recent years, including scatterometers, passive microwave radiometers, altimeters and synthetic aperture radars (SAR). However, these remote systems do not measure the desired geophysical parameters directly, but instead measure other parameters such as radar backscatter. Algorithms which convert radar backscatter to surface wind are developed and tuned using high-quality in situ measurements from ships and buoys - this reinforces the importance of understanding the characteristics of the in situ measurements.

The scatterometer produces estimates of both wind speed and direction from the measured radar backscatter from the ocean surface. Wind speed accuracy may reach ± 1.5 m/s in low to moderate wind speed conditions and the uncertainty in wind direction is at least $\pm 10^\circ$ after a directional ambiguity is removed by using neighbouring data or a good first guess field. Spatial sampling is of the order of about 25-50 km. Systematic errors derive from uncertainty in the backscatter-vector wind model function and in the optimum reference level for backscatter-derived winds. There is even some evidence that the uncertainty in optimum reference level is dynamic and a function of wave height. Further algorithm development in conjunction with reliable ground truth is needed to improve accuracy.

The altimeter and microwave radiometers do not provide information on wind direction. The radiometer provides wind speed data over a wide swath; the altimeter provides an averaged wind speed within its 5-10 km wide footprint directly underneath the satellite path. Accuracy is about $\pm 1-2$ m/s for the altimeter, and about ± 2 m/s for the radiometer for most cases. Little or no calibration has been done for high wind speed cases.

The SAR provides detailed information over a wide swath with errors in wind speed of about ± 1 m/s for low to moderate wind speeds in comparison with accurate in situ measurements (Vachon and Dobson, 1995). The wind direction may be deduced from SAR imagery under some circumstances or may be taken from a wind analysis chart. The SAR data may be used to study kilometre-scale wind speed variations and is therefore useful in conjunction with mesoscale wind models.

With regard to extreme storm conditions, one key question, which remains unanswered, is the upper limit of sensitivity to wind speed for all remote sensors. Empirical evidence to date does not support sensitivity above equivalent 10 m wind speeds of about 20 m/s which, if also true for newer systems (e.g. QUIKSCAT), would seriously limit the usefulness of satellite winds to specification of storm wind fields and extreme wind statistics. However, a recent study of NSCAT winds in a typhoon (Jones *et al.*, 1999) suggests that sensitivity ultimately may be extended beyond 30 m/s with improved scatterometer geophysical model functions and data processing.

Another limitation of remote sensing systems which needs to be considered is temporal resolution. As noted previously, several recent hindcast studies suggest that the wind field features responsible for the generation of very extreme sea states (say HS > 12 m) are relatively small scale and evolve and propagate rapidly (Cardone *et al.*, 1996). Ideally, three-hourly sampling is needed to resolve such

features. For even a wide-swath remote sensor to satisfy this requirement, it must be mounted on at least three operational polar orbiting satellites. It is doubtful that resources will be made available to support such an operational system in the foreseeable future, though the overlap of limited duration missions such as QUIKSCAT, ADEOS-2 and ERS-2 constitute in effect a useful, if sub-optimal, operational capability. Despite the limitations of dynamic range, the NSCAT experiment showed that significant improvements in NWP model forecasts may be realized (Atlas *et al.*, 1999) from an operational satellite remote wind vector sensing capability, whether achieved from active or passive systems.

Finally, we should note a new type of 'remote sensor' deployed from an aircraft - the Global Positioning System (GPS) dropwindsonde (Hock and Franklin, 1999). This device can measure the vertical wind profile below the aircraft, including the measurement of the 10 m wind speed, with an accuracy of 0.5 m/s to 1.5 m/s. The averaging interval of the measurement is only a few seconds so several successive drops are needed to produce an average wind profile. This instrument, which is already widely used operationally in North Atlantic tropical cyclones, promises to provide a powerful new tool to evaluate buoy, platform and satellite winds by virtue of its ability to provide truly unbiased estimates of the marine surface wind at wind speeds above 20 m/s.

5. CONCLUSIONS

While NRA wind fields appear to be a significant improvement over operational wind fields, if for no other reason than they are more homogeneous over time than real time products, they still suffer from poor resolution of areas of high winds in extratropical storms and from a lack of resolution of most tropical systems.

NRA wind fields may be improved by re-assimilation of measured data through an interactive, analyst-driven, kinematic approach. However, the limitations of each data source should first be considered to ensure that any biases associated with variable measurement heights, or different averaging intervals, are minimized. Similarly, the assimilation of any satellite measurements of high wind speeds, which are thought to be biased low through saturation, should be avoided.

Research programmes are under way to gain improved estimates of biases and random errors of various types of measurements.

QUIKSCAT and other advanced scatterometers may lead to a significant improvement in real-time wind field analyses and forecasts, but their value in storm conditions may continue to be limited by saturation at higher wind speeds above 20 m/s. Supplemental use of MPBL winds derived from pressure fields and inverse modelling using satellite wave measurements may remain useful tools in such regimes until in situ or remote sensors with greater proven dynamic range are developed and implemented.

REFERENCES

- Atlas, R., S.C. Bloom, R.N. Hoffman, E. Brin, J. Ardizzone, J. Terry, D. Bugnato and J.C. Jusem, 1999: Geophysical validation of NSCAT winds using atmospheric data and analyses. *J. Geophys. Res.*, 104, C5.
- Cardone, V.J., 1991: The LEWEX wind fields and baseline hindcast. *Directional Ocean Wave Spectra*, R.C. Beal, ed., The Johns Hopkins University Press, 136-146.
- Cardone, V.J., A.J. Broccoli, C.V. Greenwood and J.A. Greenwood, 1980: Error characteristics of extratropical storm wind fields specified from historical data. *J. Pet. Tech.*, 32, 873-880.
- Cardone, V.J., A.T. Cox, J.A. Greenwood and E.F. Thompson, 1994: Upgrade of tropical cyclone surface wind field model. U.S. Corps of Engineers Misc. Paper CERC-94-14.
- Cardone, V.J., H.C. Graber, R.E. Jensen, S. Hasselmann and M.J. Caruso, 1995: In search of the true surface wind field in SWADE IOP-1: ocean wave modeling perspective. *The Global Ocean Atmosphere System*, 3, 107-150.
- Cardone, V.J., J.G. Greenwood and M.A. Cane, 1990: On trends in historical marine wind data. *J. Climate*, 3, 113-127.
- Cardone, V.J., R.E. Jensen, D.T. Resio, V.R. Swail and A.T. Cox, 1996: Evaluation of contemporary ocean wave models in rare extreme events: "Halloween storm of

- October, 1991; "Storm of the Century" of March, 1993". *J. Atmos. Ocean. Tech.*, 13, 1, 198–230.
- Cox, A.T., J.A. Greenwood, V.J. Cardone and V.R. Swail, 1995: An interactive objective kinematic analysis system. *Proceedings, 4th International Workshop on Wave Hindcasting and Forecasting* (Banff, Alberta, October 16–20, 1995) 109–118.
- Dobson, E.B. and S.P. Chaykovsky, 1991: Geosat wind and wave measurements during LEWEX. *Directional Ocean Wave Spectra*, R.C. Beal, ed., The Johns Hopkins University Press, 128–135.
- Dobson, F.W., 1983: Review of reference height for and averaging time of surface wind measurements at sea. Report No. 3, *Marine Meteorology and Related Oceanographic Activities*. World Meteorological Organization, Switzerland.
- Gemmill, W.H., T.W. Yu and D.M. Feit, 1988: A statistical comparison of methods of determining ocean surface winds. *Weather and Forecasting*, 3, 153–160.
- Gilhousen, D.B., 1987: A field evaluation of NDBC moored buoy winds. *J. Atmos. Ocean. Tech.*, 4, 94–104.
- Graber, H.C., M.J. Caruso and R.E. Jensen, 1991: Surface wave simulations during the October storm in SWADE. *Proc. MTS'91 Conf. Marine Technology Society* (New Orleans, LA) 159–164.
- Graber, H.C., R.E. Jensen and V.J. Cardone, 1995: Sensitivity of wave model predictions on spatial and temporal resolution of the wind field. 4th International Workshop on Wave Hindcasting and Forecasting (Banff, Alberta, 16–20 October, 1995) 263–278. (Available from Atmospheric Environment Service, Downsview, Ontario).
- Hock, T.F. and J.L. Franklin, 1999: The NCAR GPS dropwindsonde. *Bull. Amer. Meteor. Soc.*, 80, 3, 407–420.
- Jones, W.L., V.J. Cardone, W.J. Pierson, J. Zec, L.P. Rice, A.T. Cox, W.B. Sylvester, 1999: NSCAT high-resolution surface wind measurements in Typhoon Violet. *J. Geophys. Res.*, 105, C5, 11247–11259.
- Kalnay, E., et al., 1996: The NCEP/NCAR 40-Year reanalysis project. *Bull. Amer. Meteor. Soc.*, 77, 3, 437–471.
- Kent, E.C., P.G. Challenor and P.K. Taylor, 1999: A statistical determination of the random observational error present in voluntary observing ships meteorological reports. *J. Atmos. Ocean. Tech.*, 16, 905–914.
- Khandekar, M.L., R. Lalbeharry and V.J. Cardone, 1994: The Performance of the Canadian Spectral Ocean Wave Model (CSOWM) During the Grand Banks ERS-1 SAR Wave Spectra Validation Experiment. *Atmosphere-Ocean*, 32, 1, 31–60.
- Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate*, 7, 141–157.
- Overland, J.E. and W.H. Gemmill, 1978: Marine winds in the New York Bight. *Mon. Weather Rev.*, 105, 1003–1008.
- Powell, M.D. and P.G. Black, 1989: The relationship of hurricane reconnaissance flight-level wind measured by NOAA's oceanic platforms (6th National Conference on Wind Engineering, 7–10 March 1989).
- Skey, S.G.P. and K. Berger-North, 1998: Measurements of winds and waves from a Nomad buoy in high sea states. 5th International Workshop on Wave Hindcasting and Forecasting (January 26–30, 1998, Melbourne, FL).
- Swail, V.R. and A.T. Cox, 1999: On the use of NCEP/NCAR reanalysis surface marine wind fields for a long term North Atlantic wave hindcast. Accepted in *J. Atmos. Ocean. Tech.*
- Swail, V.R., M. Parsons, B.T. Callahan and V.J. Cardone, 1995: A revised extreme wave climatology for the east coast of Canada. *Proceedings 4th International Workshop on Wave Hindcasting and Forecasting* (October 16–20, 1995, Banff, Alberta) 81–91.
- Taylor, P.K., E.C. Kent, M. Yelland and B. Moat, 1995: The accuracy of observations from ships. *Proc. International COADS Winds Workshop* (Kiel, Germany, 31 May – 2 June, 1994).
- Thompson, E.F. and V.J. Cardone, 1996: Practical modeling of hurricane surface wind fields. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, July/August 1996, 195–205.

- Vachon, P. and F. Dobson, 1995: Validation of wind speed retrieval from ERS-1 SAR images. Fourth International Workshop on Wave Hindcasting and Forecasting (16–20 October, Banff, Alberta, Canada).
- Weller, R.A., M.A. Donelan, M.G. Briscoe and N.E. Huang, 1991: Riding the crest: a tale of two wave experiments. *Bull. Amer. Meteor. Soc.*, 72, 163–183.
- Wilkerson, J.C. and M.D. Earle, 1990: A study of differences between environmental reports by ships in the voluntary observing program and measurements from NOAA buoys. *J. Geophys. Res.*, 95, C3, 3373–3385.

SECTION 3

METADATA AND DATA QUALITY

Improving global flux climatology: the role of metadata	89
Establishing more truth in true winds	98
Quality control in recent and pending COADS releases	116

IMPROVING GLOBAL FLUX CLIMATOLOGY: THE ROLE OF METADATA

Elizabeth C. Kent*, Peter K. Taylor and Simon A. Josey

1. INTRODUCTION

This paper will describe the use of metadata in the development of the Southampton Oceanography Centre (SOC) Surface Flux Climatology (Josey *et al.*, 1999). The data source for the climatology was the merchant ship weather reports within the Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a (Woodruff *et al.*, 1993). Although the quality of reports from the merchant ships, the Voluntary Observing Ships (VOSs), is known to be variable, the reports are a valuable source of data over the oceans. The metadata we shall use are collected by Port Meteorological Officers around the world and in the period up until 1994 were published annually by WMO (WMO-No. 47 e.g. WMO, 1994). Metadata for later years can be found on the WMO web site. COADS also contains data from other sources, such as moored and drifting buoys and oil platforms, but metadata for these data sources have not been collated in the same way as for the VOS data. Buoy metadata were therefore not used in the SOC climatology but should prove useful in future.

The information contained in the WMO-No. 47 metadata allows the identification of instrument types and heights for most VOS weather reports from COADS Release 1a for the 1980-1993 period. Using this information we can apply the results of the VOS Special Observing Project - North Atlantic (VSOP-NA) which identified errors in merchant ship weather observations. The importance of external sources of metadata is well recognized and the next version of COADS, Release 2 (Woodruff *et al.*, 1998), will contain metadata enhancements.

Climatological estimates of the ocean surface heat balance can be calculated from COADS individual weather reports. Bulk formulae (e.g. Smith, 1980; 1988) are used to calculate heat and momentum fluxes from the meteorological variables reported by the ships. In past climatologies, global adjustments have been made to the resulting flux fields to balance the global heat budget. Authors have justified these modifications as potentially compensating for the effect of ship measurement errors on the fluxes, and for uncertainties in the bulk formulae used to calculate the surface fluxes from the VOS reports. The approach taken to balance the heat budget has been to appeal to external information to constrain the fluxes. For example, da Silva *et al.* (1994) used inverse analysis to simultaneously tune COADS-based heat and fresh water fluxes to conform with oceanographic estimates of meridional heat and fresh water transports. This resulted in a 13 per cent increase in the latent heat flux (with a compensatory increase in the precipitation to balance the fresh water transport) and an 8 per cent decrease in the incoming solar flux. The sensible heat flux and the longwave flux were changed by smaller amounts. Even larger adjustments have been made in other studies that are hard to justify (Kent and Taylor, 1995).

Following the identification of sources of error in VOS data in the VSOP-NA project we were able to test whether the heat imbalance (about 30 Wm^{-2} excess heating of the ocean) found in VOS-derived flux climatologies is due to errors in the data. If the VSOP-NA corrections applied to the COADS data are similar in size to the global adjustments used by da Silva *et al.* (1994) then we might assume that the latter can be justified on the grounds of ship errors. If not, we will have to look for other methods of balancing the heat budget.

* Corresponding author's address:

Elizabeth C. Kent, James Rennell Division for Ocean Circulation and Climate, Southampton Oceanography Centre, European Way, Southampton, SO14 3ZH, UK

2.
CORRECTING THE VOS
DATA
2.1
THE CORRECTIONS

The VSOP-NA (Kent *et al.*, 1991, Kent and Taylor, 1991) consisted of the detailed analysis of two years of meteorological reports from 46 ships selected because they reported regularly in the North Atlantic. Port Meteorological Officers gathered detailed information about the ships and the instruments carried. Photographs or plans of the ships and of the instrumentation sites were collected where possible along with information on observing practices. In addition, extra fields were added to the ships' weather log to identify the conditions at the time of the observations. Deutscher Wetterdienst, Hamburg, keyed the information from the logbooks (over 33 000 records) into ASCII format. The reports were then merged with the output of a numerical weather prediction model by the UK Met Office. The model was used to provide a consistent standard to allow ship reports separated in space and time to be compared. Figure 1 shows one of the ships that took part in the VSOP-NA project which is typical of many of the ships providing weather reports in the North Atlantic.

An example of the VSOP-NA results for sea surface temperature (SST) is shown in Figure 2. The ships in the VSOP-NA project used three different methods of measuring the SST (Kent and Taylor, 1991). Ships recruited by Germany and the Netherlands used insulated buckets to collect samples of surface seawater and measured the temperature of the water on deck with a thermometer. Ships from France and the United States measured the temperature of seawater pumped aboard to cool the ships engines (engine room intake, ERI). Ships recruited by the UK used a combination of methods but some were fitted with a dedicated SST sensor attached to the ships hull, a hull contact sensor.

Using the model as a comparison standard (Figure 2(a)) the bucket and hull sensor data were in reasonable agreement at night, while the ERI data was comparatively warm. The hull contact data were less scattered than those from other methods. Using the hull contact data as a reference (Figure 2(b)) showed that the ERI data were on average biased high by between 0.2 and 0.4°C; a typical mean value was 0.35°C but individual ships had mean biases between -0.5°C (too cold) and +2.3°C (too warm). Figure 2(b) also indicates that the bucket values were possibly about 0.1°C cold at night but became biased warm by up to 0.4°C with increasing solar radiation.

These and other results from the VSOP-NA suggested that:

1. Sea surface temperature (SST) measurements made using engine intake thermometers were biased warm (Kent *et al.*, 1993a). This correction could be applied for the logbook reports which contain the method of SST measurement, but for reports received by radio during this period the method of SST measurement needed to be found from the WMO-No. 47 metadata.
2. Air temperature measurements were affected by solar radiation. The warm bias caused by the solar heating of the ship superstructure could be removed on average using a formula depending on the incoming solar radiation and



Figure 1—The Nedlloyd Zeelandia, one of the 46 VSOP-NA ships.

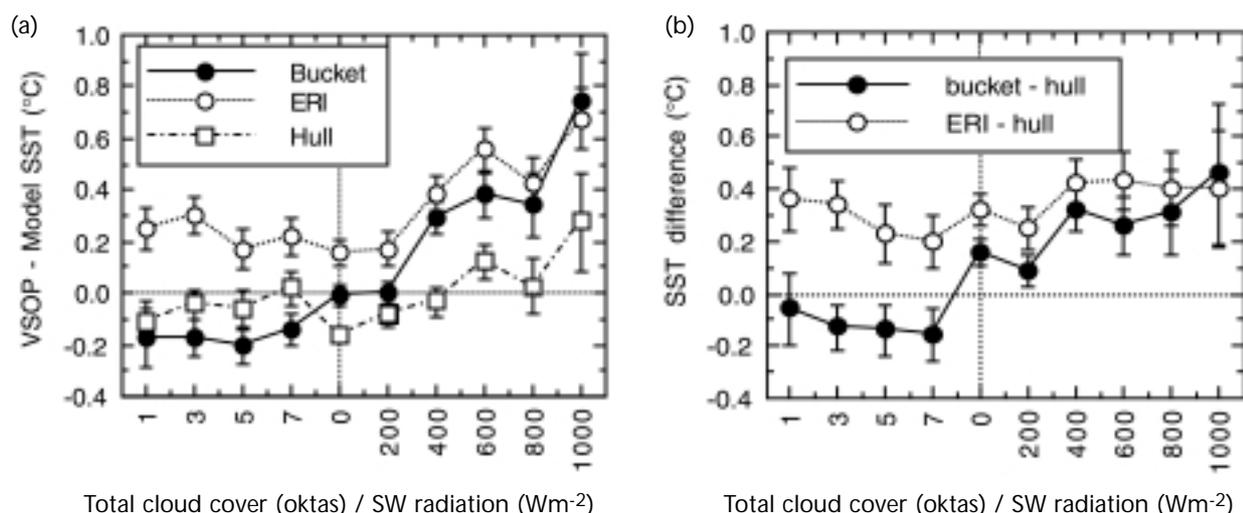


Figure 2—Comparisons of SST data obtained from the VSOP-NA ships using SST buckets, engine room intake (ERI) thermometers and hull contact sensors. Night-time data is plotted against total cloud amount and daytime data against the estimated solar radiation.

(a) mean difference (ship data - model value)
 (b) mean difference using the hull contact sensor data as a reference. (from Taylor *et al.*, 1998).

the relative wind speed (Kent *et al.*, 1993b). Both these parameters can be calculated from information in the normal weather report; no external metadata are required.

- Humidity measurements from both screens and psychrometers were unaffected on average by solar heating (Kent and Taylor, 1996).
- Humidity measurements from screens were biased high when compared with those from psychrometers, presumably due to their poorer ventilation. This bias could be removed on average by reducing the humidity from screens using an empirical formula (Kent *et al.*, 1993a). The method of humidity measurement needs to be found from WMO-No. 47 metadata.
- Height correction of instrument-based wind speed measurements to 10 m should be carried out to homogenize the wind data (Kent *et al.*, 1993a). The height of the anemometer is contained in the WMO-No. 47 metadata. It should be noted that some recent wind reports are adjusted to 10 m height on board ships by software such as TurboWin. At present there is no way of telling which reports have already been adjusted and, therefore, height adjustment has been made for all anemometer measured reports. As more ships start to use such automated software, the number of reports that do not require further adjustment will increase. An extension to the wind speed indicator (iw in the Ship Code) or other data flag to indicate which reports have been adjusted to 10 m height would be necessary to allow this correction to be applied with confidence in the future.

Additionally, visual estimates of sea state need to be converted to a wind speed using a Beaufort equivalent scale. This conversion is made on board the ship by the observer using a conversion code that is now thought to introduce biases into the data (WMO, 1970). Many alternative Beaufort equivalent scales have been suggested in the literature. Kent and Taylor (1997) found that adjusting visually observed winds to conform to the Beaufort equivalent scale of Lindau (1995) gave the best agreement between one-degree monthly mean wind speeds derived from anemometers and those derived from visual observations. The method of wind measurement, visual or by anemometer, is contained within COADS and no external metadata are required to make this adjustment.

2.2 THE METADATA

The metadata for merchant ships in the Voluntary Observing Fleet have been published annually by WMO since the 1950s (e.g. WMO, 1994) and are available in electronic format for 1973 onwards. For each ship the metadata consists of the ship's name and call sign followed by a coded list of instrument types and heights. The instrument types for pressure, air and sea temperature and humidity are listed along with information about more specialized instrumentation installed on the ship. Anemometer heights are listed for those ships carrying anemometers (although some ships with anemometers still report winds from visual observations of sea state if national observing practices so dictate) along

with the height of the observing platform which we have used as a proxy for the height of air temperature and humidity measurement.

The ASCII version of the metadata for 1973 to 1994 which was reformatted as part of the SOC flux climatology project can now be found on the internet via the COADS web site: <http://www.cdc.noaa.gov/coads/>.

Updates to WMO-No. 47 can be found on the WMO web site: <http://www.wmo.ch/web/ddbs/ddbs.html>.

2.3
COMBINING DATA AND
METADATA

The link between the COADS data and the WMO-No. 47 metadata is made using the ship's call sign, which is at present in both data sets. Since a given call sign can be transferred from one ship to another, and because changes may occur in the metadata, the matching must be done on a year-by-year basis. For each ship's meteorological report in COADS, the WMO-No. 47 database was searched to find the metadata for the ship with the appropriate call sign in that particular year. About 10 per cent extra reports were matched if reports that were unmatched with the correct year were checked with metadata for the following year, indicating that there is sometimes a time lag between a ship being recruited to make weather reports and its details being collected for WMO-No. 47. Figure 3 shows the success rate for matching a report in COADS with the ship metadata. This figure shows that although the number of reports in COADS has increased slightly over the 1980–1993 period, the number of reports from ships has declined. The deficit is made up from reports from fixed platforms and moored and drifting buoys. The composition of COADS is therefore very different towards the end of this 14-year period than at the beginning. This also leads to reduced spatial coverage since the data from platforms and buoys are usually restricted to fixed locations near the coast. The matching rate increased from less than half the ship reports at the beginning of the period, largely due to the lack of call sign information in COADS, to more than 80 per cent by the end of 1992. The step increase in the match rate in 1982 is due to the inclusion of ship call signs in the WMO format for the exchange of ship logbook data at that time.

3.
RESULTS
3.1
EFFECT OF THE VSOP-NA
CORRECTIONS

The effect of the corrections described in section 2.1 on the flux fields can now be determined (for a full description of the corrected climatological fields see Josey *et al.*, 1999). As an example we shall take the latent heat flux. Figure 4 shows the effect of the VSOP-NA corrections, the individual height corrections (as opposed to a single assumed height of 25 m for anemometer winds) and the visual scale of Lindau (1995) on the latent heat flux in January 1990.

The latent heat flux is reduced in the North Pacific, in some regions by more than 15 Wm⁻². This is due to the correction to SST from engine intakes. The SST

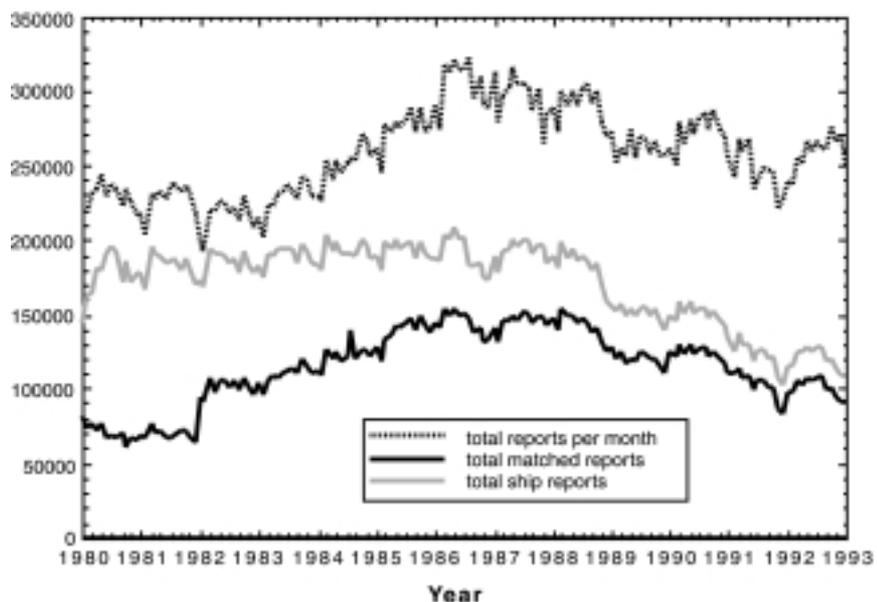
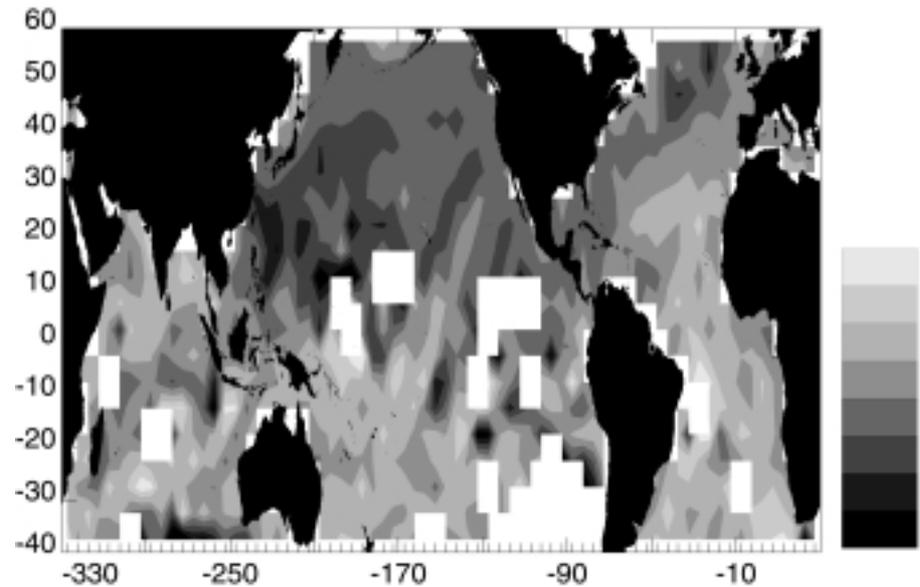


Figure 3—The number of reports per month in COADS Release 1a (top dotted line). The number of reports from ships (centre line). The number of reports for which metadata were found in WMO-No. 47 (lower dark line).

Figure 4—The effect of VSOP-NA corrections applied to the VOS data on the latent heat flux. Differences are plotted in Wm^{-2} and a negative difference represents a decrease in the heat loss from the ocean due to the corrections.



is reduced in these cases, which reduces the saturation humidity at the sea surface and hence the sea-air humidity difference results in a decrease in the latent heat flux. An additional, but smaller, decrease in the latent heat flux in this region arises from the use of individual anemometer heights. In contrast, the correction to the screen humidities decreases the air humidity, hence increasing the sea-air humidity difference and leading to an increase in the latent heat flux. It is largely this effect which leads to the increased latent heat flux values over the subtropical and South Atlantic where most of the screen-measured humidities are reported. In the North Atlantic, particularly in the north-west, the effect of the individual anemometer heights causes a decrease in the latent heat flux. The effect of the solar radiation correction to the air temperature only affects the calculation of stability and the effects on the latent heat flux are therefore small. A larger effect of this correction is seen in the sensible heat flux field (not shown). The overall effect of the individual height corrections and the Lindau (1995) scale is patchy. The effect of ship corrections on the latent heat flux is thus complex and regional and has little correlation with the magnitude of the latent heat flux. This implies that a global increase of latent heat flux in order to match the fluxes with the ocean heat transport estimates cannot be justified on the grounds of errors in the ship data.

The use of metadata has shown us that the latent heat flux errors are regional in nature. This suggests that for a regional study, globally adjusted latent heat fluxes may not be the best to use. This is investigated further in the next section.

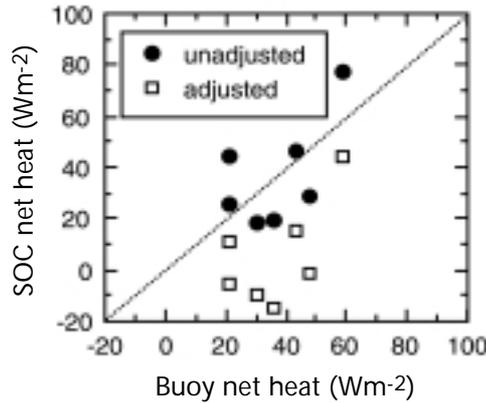
3.2 EXAMPLE OF HEAT FLUX VALIDATION

The surface fluxes in the climatology require validation against independent data. An example of this is shown in Figure 5, in which the net heat flux data from IMET research buoys (Weller *et al.*, 1998) are compared with data from the climatology (Weller and Taylor, 1999). The IMET buoys produced high-quality flux data in the subtropical Atlantic, the Arabian Sea and the tropical Pacific. The buoy and climatology agree to within $25 Wm^{-2}$. If the flux components are adjusted in a manner similar to da Silva *et al.* (1994) then the agreement between the buoys and the climatology worsens (open squares in Figure 5). In addition, White and da Silva (1999) find that zonal mean net heat flux estimates from the da Silva *et al.* (1994) unadjusted fluxes compare better with the output of reanalysis models than the globally balanced fluxes in well sampled mid-latitude regions. This again suggests that regional, and not global, corrections are required.

3.3 EFFECT OF METADATA ON RANDOM ERRORS

The use of metadata with COADS data can reduce the random errors present in the data set, which will impact on the accuracy of monthly mean values, particularly in poorly sampled regions. An example of this was demonstrated by Kent *et al.* (1999) who used the semivariogram technique following Lindau (1995) to

Figure 5—Comparison of the net heat flux calculated from the SOC climatology and from IMET buoy data. The open symbols show the effect of globally adjusting the SOC fluxes to balance the heat budget.



determine the random errors in merchant ship observations using the same meta data enhanced version of COADS as Josey *et al.* (1999). The semivariogram technique is used to extrapolate the mean square differences of many pairs of ship reports to estimate the root mean square (rms) error expected if the ships were in the same position. The spatial element of variability in the ship reports is thus removed giving a better estimate of the true random error. Figure 6 shows the mean square SST and wind speed differences for pairs of ships in January 1980 in the North Pacific averaged in 50 km ranges of ship separation. The spatial element of the mean square difference increases nearly linearly for ships less than 300 km apart and can be extrapolated to zero separation to give the observational component of the mean square difference.

Kent *et al.* (1999) found that using the individual anemometer heights from WMO-No. 47 to correct the data to the standard level of 10 m above sea level, rather than assuming an average of 25 m anemometer height, reduced their random error estimates by about 15 per cent. Figure 7 shows the random observational errors derived for wind speed after applying the individual height correction to the data. If error estimates are required for uncorrected wind speed estimates these values should be increased by 15 per cent. It is worth noting that a random error in wind speed will cause a systematic error in the wind stress. Hence the reduction of wind speed random errors, even in well-sampled regions, will improve the quality of wind stress data sets.

Although the method of calculation employed by Kent *et al.* (1999) precluded any estimate of the changes to random errors in other variables due to the metadata-derived corrections, there was an indication that air temperature random errors were reduced by the corrections applied.

Figure 6—Examples of semivariograms for SST and 10 m corrected wind speed in the North Pacific for January 1980. The dark circles are the SST mean variance ($^{\circ}C^2$) and the open circles the wind speed mean variance (m^2s^{-2}) in 50 km ranges for ship pair separations below 1000 km. The lines are a regression on the individual points for data pairs up to 300 km separation. The intercept of the plot is twice the error variance. The regression lines are:
 SST $y = 0.029 (\pm 0.001) x + 3.65 (\pm 0.1)$
 Wind $y = 0.040 (\pm 0.001) x + 13.9 (\pm 0.3)$.

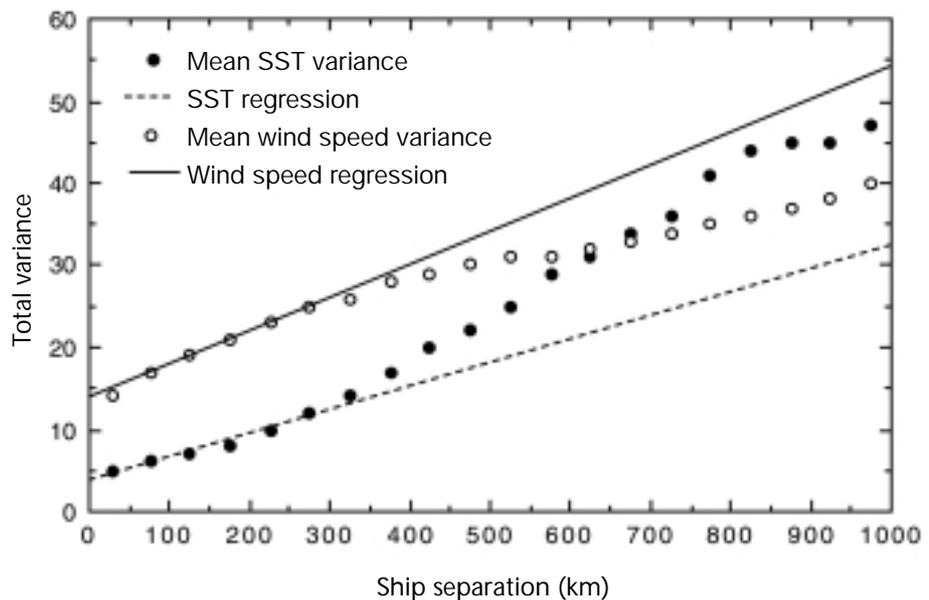
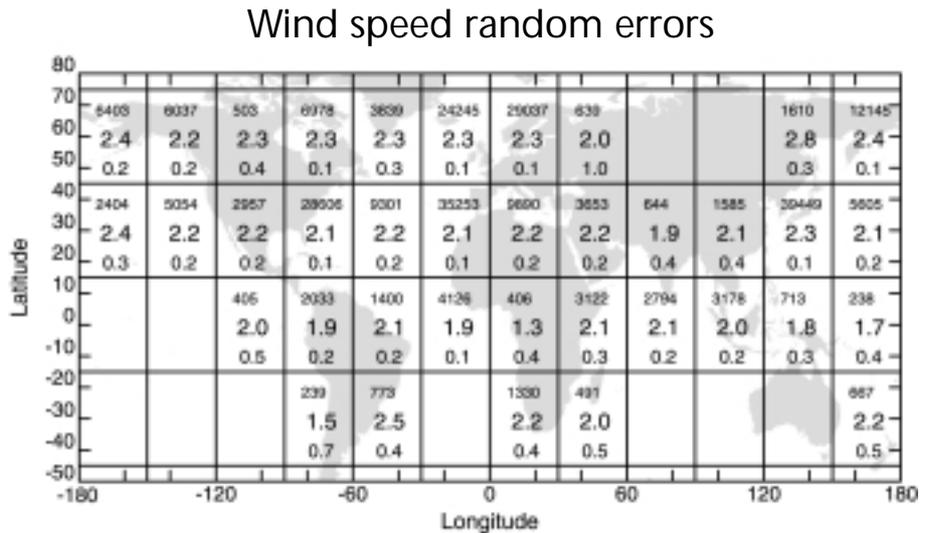


Figure 7—Wind speed random observational errors derived from COADS. The upper figure is the number of report pairs used to make the error estimate, the large central figure is the rms error (ms^{-1}) for that particular 30° region and the lower figure is the estimated uncertainty in the rms error estimate.



3.4 FUTURE USE OF METADATA

As computing and data storage capacity increase, the use of metadata with data sets such as COADS becomes easier. Although the results of the VSOP-NA suggested corrections that might be applied to merchant ship observations, these were derived from a limited subset of large merchant ships operating in the North Atlantic. Josey *et al.* (1999) applied these corrections to data from other regions in the absence of other information. The metadata-enhanced version of COADS and COADS Release 2 will allow the extension of the study of biases to other ocean basins. The semivariogram technique could be extended to look at mean differences as well as rms differences. The metadata allow the mean differences between data measured using different methods to be determined and the study of these differences under different environmental conditions. Thus, the results of the VSOP-NA could be tested for a wider range of ships and extended to other regions.

4. SUMMARY AND CONCLUSIONS

Surface flux climatologies calculated from VOS reports typically show a 30 Wm^{-2} global heat imbalance which is often removed by scaling the flux components using inverse techniques. This has been justified by appealing to errors in the VOS data. Surface fluxes calculated from the metadata-enhanced COADS showed regional and seasonal differences from those calculated without bias correction. The differences depended on the types of instruments used by ships in a particular region and are therefore affected by the choice of instrument type made by local meteorological agencies. The resulting bias corrections are not simply related to the magnitude of the flux, which suggests that global adjustments scaling the fluxes will not give a correct regional picture of the surface fluxes. This is confirmed by comparing the SOC climatology with high-quality research buoy data in the limited regions where it is available. In most cases global adjustments to the latent and shortwave fluxes typical of those required to remove the heat imbalance worsen the local agreement of the ship-derived and buoy heat fluxes. The global heat flux imbalance is, however, little affected by the corrections to the fluxes, which increase the fluxes in some regions and decrease them in others.

We cannot currently appeal to models to determine the correct fluxes. Unconstrained models can show even larger global heat imbalances and worse agreement with the research quality buoy data.

The VSOP-NA project used reports from a small subset of ships in the North Atlantic which were compared with model output. The large increase in computing resources since the project means that different types of analyses are now possible. Computationally intensive statistical techniques have been used to quantify the random errors in COADS ship reports, which will prove valuable to those wishing to apply advanced interpolation and smoothing techniques to VOS data sets. These statistical methods can be extended to examine biases between instrument types in a metadata-enhanced data set. This would extend the results

of the VSOP-NA beyond the North Atlantic, examine data from countries that did not participate in the VSOP-NA and include a wider range of ship types.

REFERENCES

- da Silva, A.M., C.C. Young and S. Levitus, 1994: *Atlas of Surface Marine Data Vol. 1 Algorithms and Procedures*. NOAA, 83 pp.
- Josey, S.A., E.C. Kent and P.K. Taylor, 1999: New insights into the ocean heat budget closure problem from analysis of the SOC air-sea flux climatology. *Journal of Climate*, 12(9), 2856–2880.
- Kent, E.C., P.G. Challenor and P.K. Taylor, 1999: A statistical determination of the random observational errors present in Voluntary Observing Ships meteorological reports. *Journal of Atmospheric and Oceanic Technology*, 16(7), 905–914.
- Kent, E.C. and P.K. Taylor, 1991: *Ships observing marine climate: A catalogue of the voluntary observing ships participating in the VSOP-NA*. World Meteorological Organization, 123 pp.
- Kent, E.C. and P.K. Taylor, 1995: A comparison of sensible and latent heat flux estimates for the North Atlantic ocean. *Journal of Physical Oceanography*, 25(6), 1530–1549.
- Kent, E.C. and P.K. Taylor, 1996: Accuracy of humidity measurements on ships: Consideration of solar radiation effects. *Journal of Atmospheric and Oceanic Technology*, 13(6), 1317–1321.
- Kent, E.C. and P.K. Taylor, 1997: Choice of a Beaufort equivalent scale. *Journal of Atmospheric and Oceanic Technology*, 14(2), 228–242.
- Kent, E.C., P.K. Taylor, B.S. Truscott and J.A. Hopkins, 1991: The accuracy of ship's meteorological observations - Results of the VSOP-NA. World Meteorological Organization, Marine Meteorology and Related Oceanographic Activities Report No. 26, 86 pp.
- Kent, E.C., P.K. Taylor, B.S. Truscott and J.S. Hopkins, 1993a: The accuracy of voluntary observing ship's meteorological observations - Results of the VSOP-NA. *Journal of Atmospheric and Oceanic Technology*, 10(4), 591–608.
- Kent, E.C., R.J. Tiddy and P.K. Taylor, 1993b: Correction of marine air temperature observations for solar radiation effects. *Journal of Atmospheric and Oceanic Technology*, 10(6), 900–906.
- Lindau, R., 1995: A new Beaufort equivalent scale. *Proceedings of the International COADS Winds Workshop* (Kiel, Germany, 31 May - 2 June 1994) Institut für Meereskunde, Kiel/NOAA, 232–252.
- Smith, S.D., 1980: Wind stress and heat flux over the ocean in gale force winds. *Journal of Physical Oceanography*, 10, 709–726.
- Smith, S.D., 1988: Coefficients for sea surface wind stress, heat flux and wind profiles as a function of wind speed and temperature. *Journal of Geophysical Research*, 93, 15,467–15,474.
- Taylor, P.K., E.C. Kent and S.A. Josey, 1998: The accuracy of sea surface temperature data from voluntary Observing Ships. GCOS Workshop on Global Sea Surface Temperature (The International Research Institute for Climate Prediction, Lamont-Doherty Earth Observatory, Palisades, New York, USA, 2-4 November 1998).
- Weller, R.A. and P.K. Taylor, 1999: *A Strategy for Marine Meteorological and Air-Sea Flux Determination*. Preprint Volume: 3rd Symposium on Integrated Observing Systems (Dallas, Texas, January 1999, American Meteorological Society).
- Weller, R.A., M.F. Baumgartner, S.A. Josey, A.S. Fischer and J.C. Kindle, 1998: Atmospheric forcing in the Arabian Sea during 1994-1995: Observations and comparisons with climatology models. *Deep Sea Research*, 45(11), 1961–1999.
- White, G.H. and A.M. da Silva, 1999: A comparison of fluxes from the reanalysis with independent estimates. Submitted to the ECMWF Reanalysis Conference (Reading, 1999).
- WMO, 1970: *The Beaufort Scale of Wind Force (technical and operational aspects)*, Reports on Marine Science Affairs, No. 3, WMO, Geneva, 22 pp.
- WMO, 1994: *International List of Selected, Supplementary and Auxiliary Ships*, WMO-No. 47, Geneva.

Woodruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley, 1998: COADS Release 2, data and metadata enhancements for improvements of marine surface flux fields. Physics and chemistry of the earth, *Proceedings of the EGS XXII General Assembly* (Vienna, Austria, 21–25 April 1997).

Woodruff, S.D., S.J. Lubker, K. Wolter, S.J. Worley and J.D. Elms, 1993: Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. *Earth System Monitor*, 4(1), 4–8.

ESTABLISHING MORE TRUTH IN TRUE WINDS

Shawn R. Smith, Mark A. Bourassa and Ryan J. Sharp, Center for Ocean-Atmospheric Prediction Studies (COAPS)*

ABSTRACT Techniques are presented for the computation and quality control of true winds from vessels at sea. Correct computation of true winds and quality control methods are demonstrated for complete data. Additional methods are presented for estimating true winds from incomplete data. Recommendations are made for both existing data and future applications.

Quality control of Automatic Weather Station (AWS) data at the World Ocean Circulation Experiment Surface Meteorological Data Center (WOCE-MET) reveals that only 20 per cent of studied vessels report all parameters necessary to compute a true wind. Required parameters include the ship's heading, course over the ground (COG), speed over the ground (SOG), wind vane zero reference and wind speed and direction relative to the vessel. If any parameter is omitted or if incorrect averaging is applied, AWS true wind data display systematic errors. Quantitative examples of several problems are shown in comparisons between collocated winds from research vessels and the NASA scatterometer (NSCAT). Procedures are developed to identify observational shortcomings and to quantify the impact of these shortcomings in the determination of true wind observations.

Methods for estimating true winds are presented for situations where heading or COG is missing. Empirical analysis of two vessels with high-quality AWS data showed these estimates to be more accurate when the vessel heading is available. Large differences between the heading and COG angles at low ship speeds make winds estimated using the course unreliable (direction errors exceeding 60°) for ship speeds of less than 2.0 m s^{-1} . The threshold where the direction difference between a course estimated and true wind reaches an acceptable level ($\pm 10^\circ$) depends upon the ship, winds and currents in the vessel's region of operation.

1. INTRODUCTION Techniques are presented to calculate and quality control true winds from automated observations collected on sea-going vessels. True wind is defined herein as a vector wind with a speed referenced to the fixed earth and a direction referenced to true north. These techniques are developed to improve the accuracy of true winds calculated by maritime Automatic Weather Stations (AWSs). The need for accurate true winds from ships arises from a desire to improve the quality of flux fields over the ocean, coupled ocean-atmospheric modelling, operational forecasting and over-water climatologies. Correct true wind calculations are provided as a tutorial and quality control procedures are developed to identify shortcomings in existing data reporting and recording practices. Methods for estimating true winds from incomplete data are shown and evaluated. Recommendations are made for both existing data and future applications.

Numerous problems relevant to true winds are identified by the quality control team at the World Ocean Circulation Experiment Surface Meteorological Data Center (WOCE-MET) using data from 20 AWS-equipped vessels. One serious problem is the incomplete or inaccurate reporting of both navigation and measured wind parameters. The parameters necessary to compute true winds include

* The Florida State University
Tallahassee, FL 32306-2840 USA

Reprinted with the kind permission of the American Meteorological Society

© 1999 American Meteorological Society

the ship's heading, the course and speed over the ground, the wind vane zero reference, and the wind direction and speed relative to the vessel. Only 20 per cent of the studied vessels report all six parameters. Further investigation reveals an underlying confusion concerning the definition of true winds. Meteorologists, oceanographers and members of the merchant marine typically define true wind differently and, as a further complication, the convention is rarely reported with the wind data. Of the 20 vessels studied, nine report their winds using a meteorological definition, one uses an oceanographic definition, and the remaining ten (50 per cent) report no definition. Additional problems include the placement and orientation of wind instrumentation, flow distortion (Yelland *et al.*, 1998), averaging methodology, and confusion over how to correctly compute true winds. As a result of the above problems, we can confirm the accuracy of reported true winds on only four of 20 vessels studied.

Solutions to problems with wind observations are presented for both future applications and existing (often incomplete) data sets. The presented techniques are a direct result of the work carried out by WOCE-MET personnel to identify, collect and quality control 181 ship months of AWS data from international research vessels (R/Vs). Our focus is on high temporal resolution automated data, although most techniques can be applied to manual observations collected at standard synoptic times. Shortcomings in the observations archived by WOCE-MET lead to the obvious conclusion that future data collection and reporting must include all the parameters required to compute a true wind. Furthermore, quality control must be applied to navigation data, measured winds and calculated true winds to identify problems. When all necessary parameters are reported, and the methodology and quality control procedures outlined herein are applied, an accurate meteorological true wind can be computed.

Procedures are outlined to estimate true winds when existing data sets lack either the heading or course angles. The limitations of these techniques are evaluated by comparing the estimates to correctly computed true winds. Estimates computed using a heading to approximate the course of the vessel are found to be superior to those constructed using the course to approximate the heading. Large differences between the heading and course angles at low ship speeds make winds estimated using the course unreliable (direction errors exceeding 60°) for ship speeds of less than 2.0 m s^{-1} . The threshold where the direction difference between a course estimated and true wind reaches an acceptable level (i.e. $<10^\circ$) can be determined empirically and depends upon the ship, winds and currents in the vessel's region of operation. These techniques produce true wind estimations from incomplete data sets. The range of conditions for which these techniques are valid is also examined.

2. CAUSES OF INACCURACY

Inaccuracies in true winds result from many problems, particularly the confusion surrounding the definition of a true wind and the parameters needed to calculate that wind. There are also problems associated with the location and calibration of instruments, averaging, and recording of both wind and navigation measurements. We begin by defining all essential parameters related to true winds and their computation. Definitions typically used by meteorologists, oceanographers and the merchant marine are discussed. We end this section with descriptions of typical problems found in the WOCE automated data.

A. DEFINITIONS

Navigational and wind parameters defined by meteorologists, oceanographers and the merchant marine are outlined in Table 1. Each group defines a course over the ground, ship speed over the ground, heading, platform-relative wind, apparent wind and true wind. For each measured parameter, the velocity and direction are referenced either to the ship or the fixed earth. The ship's directional reference frame has zero degrees at the bow of the vessel with angles increasing in a clockwise direction, while the earth's reference frame has true north corresponding to zero degrees with angles increasing in a clockwise direction. Each directional parameter has positive values defined with a direction to, or from, which the wind or ship is moving.

Table 1—Definitions of wind and navigation parameters for the three most common sources of observations. Differences from the meteorological conventions are emphasized with italics. The merchant marine has two definitions of apparent wind: (1) the wind experienced on the deck of the ship with a direction referenced to true north (consistent with the meteorological definition); and (2) the wind measured by the anemometer (similar to the meteorological platform-relative wind). The merchant marine also has two definitions of true wind: (1) relative to true north; and (2) relative to the bow of the ship. The use of a zero reference angle (zero ref. ang.) measured with respect to the bow is common to all three groups.

<i>Meteorological definitions</i>	<i>Velocity frame of reference</i>	<i>Directional frame of reference</i>	<i>Direction convention</i>
Ship COG and SOG	fixed earth	true north	moving to
Ship heading	fixed earth	true north	moving to
Platform-relative winds	ship	zero ref. ang.	moving from
Apparent winds	ship	true north	moving from
True winds	fixed earth	true north	moving from
<i>Oceanographic definitions</i>			
Ship COG and SOG	fixed earth	true north	moving to
Ship heading	fixed earth	true north	moving to
Platform-relative winds	ship	zero ref. ang.	<i>moving to</i>
Apparent winds	ship	true north	<i>moving to</i>
True winds	fixed earth	true north	<i>moving to</i>
<i>Merchant marine definitions</i>			
Ship COG and SOG	fixed earth	true north	moving to
Ship heading	fixed earth	true north	moving to
Platform-relative winds	ship	zero ref. ang.	moving from
Apparent winds (1)	ship	true north	moving from
Apparent winds (2)	ship	<i>bow of ship</i>	moving from
True winds (1)	fixed earth	true north	moving from
True winds (2)	fixed earth	<i>bow of ship</i>	moving from

Course over the ground (COG) is defined as the direction (relative to true north) in which the vessel actually moves over the fixed earth (Bowditch, 1984). Course, which differs from the COG, is defined as the "horizontal direction in which the vessel is steered" (Bowditch, 1984). For the purpose of computing true winds, the COG is the essential measurement. The speed at which the vessel moves in the direction of the COG is known as the speed over the ground (SOG). The accuracy of the COG and SOG depends on the navigation system. The older NAVSAT (TRANSIT) system and the Global Positioning System (GPS) indicate different values for COG and SOG (Bowditch, 1984). Of the 12 studied vessels that reported a COG, eight used GPS, one utilized an integrated inertial navigation-GPS, and the other three systems were unknown.

Heading is defined as the direction to which the bow is pointing relative to true north (Bowditch, 1984). Without this parameter, true winds cannot be computed. The heading is necessary to orient the shipboard anemometer's wind direction to true north. The heading and COG are not identical. For example, some R/Vs can be propelled to astern, resulting in a COG that is 180° opposite the heading. Differences between COG and heading are also the result of currents, wind, and steering error (Bowditch, 1984), and they are greatly reduced when the vessel is moving forward at a moderate or greater speed.

In addition to the ground referenced navigation (COG, SOG and heading), a common practice is to measure the motion of the vessel through the water. This water-relative motion is a vector with components along, and perpendicular to, the axis of the ship. The fore to aft component of this motion (SOW_{FA}) is defined in all the observational data sets provided to WOCE-MET as the speed over the water. As defined, the SOW_{FA} is the speed of the vessel in the direction of the heading. The component of the water-relative motion along the beam of the ship can be measured by a two-axis speed log; however, this component was only provided by one of the 20 studied vessels so we limit our discussion to the SOW_{FA} .

Most meteorologists, oceanographers and members of the merchant marine use similar navigational definitions; however, differences in wind definitions are common. Platform-relative wind is defined as the wind vector measured relative to the ship. The only variation between meteorologists, oceanographers and the

merchant marine occurs with the platform-relative wind direction. Both meteorologists and the merchant marine report the direction from which the wind is blowing, while oceanographers usually report a direction to which the wind is blowing (Table 1).

In measuring a platform-relative wind, the zero reference angle is defined as the angle between the zero line of the wind vane and the bow of the vessel (measured clockwise from the bow). A zero reference angle becomes necessary when operational constraints preclude orienting the wind vane's zero line to the bow. For example, when mounting a vane high on a mast spar, it may be easier to orient the vane's zero line along the spar and then measure the angle between the spar and the fore to aft centerline of the vessel (hereafter, this direction will be referred to as the bow). Furthermore, many wind vanes have a potentiometer dead space at 360° (Fritschen and Gay, 1979). In this case, orienting the vane with 180° toward the bow is practical since the majority of the platform-relative winds will be from the bow when the vessel is underway. The zero reference angle must be known to adjust the measured platform-relative winds to the ship's directional reference frame (i.e. bow = 0°). Wind vane installations are specific to each vessel or experimental design and must be known to correctly compute true winds.

The apparent wind is a wind vector with a speed referenced to the vessel and a direction referenced to true north. The apparent wind direction can be computed by adding the heading and zero reference angle to the platform-relative wind direction (the apparent wind speed equals the platform-relative wind speed). Meteorologists and the merchant marine again provide the direction from which the apparent wind blows while oceanographers typically record the direction to which the apparent wind blows (Table 1). The merchant marine also have an alternative definition for the apparent wind (Bowditch, 1984) which is identical to the meteorological platform-relative wind. The purpose of this second definition is not clear in the context of motor-powered vessels and leads to obvious confusion.

The true wind is generically defined as a vector wind with a speed referenced to the fixed earth and a direction referenced to true north. The meteorological definition of true wind (Table 1) references the direction from which the wind is blowing (Huschke, 1959), while oceanographers often reference the direction to which the wind is blowing (Hosom *et al.*, 1995). The merchant marine utilizes two true wind definitions: one identical to the meteorological definition and the other with the true wind direction reported relative to the bow of the ship (Bowditch, 1984). The authors' experience with WOCE data indicates that the lack of a standard true wind definition or documentation of a specific definition is partially responsible for large discrepancies found in automated true wind data and in bridge measurements reported primarily by Voluntary Observing Ships (VOSs) (Pierson, 1990; Wilkerson and Earle, 1990; Kent *et al.*, 1993).

B. PROBLEMS COMMON TO MARINE WIND MEASUREMENT

Additional problems with wind data from AWS-equipped R/Vs are related to the wind instrumentation, approximations regarding navigation data, and calculation methodology. The calibration, orientation (see zero reference angle above), and location of the wind sensor are all very important to true wind calculations. Ideally, wind sensors are located in a region where the airflow is not seriously distorted by the measurement platform. In practice, disturbance of the flow at the instrument location by upwind or downwind structures (i.e. flow distortion) can only be minimized. The entire structure of the vessel and the mounting platform cause some degree of flow distortion; thus, the primary concern is siting the anemometer in a region that minimizes flow distortion caused by these structures (Kahma and Leppäranta, 1981; Rahmstorf, 1989; Yelland *et al.*, 1994; Yelland *et al.*, 1998). Recommended wind sensor locations range from high on the main superstructure to far out ahead of the bow. The solution attempted on several vessels (e.g. R/V Wecoma, R/V Meteor) is to install multiple sensors and have an automated routine extract the data from the instrument best exposed to the wind.

Errors associated with the navigation assumptions are also troubling for true wind calculations. Three essential navigation parameters (COG, SOG and heading) must be accurately recorded. Clear definitions of which navigation

values have been measured are also essential. For example, simply reporting a 'course' is ambiguous and can easily be mistaken to mean either the direction in which the vessel is steered, the course made good (Bowditch, 1984), or the COG. Reporting only a vessel's 'speed' causes similar confusion because the speed could be referenced to the water or the earth. Furthermore, if the navigation sensors are not properly calibrated (Hartten, 1998), then use of the measurements in calculations will lead to erroneous true winds. Finally, some measure of the navigation data's quality is necessary since positions are frequently reported in the wrong hemisphere, over land, or at a distance too far removed from the previous position to represent realistic ship movement. Poorly calibrated, missing or incorrectly measured navigational parameters lead to errors in calculated true winds.

Finally, multiple methods for calculating a true wind are employed in a wide range of applications. For example, most merchant marine vessels use graphical calculators, whereas R/Vs often rely on a series of equations encoded in an AWS. In the absence of standard reporting, meteorologists, oceanographers and members of the merchant marine tend to calculate and report true winds in the convention most suited to their operational needs. True winds are routinely exchanged without an explicit statement of the recording convention or calculation methodology. As a result, the differences in calculations and definitions are not known to the user of the true winds.

3. METEOROLOGICAL TRUE WIND

For centuries, requirements for ship operation, and more recently operational weather forecasting, have relied on a knowledge of the meteorological true wind. The World Meteorological Organization (WMO) requires VOSs to report true winds in the meteorological sense (WMO, 1996). The authors recommend that the meteorological (first merchant marine) definition be used to record true winds on automated systems, including those on non-VOS ships. Alternatively, useful true winds can be computed if the recording convention is reported.

A. CORRECT COMPUTATION

Calculating the meteorological true wind from a moving vessel requires the observed wind to be adjusted for the mean horizontal motion of the ship. For example, consider a woman facing forward on the bow of a stationary ship on a calm day. If the ship starts to move forward, the woman will feel a fresh wind (the apparent wind) on her face. The wind induced by the ship's motion (M) must be removed from the apparent wind (A) to compute a meteorological true wind (T):

$$T = A - M \tag{1}$$

The apparent wind is calculated by adding the heading and zero reference angle to the platform-relative wind direction, thereby orienting the wind measured on the vessel to true north. The motion-induced wind has the same magnitude as the course vector (C) with the opposite sign:

$$M = - C \tag{2}$$

Note that C is the vector motion of the ship over the fixed earth (i.e. direction equals COG, magnitude equals SOG). From (1), a true wind results by adding the course vector to the apparent wind vector:

$$T = A - (- C) = A + C \tag{3}$$

In the example above, the breeze felt by the woman on the bow would be cancelled by the vector addition of the forward motion of the vessel.

The computation of a true wind is often misinterpreted as removing the ship's course vector from the apparent wind vector. This error causes a distinct stair-step pattern (Figure 1) in the incorrectly calculated true wind speed (red) that is associated with the ship's forward speed (black). In this case, the incorrectly calculated true wind speed differs from a correctly calculated true wind speed

(green) by up to 8 m s^{-1} when the vessel is moving at speeds of greater than 2 m s^{-1} . Similar stair-step patterns occur in true wind data when other 180° errors are recorded in the platform-relative wind data (e.g. failure to report an oceanographic convention or a wind vane installed with the zero reference toward the stern). In general, a 180° error yields wind speeds that differ from the correct true wind speed by less than or equal to double the ship's speed.

The computation of meteorological true winds using an automated system requires that the vector equations be broken down into components. A detailed methodology is presented in Appendix A. Appendix B contains techniques to convert between meteorological, oceanographic and merchant marine conventions.

B. REQUIREMENTS FOR PRACTICAL APPLICATION

In many applications, from flux calculations to data ingestion in a general circulation model, it is necessary to have more information than the true wind speed and direction. For example, many applications require that the wind speed be adjusted to the meteorological standard height of 10 m above the surface. Other applications require winds relative to the surface current (e.g. scatterometry, stresses, and forcing of ocean models), while meteorological forecasts require earth-relative winds. The calculation of surface fluxes (of momentum, sensible heat and latent heat) and atmospheric stability require additional observations including the air temperature, the skin temperature of the water (approximated by the near surface temperature), and a measure of the humidity (Liu *et al.*, 1979). Observations of pressure are also useful to convert typical humidity measurements to specific humidity, which is used in height adjustments and flux calculations.

In recent years, the influence of sea state on fluxes and drag coefficients has become of interest (Smith *et al.*, 1992; Donelan *et al.*, 1997; Bourassa *et al.*, 1999). There is some controversy regarding the dependence on sea state. Several flux parametrizations require wave age or the phase speed of the dominant waves (e.g. Smith *et al.*, 1992; Bourassa *et al.*, 1999). Recently, the direction of the wind relative to the direction of wave propagation has been shown to have a large impact on the surface stress and drag coefficients (Donelan *et al.*, 1997; Bourassa *et al.*, 1999).

Essential metadata, such as the height of the sensors, should be recorded for use in height adjustment and the calculation of fluxes. In theory, the height of the temperature and humidity measurements must be the same, but these can differ from the height of the anemometer (Liu *et al.*, 1979). In practice, the height of the temperature and humidity observations has little influence on the height adjustment of winds; however, these heights can have a serious impact on the calculation of fluxes (e.g. stress and latent heat). In most cases, the lack of metadata prevents the accurate calculation of surface fluxes. One of the most common errors in ten-metre wind speed is due to the incorrect specification of anemometer heights. In several cases this height was given relative to the deck rather than

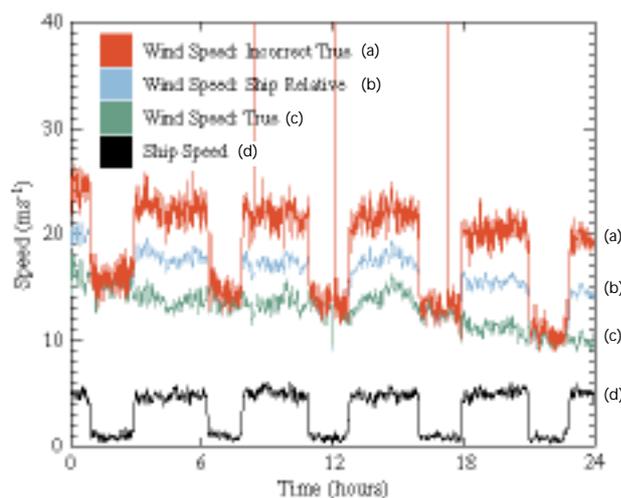


Figure 1—Example of accurate true wind calculation (c) vs. incorrect calculation (a) for the R/V Knorr. Note that both the platform-relative wind (b) and the incorrect true wind have a signal of the ship's earth-relative speed (d).

relative to the water surface (since the waterline of vessels changes according to the load, this error is understandable). Ideally, data records would include the height of the deck above sea level; however, such information is available from only very few highly specialized R/Vs.

4. EVALUATING THE QUALITY OF WIND DATA

Our experience has shown that missing data due to instrument malfunctions, encoding errors, approximations and oversights are common occurrences in automated data. Techniques to retrieve useful true wind information from these data sets are discussed in section 5 of this paper. The application of quality control procedures to identify problems is an essential first step. In this section, the quality control methods discussed include automated and visual inspections for erroneous data values and the identification of errors caused by a vessel's acceleration. A brief note is included concerning the unavoidable problem of flow distortion. After identifying problems with wind and navigation data, techniques for estimating true wind (section 5) can be applied to incomplete data sets.

A. IDENTIFYING ERRONEOUS VALUES

WOCE-MET utilizes a two-step process to quality control both true wind data and the variables necessary to calculate a true wind (Smith *et al.*, 1996). The first step is automated and identifies erroneous ship positions and physically unrealistic observations. A position check verifies that the latitude and longitude values are over water, while a speed check verifies that the vessel has not moved forward at a rate greater than 15 m s^{-1} . A range check of realistic wind directions (0 to 360°) and wind speeds ($< 40 \text{ m s}^{-1}$) is also performed. This latter check may highlight realistic extreme winds; thus, WOCE-MET personnel visually verify all flags added by the automated quality control.

Visual inspection of the data, though time consuming, is essential. The analyst adds flags for spikes, known instrument malfunctions, discontinuities, and values that are highly inconsistent with the surrounding trend. This latter contingency requires knowledge of the behaviour of wind data from vessels and is therefore subjective. Automated tests for discontinuities and spikes are available (Vickers and Mahrt, 1997), but we find visual inspection to be adequate. Based on 82 ship months of automated meteorological true winds, the two-level quality control applies flags to an average of 5 per cent of wind speeds and 6 per cent of wind directions. On some vessels, the visual inspection determines that all true wind directions and speeds are incorrect. Removing or correcting these flagged true wind values is essential before performing any application using the data.

The two-level quality control employed by WOCE-MET has proven invaluable. For two of the four vessels reporting all values required to compute true winds, the visual inspection allowed the analyst to determine that the platform wind direction was reported opposite the desired meteorological direction. When problems of this type were located, the platform wind was corrected and new meteorological true winds were calculated. The impact of fixing 180° errors was evident (Table 2) when the true wind values were compared to independent wind measurements from the NASA Scatterometer (NSCAT) (Bourassa, *et al.*, 1997). Table 2 reveals that the correction of the 180° error decreased the root-mean-square (rms) difference between collocated (within 25 km and 20 minutes) NSCAT and ship winds by 44 per cent for speed and by 33 per cent for direction. The correlation coefficient for collocated wind speed improves by 74 per cent. Visual inspection of the wind data is necessary and in some cases leads to a much larger set of useable wind values.

Table 2—Rms differences between collocated ship and NSCAT true winds for ship data with a 180° error in the platform wind and for the corrected true wind. Also presented is the improvement in the correlation coefficient for collocated wind speeds.

<i>True wind</i>	<i>Rms wind speed difference</i>	<i>Rms wind direction difference</i>	<i>Correlation coeff. for wind speed</i>
with 180° error	3.2 m s ⁻¹	21°	0.51
Corrected	1.8 m s ⁻¹	14°	0.89
Percentage change	-44	-33	74

B. AVERAGING TECHNIQUES AND ACCELERATION PROBLEMS

The choice of averaging techniques impacts the accuracy of true winds. Ideally, observations (platform-relative winds) would be made over short intervals and used to calculate true winds corresponding to those times. The true winds would then be averaged and stored. At this time, observational equipment and data processing are equal to the task (e.g. the TAO buoy array; Hayes *et al.*, 1991); however, this ideal is rarely achieved.

The averaging time for platform-relative winds should be sufficiently short so that navigational and ship-relative wind observations are approximately constant. The size of averaging periods depends on accuracy requirements and operational constraints for the vessel. For example, R/Vs spend a relatively large fraction of their operating time accelerating or decelerating. It will be shown that these changes in velocity can be identified in one-minute averages; therefore, shorter averaging times (perhaps <10 seconds) are recommended for the navigational parameters and platform-relative winds. The requirements for storage and post processing could be copious; therefore, we recommend that this averaging be processed by the shipboard instruments and that an average of the true winds (over 15-300 seconds) be recorded. A further advantage of averaging true winds as vectors, rather than speeds and directions, is the elimination of problems with the 360-0° breakpoint. At this time, such procedures are rarely implemented.

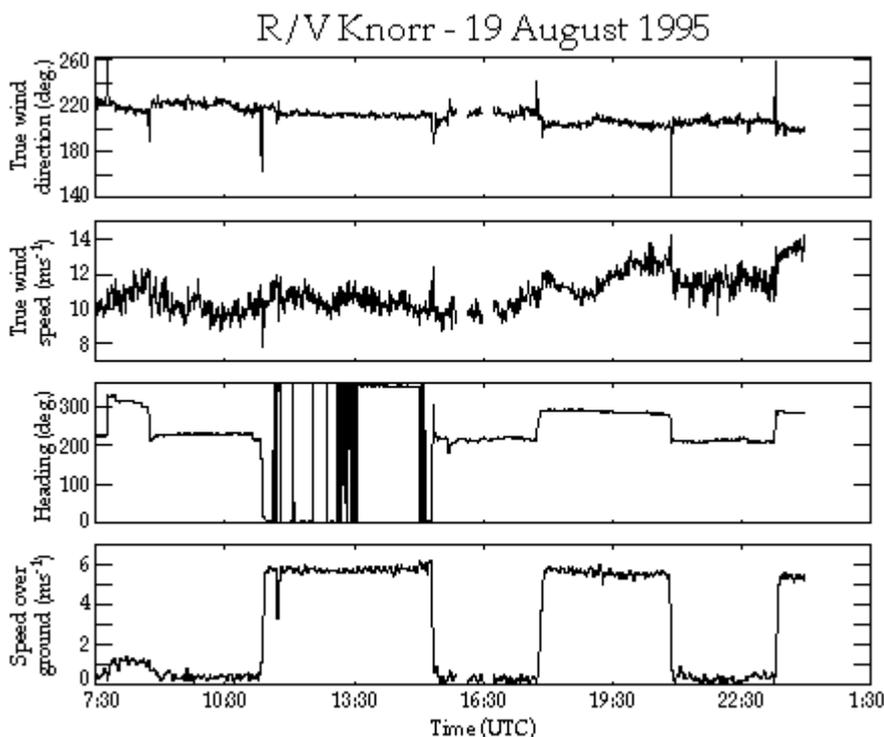
Typically, platform-relative winds and the navigational parameters utilized in the true wind calculations (section 3a and Appendix A) are averaged over intervals ranging from one minute to one hour. Since the true wind equations are nonlinear, they are accurate only when all the input parameters are approximately constant over the averaging period. When appropriate averaging cannot be applied, and the observations are too variable, the true winds should be flagged as suspect. One noticeable and regular manifestation of this problem occurs when R/Vs accelerate or decelerate. The impact of changing ship velocities is examined for an Improved METeorology (IMET) system (Hosom *et al.*, 1995) on the R/V Knorr, which records platform-relative winds in one-minute intervals. These acceleration errors manifest themselves as spikes in the true wind speed and direction data (Figure 2). The magnitude of the error in individual calculations is dependent on the rate of acceleration; however, for the R/V Knorr the spikes can approach 2 m s⁻¹ and 60°.

We have found, empirically, that quality control criteria can be based on the standard deviation of the ship's velocity (σ_v) determined from one-minute observations within a longer averaging period (six minutes in the following example):

$$\sigma_v = \left\{ \frac{1}{N} \sum_{i=1}^N \left[(U_i - \bar{U})^2 + (V_i - \bar{V})^2 \right] \right\}^{\frac{1}{2}} \quad (4)$$

where N is the number of observations and the overbar indicates averages of these N observations. For many applications the uncertainty due to acceleration is relatively small and can be ignored. However, satellite measurements of the near surface winds by NSCAT are sufficiently accurate (Bourassa *et al.*, 1997) that this additional uncertainty is apparent when comparing ship-based winds to remotely-sensed winds. The impact of this criterion is shown in the mean and rms differences between winds from NSCAT and the R/V Knorr. Without this criterion, there are 18 collocations (closest observations within 25 km and 20 minutes) with a mean difference (satellite minus ship) of -0.8 m s⁻¹, and an rms difference of 2.0 m s⁻¹. When observations with $\sigma_v > 1.0$ m s⁻¹ (12 collocations with accelerations that are considered too rapid and prolonged) are flagged and removed, the mean difference changes to 0.55 m s⁻¹, and the rms difference drops to 1.3 m s⁻¹. These findings are consistent with an assessment of the NSCAT-1 model function in comparisons with the NDBC buoys by M. Freilich and R.S. Dunbar (1997, personal communication) and the TAO buoys by K. Kelly (1997, personal communication). The change in the mean is statistically significant, corresponding to 4.3 standard deviations of the mean. The almost 50 per cent overestimation of the rms difference, prior to this quality control

Figure 2—Spikes that occur in true wind direction and speed caused by the acceleration of the vessel. Displayed are the true wind direction and speed, heading, and speed over the ground from the automated weather system on the R/V Knorr (0730 to 2359 UTC, 19 August 1995).



criterion, shows that there are applications where changes in the vessel's velocity can result in substantial averaging-related errors in the calculated true winds.

C. FLOW DISTORTION

Another problem that occurs with ship-based winds is flow distortion. Structures (i.e. the entire ship, and to a lesser extent the measuring device) cause air to deviate from the path it would take if the structures were not present. Flow distortion occurs in the wake of structures, around structures and upwind of structures. The resulting change in wind characteristics (speed, direction and the variation of these quantities) is highly dependent on the shape of the vessel, instrument position, and wind direction relative to the vessel's heading. Recently, computational fluid dynamics has been successfully applied to correct for the impacts of flow distortion (Yelland *et al.*, 1998). Other techniques are relatively simple, however, they are much more crude. The range of directions, over which the influence of flow distortion is a relatively strong function of platform-relative wind direction, can be estimated by binning the wind speed as a function of this direction (Thiebaut, 1990). We have found a similar result with the variation in the wind speed. However, neither of these approaches indicates the impact of the flow distortion or the angles at which the impact is a minimum. These techniques can only be used to isolate angles at which the impact of flow distortion is approximately constant, which can be advantageous for data analysis.

5. ESTIMATING TRUE WINDS FROM INCOMPLETE DATA

Incomplete observations from 16 of 20 studied vessels left only four with all values required to calculate meteorological true winds (Table 3). Consequently, we investigated methods for estimating meteorological true winds when some of the navigation parameters were missing. The two most common occurrences of missing navigation data are vessels reporting only a COG and SOG (no heading), or a heading and SOW_{FA} (no COG; Table 3). In these cases, if the platform-relative winds and zero reference angle are known, estimates for the true winds can be made; however, the underlying assumptions can lead to serious errors. Empirical studies reveal the conditions under which these estimations are practical.

If the heading is missing, a true wind can be estimated by replacing the heading with the COG. This estimate is hereafter called a course-estimated wind. Thus, the apparent wind direction is calculated by summing the COG angle, the zero reference angle and the platform-relative wind direction. The accuracy of

Table 3—Based on 20 vessels equipped with automated wind systems, the number that report (•) the parameters needed for computing a meteorological true wind or an estimate of the true wind (heading or course missing).

<i>Platform relative wind</i>	<i>Heading</i>	<i>Course over ground</i>	<i>Speed over ground</i>	<i>Speed over water</i>	<i>Number of ships reporting</i>
•	•	•	•	-	4
•	-	•	•	-	8
•	•	-	-	•	6

this estimate is questionable at low ship speeds where the course-estimated wind direction (1) in Figure 3 (a) deviates wildly from the actual true wind (2).

The range of SOG where the course-estimated winds are valid can be determined empirically and depends upon the vessel and its region of operation. As an example, we determined this range using two vessels that reported all necessary values to WOCE-MET. Differences between the course-estimated and true wind direction were computed using quality controlled observations from the R/V Thompson (8.9 months) and the R/V Knorr (4.7 months). The direction differences were separated into 0.5 m s^{-1} SOG bins and an rms difference was calculated for each bin. Rms differences and the number of values in each SOG bin, from 0 to 9 m s^{-1} , are presented for the R/V Thompson and R/V Knorr (Figure 4). For low ship speeds (SOG < 2 m s^{-1}), both vessels exhibit an rms difference of near or greater than 60° . Direction differences drop below 20° when the SOG exceeds 2.5 m s^{-1} for the R/V Thompson and 4.0 m s^{-1} for the R/V Knorr. Determining a

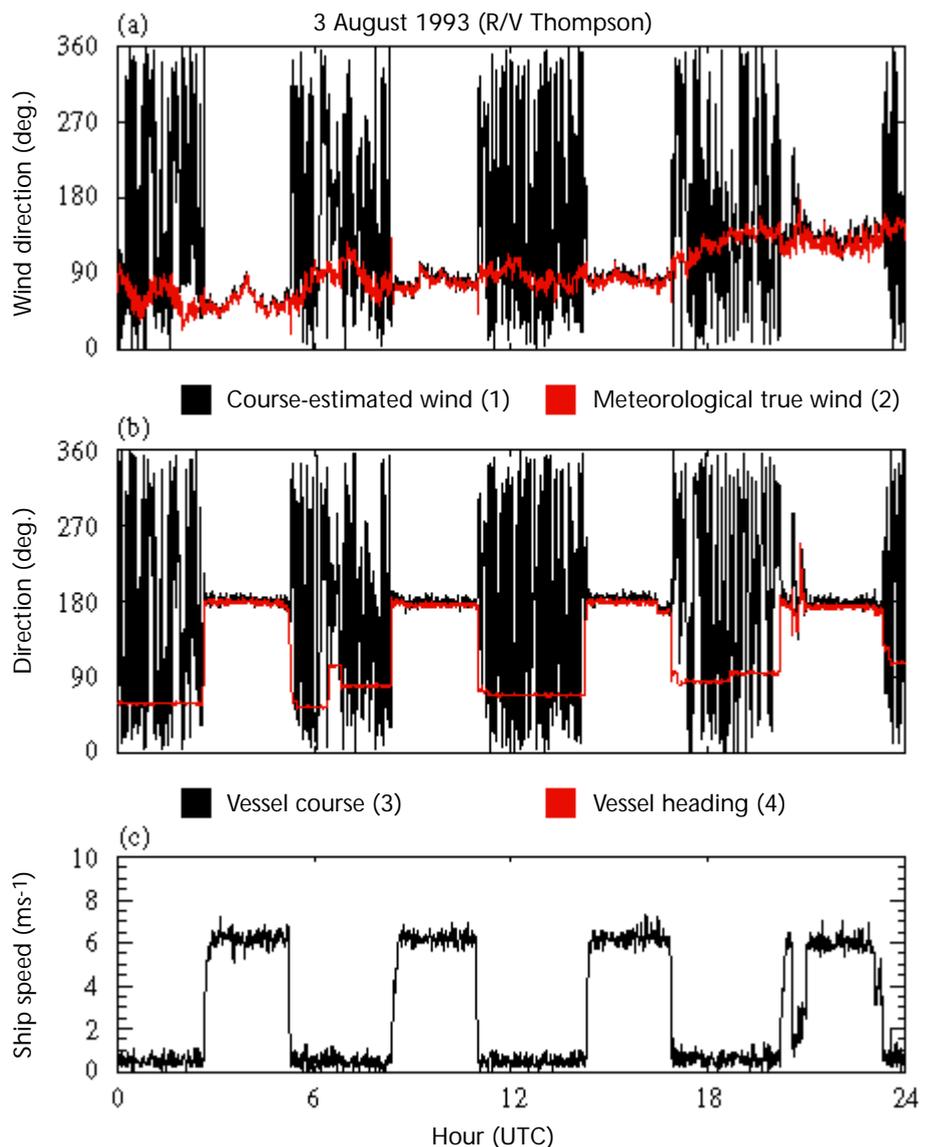


Figure 3—Time series plots of (a) course-estimated (1) vs. meteorological true (2) wind direction, (b) vessel course (3) vs. heading (4), and (c) vessel speed over the ground from the R/V Thompson (3 August 1993).

threshold SOG for which the rms difference is within an acceptable range is user dependent. For an rms wind direction difference of less than 10°, the threshold SOG is 3.5 m s⁻¹ for the R/V Thompson and 5 m s⁻¹ for the R/V Knorr (Figure 4). In summary, the primary limitation of the course-corrected wind estimates is that they are sensitive to the SOG; becoming unreliable at low ship speeds.

Inaccuracies in the course-estimated winds are directly related to measuring only SOG and COG without a heading. Eight of twenty studied vessels relied solely on single receiver GPS systems to measure their geographical position and to provide values of SOG and COG. A single receiver GPS is not designed to measure heading; therefore, it cannot always estimate the heading with sufficient accuracy. This problem is exaggerated at low ship speeds when current and wind forces on the ship can cause large differences between heading and COG (Figure 3 (b) and (c)). As a result, the true winds reported by these eight vessels are course-estimated winds. Furthermore, errors in course-estimated winds are increased if the latitude and longitude are not recorded to at least the fourth decimal place. The ability to measure the orientation of the vessel using only GPS technology can be improved using a multiple receiver GPS, but for vessels with single receiver GPS we recommend the addition of a gyrocompass to record the heading.

When a vessel relies on navigation without the aid of technology referenced to the fixed earth, it is common practice to measure only the heading of the vessel and the SOW_{FA}. When only heading and SOW_{FA} are measured, an earth-relative wind cannot be computed. Instead, an estimate referenced to the water can be created by replacing the course vector in Equation (3) with a heading vector, H, where the |H| equals SOW_{FA} and the direction of H is the direction in which the bow is pointing (referenced to true north). Unlike the course-estimated winds, frequency diagrams (not shown) of this heading-estimated wind minus the meteorological true wind reveal differences with no dependence on forward ship speed. Instead, the heading-estimated wind deviates from a true wind only when the SOW_{FA} and SOG are different from one another. A time series plot for 12 hours of wind data from the R/V Knorr illustrates the differences that can occur (Figure 5). In this case, a 2 m s⁻¹ difference in the SOW_{FA} and SOG (Figure 5 (b)) results in an average direction error of 25° (Figure 5 (a)). Variations between the SOW_{FA} and SOG are related to currents.

In summary, an examination of a total of 13.6 ship months of automated observations from two vessels shows that, when the computation of a meteorological true wind is not possible, heading-estimated winds are superior to course-estimated winds. The accuracy of the heading-estimated winds is limited by the difference between SOW_{FA} and SOG. When these speeds are not significantly different, the heading-estimated and true wind directions are nearly identical. When an operator only records the SOG and COG, the potentially large

Figure 4—Rms differences of course-estimated minus true wind direction (filled triangles) from the R/V Thompson and R/V Knorr. Rms differences are calculated for 0.5 m s⁻¹ bins of the vessel's speed over the ground using the absolute value of the wind direction differences (i.e. range 0-180°). The number of values in each bin (open squares) is presented in units of a thousand.

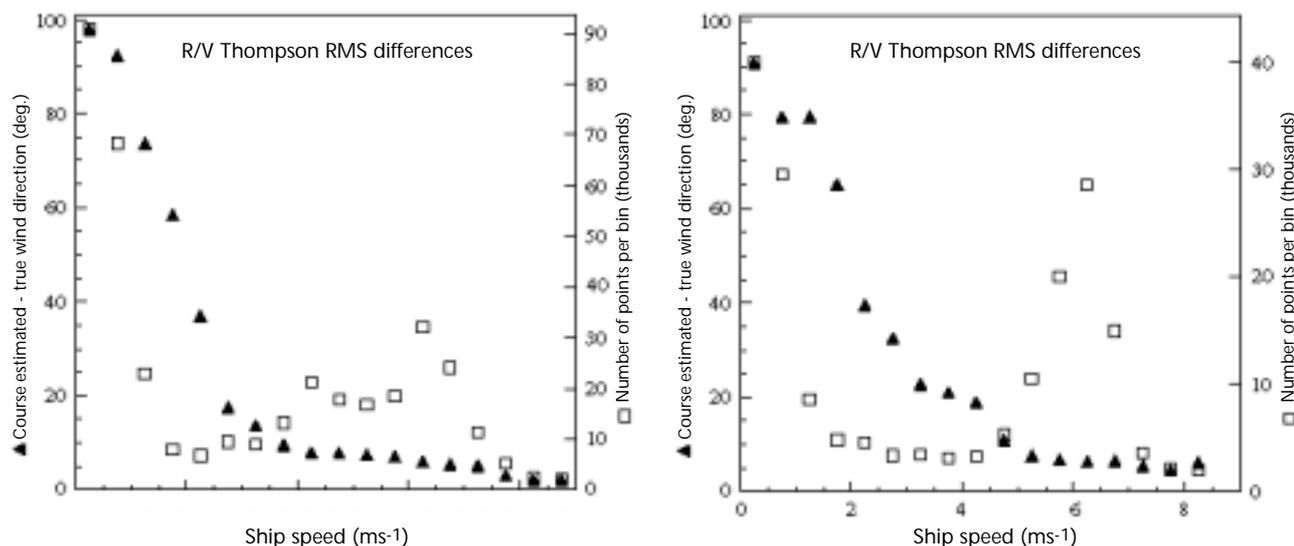
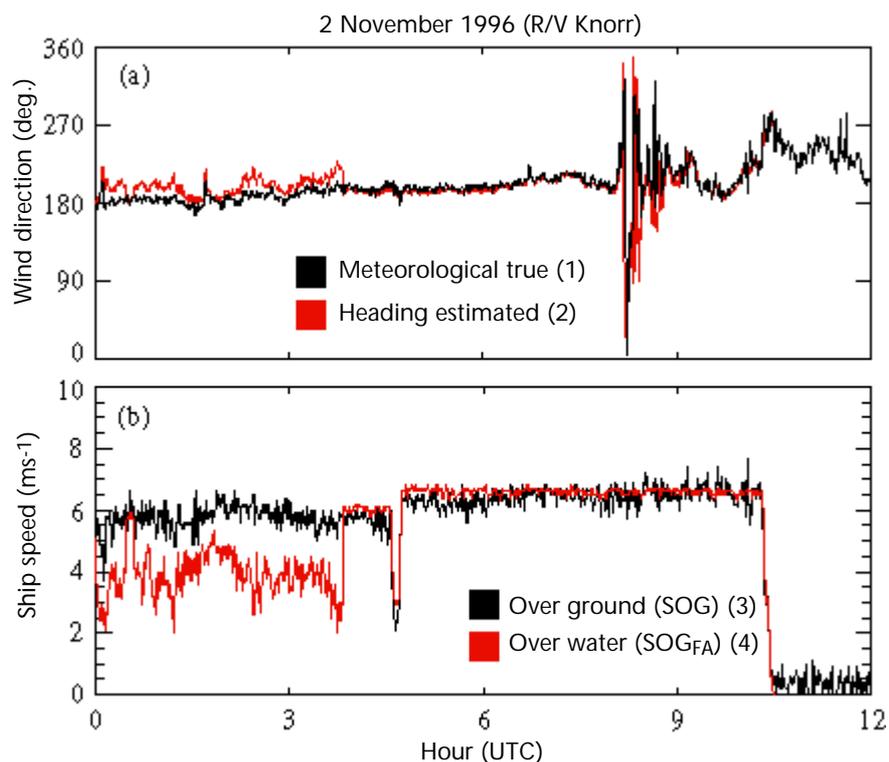


Figure 5—Time series plots of (a) heading-estimated (2) vs. meteorological true (1) wind direction and (b) vessel speed over the water (4) vs. the speed over the ground (3) from the R/V Knorr (2 November 1996). Note that the speed over the water drops out after 10:30 UTC due to an instrument malfunction.



differences between the COG and heading, at low ship speeds, result in large differences between course-corrected and true winds.

6. SUMMARY AND RECOMMENDATIONS

Problems in computing true winds from automated systems have been identified and solutions are demonstrated. Principal problems include confusion related to inconsistencies in definitions for true wind used by meteorologists, oceanographers and the merchant marine, and the lack of standard reporting of both wind and navigation measurements (or the convention used). The primary recommendation is to set a standard for reporting the six values needed to compute a true wind: COG, SOG, heading, zero reference, and platform-relative wind direction and speed. Additional metadata, especially the height of the wind sensor relative to the water surface, must also be reported.

Accurate meteorological true winds result from the vector sum of the ship's motion relative to the fixed earth and the apparent wind. Details of this calculation are outlined in Appendix A. Conversions from the meteorological true wind to oceanographic and merchant marine definitions are detailed in Appendix B. The true wind calculations and conversions presented can be applied to AWS and non-automated wind measurements.

True winds must be quality controlled before application to identify errors. At WOCE-MET, a two-level quality control system comprised of an automated pre-processor and a detailed visual examination has proven effective in identifying both minor (e.g. out of range values, spikes, ship acceleration) and major errors (e.g. incorrectly oriented platform-relative wind).

When dealing with incomplete data sets (e.g. approximately 80 per cent of examined AWS data), true winds can be estimated within determinable limitations. A better estimate for a wide range of forward ship speeds can be obtained when a heading and SOW_{FA} are measured, rather than an estimate derived from SOG and COG when no heading is available. The heading-estimated wind varies from a true wind only when the SOW_{FA} and SOG are significantly different. The uncertainty in course-estimated winds has a strong dependency on the forward ship speed. Empirical studies show course-estimated wind directions to be unreliable ($rms > 60^\circ$) when the $SOG < 2.0 \text{ m s}^{-1}$. Useful estimates can be obtained at higher ship speeds; however, the threshold SOG depends upon the ship, the vessel's operating area and the users desired level of uncertainty.

The following recommendations are made for future automated observing systems, thereby avoiding the need to estimate true winds. The standard set of measurements needed to compute a true wind (i.e. SOG, COG, heading, wind relative to the vessel) must all be logged at the same frequency as the standard meteorological variables. Averaging should be applied to true winds calculated from shorter term (0.5 to 10 sec.) observations. Essential metadata (e.g. zero reference and instrument heights) must be reported. When it is essential that measurements be collected without losing data, redundancy should be planned for both the navigation and wind measurements. When an instrument fails in a redundant system, alternate measurements can be used in the computation of true wind, or an estimate of the true wind can be created. For example, three years of AWS wind data on one studied vessel were lost due to the failure of a navigational compass in the wind sensor. If the ship's gyrocompass heading had been archived in the meteorological data stream, this loss could have been avoided. If the marine community utilizes the techniques and recommendations herein, a superior quality of high temporal resolution true wind observations can be computed from automated platforms on vessels at sea.

ACKNOWLEDGEMENTS

The authors wish to thank Ms Leslie Hartten for her comments on ship wind measurements, Ms Masha A. Medvedeva for her contributions to the wind analysis, and the anonymous reviewers for their constructive comments. The authors also acknowledge the professional support of Mr David M. Legler. We further express our gratitude to the personnel both at sea and on shore who provided the data for this study.

COAPS receives its base funding from the Physical Oceanography Section of the Office of Naval Research. WOCE-MET and this research are funded by the Physical Oceanography Section of the National Science Foundation, grant OCE-314515, and from the NASA JPL NSCAT Project, contracts 957649 and 980646.

APPENDIX A

DETAILS FOR AUTOMATED CALCULATION OF TRUE WINDS

This appendix is a tutorial containing algorithms for calculating meteorological true winds from ship observations. The mathematics and all necessary variables (Table A1) are discussed and an example provided.

All calculations are performed in the mathematical coordinate system which has an angle of zero degrees on the positive x-axis with angles increasing in a counter-clockwise direction. Each vector direction, originally defined using the meteorological conventions (Table 1), is converted to mathematical coordinates prior to other calculations (see Table B1 to convert from other conventions). Primes (') denote values in mathematical coordinates.

Table A1—Variables needed for true wind vector (T) calculation.

<i>Parameter</i>	<i>Type</i>	<i>Symbol</i>	<i>Direction reference frame</i>	<i>Velocity reference frame</i>
Ship heading	scalar	h_{θ}	true north	--
Ship course over ground	vector	C	true north	fixed earth
Platform-relative wind (direction <i>from</i> which wind is blowing)	vector	P	zero reference on ship	ship
Zero reference angle for platform-relative wind	scalar	R_{θ}	bow of ship	--

Platform-relative winds (**P**) and navigational data are used to calculate apparent (**A**) and true (**T**) winds. The direction of the apparent wind in the mathematical coordinates is:

$$A_{\theta}^N = 270^{\circ} - (h_{\theta} + R_{\theta} + P_{\theta}) \quad (\text{A1})$$

where h is the vessel's heading, R is the zero reference, and the subscript θ designates an angle. The magnitude of **A** is the same as the magnitude of **P**. Use of the heading instead of the COG in Equation (A1) is essential because the bow is rarely oriented in the direction of ship motion over the fixed earth. As an example, consider the bow of a ship oriented directly to the east ($h_{\theta} = 90^{\circ}$). If there is either a strong current or wind from the north, then the vessel will be pushed to the south, resulting in a COG greater than 90° .

Most ships utilize the bow as the zero reference for the platform-relative wind, but there are exceptions to this practice. When another point on the ship is used as a zero reference for the wind vane, the angle between this reference and the bow (R_{θ} in Figure A1) must be included in Equation (A1) to correctly calculate the apparent wind direction.

The COG of the vessel (C_{θ}) in mathematical coordinates is:

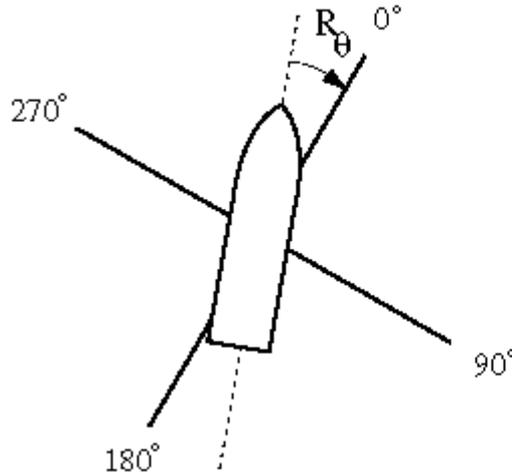
$$C_{\theta}^N = 90^{\circ} - C_{\theta} \quad (\text{A2})$$

The true wind is then computed by summing the vector components of the apparent wind and ship motion:

$$T_u = T_u^N = |A| \cos(A_{\theta}^N) + |C| \cos(C_{\theta}^N) \quad (\text{A3a})$$

$$T_v = T_v^N = |A| \sin(A_{\theta}^N) + |C| \sin(C_{\theta}^N) \quad (\text{A3b})$$

Figure A1—Platform-relative coordinate system with a zero reference angle, R_θ , not oriented to the bow.



where positive T_u and T_v are the eastwards and northwards components of the true wind in the earth reference frame. The true wind speed (T) and direction (T_θ) can then be calculated:

$$(A4)$$

$$|T| = (T_u^2 + T_v^2)^{1/2}$$

and

$$T_\theta = 270^\circ - \text{atan} \left[\frac{T_v}{T_u} \right] \quad (A5)$$

The 270° in Equation (A5) converts the value of $\text{atan} (T_v T_u^{-1})$ to a direction *from* which the wind is blowing (meteorological convention) in the earth coordinate system. For Equation (A5) to return a correct angle, the atan function must have a range from -180° to 180° to determine the vector's trigonometric quadrant (e.g. the FORTRAN 'atan2' function). Also, any program using Equation (A5) must have a check to avoid dividing by zero.

As an example calculation consider a ship (Figure A2) with a heading (h_θ) of 30.0° and a COG (C_θ) of 45.0° both referenced to true north (0° in a fixed earth reference frame). The vessel is travelling at a SOG (C) of 5.0 m s^{-1} . The platform-relative wind, with the bow as the zero reference angle ($R_\theta = 0.0^\circ$), is blowing *from* a direction (P_θ) of 250.0° with a magnitude (P) of 10.0 m s^{-1} . The conversion to mathematical coordinates using Equations (A1) and (A2) results in $A_\theta^N = 350.0^\circ$ and $C_\theta^N = 45.0^\circ$. Computing the true wind components using Equations (A3a) and (A3b) gives $T_u = 13.4 \text{ m s}^{-1}$ and a $T_v = 1.8 \text{ m s}^{-1}$. The meteorological true wind

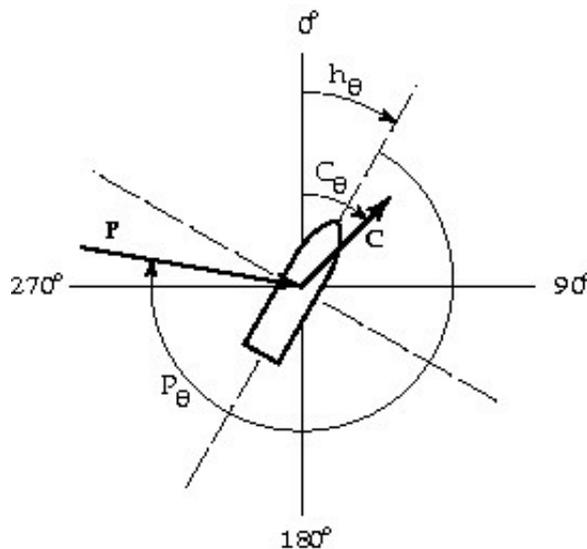


Figure A2—Schematic representation of the vectors and angles involved in the true wind problem. See text for explanation of symbols.

Vessel course over the ground (deg.)	Vessel speed over the ground (m s ⁻¹)	Vessel heading (deg.)	Platform wind direction (deg)	Platform and apparent wind speed (m s ⁻¹)	Apparent wind direction (deg.)	True wind direction (deg.)	True wind speed (m s ⁻¹)
					0.0		
0.0	0.0	90.0	5.0	90.0	90.0	5.0	
0.0	0.0	90.0	90.0	5.0	180.0	180.0	5.0
0.0	5.0	0.0	360.0	5.0	360.0	0.0	0.0
0.0	5.0	0.0	0.0	0.0	0.0	180.0	5.0
180.0	5.0	180.0	180.0	5.0	360.0	360.0	10.0
90.0	5.0	90.0	90.0	5.0	180.0	225.0	7.1
90.0	5.0	45.0	135.0	5.0	180.0	225.0	7.1
225.0	5.0	225.0	270.0	5.0	135.0	90.0	7.1
270.0	3.0	270.0	90.0	4.0	360.0	36.9	5.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table A2—Sample input and output for the true wind calculation. For simplification, the zero reference is the bow of the vessel ($R_0 = 0^\circ$). The table is divided into input (light face) and output (bold) values. Note that both the apparent and true wind directions are referenced to true north, and all wind directions are angles from which the wind is blowing. Also, the WMO convention is utilized for calm (direction = 0°) and north winds (direction = 360°).

speed from Equation (A4) is 13.5 m s^{-1} and the true wind direction, Equation (A5), is blowing from 262.3° .

Table A2 provides sample input and the output that should be returned from a meteorological (and 1st merchant marine) true wind algorithm. Any algorithm used to calculate true winds should duplicate these results. FORTRAN, C, and IDL (Interactive Data Language) routines for computing meteorological true winds are available at the URL: <http://www.coaps.fsu.edu/WOCE>.

APPENDIX B

CONVERSIONS BETWEEN DIFFERENT CONVENTIONS

As previously stated, the meteorological, oceanographic and merchant marine conventions are different when defining true wind vectors. In the meteorological convention, wind direction is defined as the direction from which the wind is blowing in the earth reference frame. The merchant marine has an identical convention, but they also define a true wind with a direction referenced to the bow of the ship. In the oceanographic convention, wind direction is defined with a direction toward which the wind is blowing. This confusion presents problems when using true wind data. Table B1 provides the conversions between each of the conventions. After all direction conversions, the modulus of the returned value must be taken with respect to 360° to ensure a direction between 0° and 360°.

Table B1—Conversion table to change between conventions for apparent and true wind directions. Note that the merchant marine utilizes two definitions for apparent and true wind (Table 1). The first merchant marine definitions are identical to those used in meteorology and should be treated identically (merchant marine (1) = Meteorology) when using this conversion table. The merchant marine listings in this table refer to the second (2) apparent and true wind definitions from Table 1. After each direction conversion, a modulus with respect to 360° must be performed to ensure a value in the range of 0° to 360°.

<i>Given</i>	<i>Meteorology oceanography</i>	<i>Oceanography meteorology</i>
Apparent wind	Add 180°	Add 180°
True wind	Add 180°	Add 180°
<i>Given</i>	<i>Meteorology merchant marine (2)</i>	<i>Merchant marine (2) meteorology</i>
Apparent wind	Subtract heading of ship	Add heading of ship
True wind	Subtract heading of ship	Add heading of ship
<i>Given</i>	<i>Merchant marine (2) oceanography</i>	<i>Oceanography merchant marine (2)</i>
Apparent wind	Add heading of ship and add 180°	Subtract heading of ship and add 180°
True wind	Add heading of ship and add 180°	Subtract heading of ship and add 180°

REFERENCES

- Bourassa, M.A., M.H. Freilich, D.M. Legler, W.T. Liu and J.J. O'Brien, 1997: Wind observations from new satellite and research vessels agree. *EOS Trans. Amer. Geophys. Union*, 78(51), 597, 602.
- Bourassa, M.A., D.G. Vincent, W.L. Wood, 1999: A flux parameterization including the effects of capillary waves and sea state. *J. Atmos. Sci.*, 56, 1123-1139.
- Bowditch, N., 1984: *American Practical Navigator: An Epitome of Navigation*. DMA Stock No. NVPUB9V1, LC no. VK555.A48, Defense Mapping Agency, Washington, D.C., 1414 pp.
- Donelan, M.A., W.M. Drennon and K.B. Katsaros, 1997: The air-sea momentum flux in conditions of wind sea and swell. *J. Phys. Oceanogr.*, 27, 2087 - 2099.
- Fritschen, L.J., and L.W. Gay, 1979: *Environmental Instrumentation*. Springer-Verlag, 216 pp.
- Hartten, L.M., 1998: Reconciliation of surface and profiler winds at ISS sites. *J. Atmos. Oceanic Technol.*, 15, 826-834.
- Hayes, S.P., L.J. Mangum, J. Picaut, A. Sumi and K. Takeuchi, 1991: TOGA-TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bull. Amer. Meteor. Soc.*, 72, 339-347.
- Hosom, D.S., R.A. Weller, R.E. Payne and K.E. Prada, 1995: The IMET (Improved Meteorology) ship and buoy systems. *J. Atmos. Oceanic Technol.*, 12, 527-540.
- Huschke, R.E. (ed.), 1959: *Glossary of Meteorology*. American Meteorological Society, 638 pp.
- Kahma, K.K. and M. Leppäranta, 1981: On errors in wind speed observations on R/V *Aranda*. *Geophysica*, 17, 155-165.
- Kent, E.C., P.K. Taylor, B.S. Truscott and J.S. Hopkins, 1993: The accuracy of Voluntary Observing Ships' meteorological observations - results of the VSOP-NA. *J. Atmos. Oceanic Technol.*, 10, 591-608.
- Liu, W.T., K.B. Katsaros and J.A. Businger, 1979: Bulk parameterization of air-sea exchanges of heat and water vapor including the molecular constraints at the interface. *J. Atmos. Sci.*, 36, 1722 - 1735.
- Pierson, W.J., Jr., 1990: Examples of, reasons for, and consequences of the poor quality of wind data for the marine boundary layer: Implications for remote sensing. *J. Geophys. Res.*, 95, 13313-13340.
- Rahmstorf, S., 1989: Improving the accuracy of wind speed observations from ships. *Deep-Sea Res.*, 36, 1267-1276.
- Smith, S.D., R.J. Anderson, W.A. Oost, C. Kraan, N. Maat, J. DeCosmo, K.B. Katsaros, K.L. Davidson, K. Bumke, L. Hasse and H.M. Cadwick, 1992: Sea surface wind stress and drag coefficients: the HEXOS results. *Bound.-Layer Meteorol.*, 60, 109-142.
- Smith, S.R., C. Harvey, and D.M. Legler, 1996: *Handbook of Quality Control Procedures and Methods for Surface Meteorology Data*. WOCE Report 141/96, COAPS Report 96-1, 56 pp. [Available from WOCE Data Assembly Center, COAPS, The Florida State University, Tallahassee, Florida, 32306-2840, USA].
- Thiebaux, M.L., 1990: Wind tunnel experiments to determine correction functions for shipborne anemometers. Canadian Contractor Report of Hydrography and Ocean Sciences 36, 57 pp. [Available from Bedford Institute of Oceanography, Halifax, Nova Scotia, B3J 2S7, Canada].
- Vickers, D. and L. Mahrt, 1997: Quality control and flux sampling problems for tower and aircraft data. *J. Atmos. Oceanic Technol.*, 14, 512-526.
- Wilkerson, J.C. and M.D. Earle, 1990: A study of differences between environmental reports by ships in the Voluntary Observing Program and measurements from NOAA buoys. *J. Geophys. Res.*, 95(C3), 3373-3385.
- World Meteorological Organization (WMO), 1996: *Guide to Meteorological Instruments and Methods of Observation*. WMO Rep. 8.
- Yelland, M.J., B.I. Moat, P.K. Taylor, R.W. Pascal, J. Hutchings and V.C. Cornell, 1998: Wind stress measurements from the open ocean corrected for airflow distortion by the ship. *J. Phys. Oceanogr.*, 28, 1511-1526.
- Yelland, M.J., P.K. Taylor, I.E. Consterdine and M.H. Smith, 1994: The use of the inertial dissipation technique for shipboard wind stress determination. *J. Atmos. Oceanic Technol.*, 11, 1093 - 1108.

QUALITY CONTROL IN RECENT AND PENDING COADS RELEASES

K. Wolter, S.J. Lubker and S.D. Woodruff, NOAA-CIRES Climate Diagnostics Center, Boulder, CO, USA

ABSTRACT In the context of the long-term monitoring of global climate, the Comprehensive Ocean-Atmosphere Data Set (COADS) offers the most complete marine surface data collection (1784-1997) currently available for global climate research. Its long duration and the international nature of its data sources necessitate particularly careful quality control (QC) measures. The final portion of COADS QC — flagging statistical outliers and removing them from the computation of areal averages — is referred to as 'trimming'. Based on a review of the trimming impacts of COADS Release 1 (1854-1979), recent COADS Releases 1a (1980-97), 1b (1950-79) and 1c (1784-1949) were modified to allow for a better representation of large-scale climate anomalies such as major ENSO events. This paper summarizes these changes in COADS QC, and discusses related trimming modifications for near-real-time products and future COADS Releases.

1. INTRODUCTION Preceding most land-based climate records, marine weather observations have been systematically archived for well over a century. The Comprehensive Ocean-Atmosphere Data Set (COADS) (Woodruff *et al.*, 1987) has made this data openly available for global climate research. The original COADS Release 1 (Slutz *et al.*, 1985) covered the world ocean from 1854-1979 with about 72 million marine reports. The most recent COADS update for 1980-97 (Release 1a; Woodruff *et al.*, 1993) adds almost 67 million unique reports to this data bank. Releases 1b and 1c (re-)processed data for the periods 1950-79 and 1784-1949, respectively, which included new data and other improvements (Woodruff *et al.*, 1998; and Woodruff *et al.*, this publication) compared to Release 1. While marine reports were originally derived from observations on board ships of opportunity, the last two decades have seen a steep increase in other sea-borne platforms, most notably moored and drifting buoys launched in international efforts like FGGE, EPOCS and TOGA (Woodruff *et al.*, this publication).

Although individual marine reports are available, the most commonly used products of COADS are monthly summaries for $2^\circ \times 2^\circ$ and, more recently, $1^\circ \times 1^\circ$ latitude-longitude boxes. These summary statistics include the monthly average and median for eight *observed* variables (including sea surface temperature (SST)) and a number of *derived* variables, such as relative humidity, wind stress and heat flux terms (Woodruff *et al.*, 1998).

Surface marine data deviate from land data in several ways. The number of observations per month is not constant in time and space. Marine observations originate from a variety of vessels plying a given ocean region, often from different countries and with different instrumentation, resulting in non-trivial sampling errors. On the other hand, oceanic surfaces are vastly more homogeneous in their physical properties than land surfaces, so they can be more reliably sampled with only a few measurements. In addition, the large heat capacity of water contributes to the high daily to seasonal persistence of SST.

Given the relative importance of each marine observation, comprehensive quality control (QC) procedures have been applied to COADS. In particular, the process of flagging and removing statistical outliers from the computation of 2° box averages is referred to as 'trimming' (Slutz *et al.*, 1985). Such outliers were originally defined as individual observations that reside outside the long-term median plus/minus 3.5 'standard deviations' (σ ; defined separately for positive and negative departures; see next section for details) for each 2° box. In the

standard version of COADS Releases 1a-c, the original (3.5σ) trimming limits were kept, but only ship data were admitted (as far as that can be ascertained). In the *enhanced* version, the trimming limits have been expanded to include all observations up to 4.5σ away from g , using most ocean surface-based platforms.

In observational data sets, the separation and trimming of statistical outliers from climate extremes is prone to errors. It can either happen that an extreme, but valid, observation is erroneously excluded (statistical Type I error), or that a 'bogus', erroneous observation (a 'true outlier') is included (statistical Type II error). The original COADS trimming procedure was explicitly designed to minimize Type II errors (Woodruff *et al.*, 1987) since these could distort average fields of affected variables. On the other hand, large climate anomalies entail observational distributions close to COADS trimming limits, meaning that the attempt to remove Type II errors through trimming may have led to Type I errors which artificially reduce climatic variability through the COADS record, as first discussed in Wolter *et al.* (1989) and summarized in Wolter (1997), the main topic of which is how to reconcile the competing objectives of trimming erroneous observations while preserving the authenticity of the climate record.

The current paper reflects the fact that most of the research in COADS QC revolves around SST. Nevertheless, many of our comments address more general trimming issues that apply to other COADS variables as well. A brief review of COADS QC is given in section 2. Evidence of trimming errors in Release 1 is presented in section 3. Alternative trimming procedures have been applied to COADS Releases 1a-c, and are summarized in section 4. Pending COADS QC issues are introduced in section 5.

2. COADS QUALITY CONTROL

COADS is intended as a database both for climate and weather studies. While individual records are available as a COADS product, so that synoptic studies can be performed on the complete, uncensored record, COADS monthly summaries provide the basis for many climate studies. For the latter, it is the main objective of COADS to record the climate history of the surface of the world ocean as faithfully as possible. This includes large-scale, extreme climate anomalies such as the 1982-83 ENSO event that lasted well over a month. In a nutshell, the philosophy of COADS QC is to include all observations that reflect typical conditions in a given 2° box, while trying to exclude brief weather extremes and erroneous observations that would seriously distort the mean of a poorly sampled month.

Monthly summaries in COADS Release 1 blend all surface-based marine observations in a single product, be they ship-, buoy-, or ice-based. However, ship data constitute the vast bulk of Release 1 before the 1970s (Table 3-1c, Slutz *et al.*, 1985). In contrast, Releases 1a-c distinguish between exclusive ship records in the *standard* set (as much as this could be ascertained; Woodruff *et al.*, 1993, 1998) and the blended, or *enhanced* set which continues to include a large variety of in situ observations. Otherwise, COADS monthly summaries do not differentiate among observational techniques or platforms. However, such 'metadata' are retained, when available, in COADS individual reports, e.g. the 'platform type' flag in Releases 1a-c.

In general, COADS data are not adjusted for inhomogeneities or biases, since increases in the number of observations typically outweigh concerns about these errors. However, platform types or data sources that are found to introduce significant biases are excluded from monthly summaries. For instance, wind data from drifting buoys and fishing fleet observations were excluded due to their systematic bias towards low wind speeds (Woodruff *et al.*, 1993). There are remaining differences in wind speed between *standard* and *enhanced* COADS Release 1a data (Figure 3, Woodruff *et al.*, 1993), which for regions with moored buoys appear to indicate a localized low bias in the *enhanced* COADS version. In addition, known errors associated with a variety of data sources are corrected on an ongoing basis, as documented in Woodruff *et al.* (this publication) and on pertinent COADS web pages (<http://www.cdc.noaa.gov/coads/e-doc/other/>).

Given these general principles, the QC procedures applied to the original COADS Release 1 were as follows:

- All marine reports were subjected to an NCDC procedure that checked for observations outside climatological means plus/minus 5.8 standard deviations, developed for 5° boxes on older NCDC files, as well as checks for consistency and code legality. This QC procedure assigned a flag to each report that was used in the selection among duplicate reports. Owing to overlapping data sources, around one quarter of the original input reports were rejected during duplicate elimination processing (Woodruff *et al.*, 1987).
- Observations exceeding the 5° × 5° QC limits were not transferred into the Compressed Marine Reports (CMR) that formed the input records for Release 1 statistics. For missing 5° × 5° limits, observations were transferred to CMR without this QC. In addition, observations exceeding global physical limits for each variable (e.g. air temperatures outside the range of -88°C and +58°C) were excluded from CMR.
- Trimming - the final QC procedure in COADS - was designed to reject statistical outliers and questionable weather observations from 2° monthly summaries. After some smoothing across time and space, upper (lower) trimming limits for each calendar month were defined as the long-term median g plus (minus) 3.5 upper (lower) standard deviations σ_u (σ_l), which derive from the difference between the 5th (1st) sextile and the median of all observations within the trimming period (1854-1909, 1910-49 or 1950-79). Thus, systematically skewed observational distributions were taken into account in the trimming process. Table 1 documents typical values for σ_u and σ_l in sea level pressure (SLP), SST, and wind speed (W), confirming the well-known positive skewness of wind speed measurements, and the mid-latitude negative skewness of SLP observations. All observations outside the trimming limits were excluded from the computation of trimmed monthly 2° × 2° summaries in COADS. Derived COADS variables were trimmed indirectly by using only trimmed observed variables as input (Slutz *et al.*, 1985, Figure A4-1).

Interim COADS summaries (1980-91) and the *standard* version of Releases 1a (1b) for 1980-97 (1950-79) have been trimmed based on the 1950-79 trimming limits. These limits were expanded to include all observations up to $4.5\sigma_{u/l}$ away from the median for the *enhanced* version of Releases 1a and 1b (Woodruff *et al.*, 1993, 1998). In the *standard* version of COADS Release 1c (Woodruff *et al.*, this publication), the original Release 1 trimming limits for 1854-1909 were used for all ship-based data before 1910, and 1910-49 limits for that same period. Analogous to Release 1a and 1b, the *enhanced* version of Release 1c employs expanded $4.5\sigma_{u/l}$ trimming limits (see also: <http://www.cdc.noaa.gov/coads/r1c.html>).

3. EVIDENCE OF TRIMMING PROBLEMS IN COADS

If the climate of a given area were stationary, standard COADS trimming limits of $\pm 3.5\sigma$ around the median would typically remove 1 in 2500 observations (i.e. for normal distributions; Slutz *et al.*, 1985). In reality, two independent factors increase the trimming removal rates by inflating the tails of the observational distributions: first, a variety of error-generating mechanisms introduce mostly *random* outliers, and second, climate variability translates into *systematic* shifts of the observational distribution, since this year's climate is different from that of last year.

Table 1—Upper and lower standard deviations ($\sigma_{u/l}$) for SLP, SST and wind speed (W) in January and July of 1950-79, zonally averaged for 10° latitude belts of interest (using Release 1b data). Mid-latitude data imply negative skewness for SLP and widespread positive skewness for W. Units are in 0.01mb for SLP, 0.01°C for SST, and 0.01m/s for W. Underlined values indicate a 10 per cent surplus in σ_u vs. σ_l , or σ_l vs. σ_u .

Considering the first factor, erroneous outliers may be created by missing or altered digits, wrong conversion into SI units, or simply by replacing the true value with zero. The second factor is related to the ratio of interannual variability

Lat.	SLP (Jan)		(Jul)		SST (Jan)			(Jul)		W (Jan)		(Jul)	
	σ_u	σ_l	σ_u	σ_l	σ_u	σ_l/σ_u	σ_l	σ_u	σ_l	σ_u	σ_l		
50-30S	194	<u>212</u>	267	<u>296</u>	42	42	34	34	<u>128</u>	112	<u>130</u>	114	
30-10S	118	124	135	141	45	45	45	44	<u>131</u>	118	<u>149</u>	134	
10N/S	101	101	91	92	45	43	48	46	121	111	125	114	
10-30N	234	251	155	166	85	84	76	74	<u>255</u>	227	<u>208</u>	189	
30-49N	732	<u>834</u>	377	<u>429</u>	127	124	144	146	<u>440</u>	362	<u>290</u>	244	
50-69N	752	799	430	470	73	71	100	100	<u>341</u>	284	<u>257</u>	216	

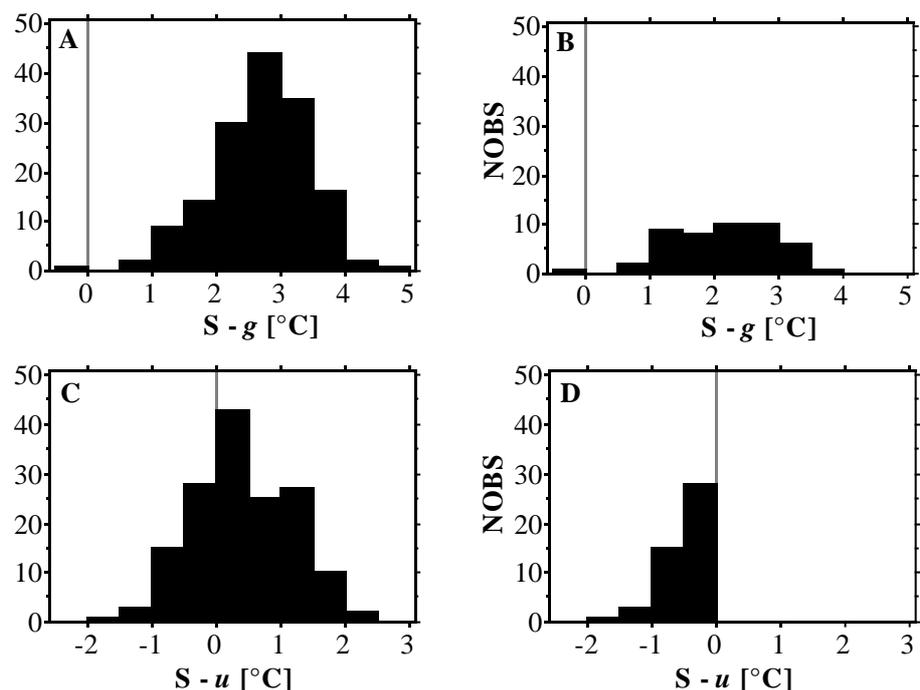
versus sub-monthly variability. If short-term variability within a calendar month is large, trimming limits tend to be wide and accommodate large interannual variability as well. If short-term variability is small, it may result in narrow trimming limits, which can screen out valid observations associated with large interannual climate anomalies (statistical Type I errors; Wolter, 1992). Note that systematic skewness-related departures from the normal distribution are addressed via upper and lower standard deviations (Table 1), and are not considered to be an important trimming problem.

The most prominent examples of Type I trimming errors were found for tropical Pacific SST which, in association with persistent ENSO conditions, can lead to significant distortions of the climate record for those regions (e.g. Wolter, 1997). Here, we reproduce a figure from Wolter *et al.* (1989) for January 1878, at the zenith of the biggest known El Niño of the 19th century (Kiladis and Diaz, 1986). Within the general area chosen here (10°N - 10°S , 80° - 180°W), two thirds of all observations were rejected, although the most anomalous observation was less than 5°C above its applicable long-term median g (Figure 1(a)). Note how 'normal' the distributions of SST anomalies appear in Figures 1(a) and (c), which include all observations, and how 'truncated' they appear if only observations within the trimming limits are considered (Figures 1(b) and (d)). The highest reported untrimmed SST within the domain in January 1878 is just above 30°C for a 2° box and 31°C for an individual observation, certainly within plausible limits of tropical SST. In sum, excessive trimming reduces the SST anomaly by up to 1°C for the ENSO event of 1878. Similar errors were found for the peak of the 1982-83 El Niño (Wolter, 1997), which, until recently, was considered as the strongest El Niño of the 20th century, justifying a major revision of COADS trimming procedures.

Excessive trimming can also be found for other variables, and outside the tropics, for instance near polar ice-edges where marine observations are rare to begin with. However, tropical Pacific SST trimming losses have received the most attention due to their systematic, direct impact in the assessment of ENSO events and associated global climate anomalies, while other trimming losses appear less systematic and widespread, and with unknown ramifications for global climate.

Wind variables could probably be trimmed better than in the current procedure which analyses zonal and meridional wind components separately, removing all wind information if either component fails the test. This procedure can be argued to censor winds from the four cardinal wind directions more frequently than winds from other points of the compass. Of course, there are

Figure 1—Histograms of general ENSO area [10°N - 10°S , 80° - 180°W] SST anomalies for January 1878. Frequencies are given for individual SST observations (S) minus the applicable local long-term median (g) for January 1854 - 1909, computed for each 2° box separately: (a) untrimmed, (b) trimmed. Analogous frequencies are shown for S minus the applicable local upper trimming limit (u) for the same period: (c) untrimmed, (d) trimmed. Data are from COADS Release 1. Bins are partitioned into $1/2^{\circ}\text{C}$ intervals, with observations being greater than the lower boundary, and smaller than or equal to the upper boundary (after Wolter *et al.*, 1989).



many other problems associated with wind estimation and measurements (Woodruff *et al.*, 1998) that will need to be addressed in future COADS updates (see also section 5).

Prior to the creation of COADS Release 1a, tests were run on 1970-89 data to delineate excessive trimming removals for all primary COADS variables. In turn, this information was used to determine the scope of trimming procedure changes appropriate for Release 1a. Differentiated by variable and decade (Table 2), the fraction of global $3.5\sigma_{u/l}$ trimming removals in the number of observations is highest in SST (generally above 2 per cent), followed by air temperature and wind (well above 1 per cent), and lowest for SLP and relative humidity (about 1 per cent). Lower trimming rates for non-temperature variables are probably due to the inherently larger variability and noisiness of these fields, resulting in wider trimming limits.

Trimming removal rates in 1980-92 are more than cut in half by widening the trimming limits from $3.5\sigma_{u/l}$ to $4.5\sigma_{u/l}$ (using the identical set of *enhanced* input data; Table 2). Based on the examination of histograms of individual observations in selected areas, global trimming removal rates of up to 1 per cent appear justifiable, presumably associated with digitization and communication errors. Any remaining trimming losses are inferred to be statistical Type I errors.

4. ALTERNATIVE TRIMMING IN COADS RELEASES 1A-C

Given the extent of the trimming problems discussed in the previous section, it was decided to modify the COADS trimming procedure for Release 1a, while preserving continuity with Release 1. In statistical terms, the balance of Type I to Type II errors was shifted from mainly minimizing Type II errors to minimizing errors of both types.

After some experimentation with different trimming approaches, we decided to inflate the existing limits from $3.5\sigma_{u/l}$ to $4.5\sigma_{u/l}$. This was the method implemented in the *enhanced* version of COADS Release 1a, because $4.5\sigma_{u/l}$ trimming limits appeared to be far more accommodating to extreme climate states than $3.5\sigma_{u/l}$ limits, while still removing the vast majority of true statistical outliers. This was verified with histograms, and with correlation maps of trimming losses compared to regional anomalies. A more drastic inflation of trimming limits to $5\sigma_{u/l}$ was rejected, because the increased risk of admitting statistical outliers was not outweighed by better accommodation of climate extremes.

Given the interest in a data set for 1980-92 that was compatible with COADS Release 1 (1854-1979), which consists mostly of ship data, a *standard* set of the monthly trimmed 2° box summaries was produced for COADS Release 1a that keeps the original trimming limits established for 1950-79 and admits only ship data (as far as it can be ascertained). In contrast, 1950-79 trimming limits for the *enhanced* set have been expanded to include all observations up to $4.5\sigma_{u/l}$ away from *g*. It includes most ocean surface-based observational platforms. Further processing details are given in Woodruff *et al.* (1993), while Wolter (1997) documents the improved spatial patterns of global trimming losses in SST (and SLP) fields.

Subsequently, *standard* and *enhanced* sets with $3.5\sigma_{u/l}$ and $4.5\sigma_{u/l}$ trimming limits, respectively, were created in the same fashion for 1950-79 data for COADS Release 1b data (Woodruff *et al.*, 1998), even though this period was characterized by less extreme ENSO events and a large majority of conventional ship-board observations (Woodruff *et al.*, this publication). However, consistently trimmed

Table 2—Global trimming removal rates (in percentage of total number of observations for each variable) for air temperature (AT), SLP, relative humidity (RH), SST, and zonal and meridional (U&V) wind. These are listed for COADS Release 1 data from the 1970s (Table 3-3 in Slutz *et al.*, 1985), for interim data from 1980-91, and for Release 1a data from 1980-92. The enhanced $3.5\sigma_{u/l}$ portion of this table is based on an analysis of all individual observations (i.e. an enhanced platform mix flagged in the standard ($3.5\sigma_{u/l}$) manner). For details (see Wolter, 1997).

	AT	SLP	RH	SST	U&V
1970-79, $3.5\sigma_{u/l}$	1.5	0.9	0.4	2.3	1.4
1980-91, interim $3.5\sigma_{u/l}$	2.5	1.6	0.6	3.4	1.9
1980-92, standard $3.5\sigma_{u/l}$	1.9	1.3	0.4	2.9	2.2
1980-92, enhanced $3.5\sigma_{u/l}$	1.9	1.5	0.4	2.6	2.1
1980-92, enhanced $4.5\sigma_{u/l}$	0.8	0.6	0.1	1.1	0.8
1980-92, enhanced ($3.5\sigma_{u/l}$ - $4.5\sigma_{u/l}$)	1.1	0.9	0.3	1.5	1.3

versions of COADS from 1950 through to 1997 are useful in the assessment of interannual and decadal climate variability (e.g. IPCC, 1996).

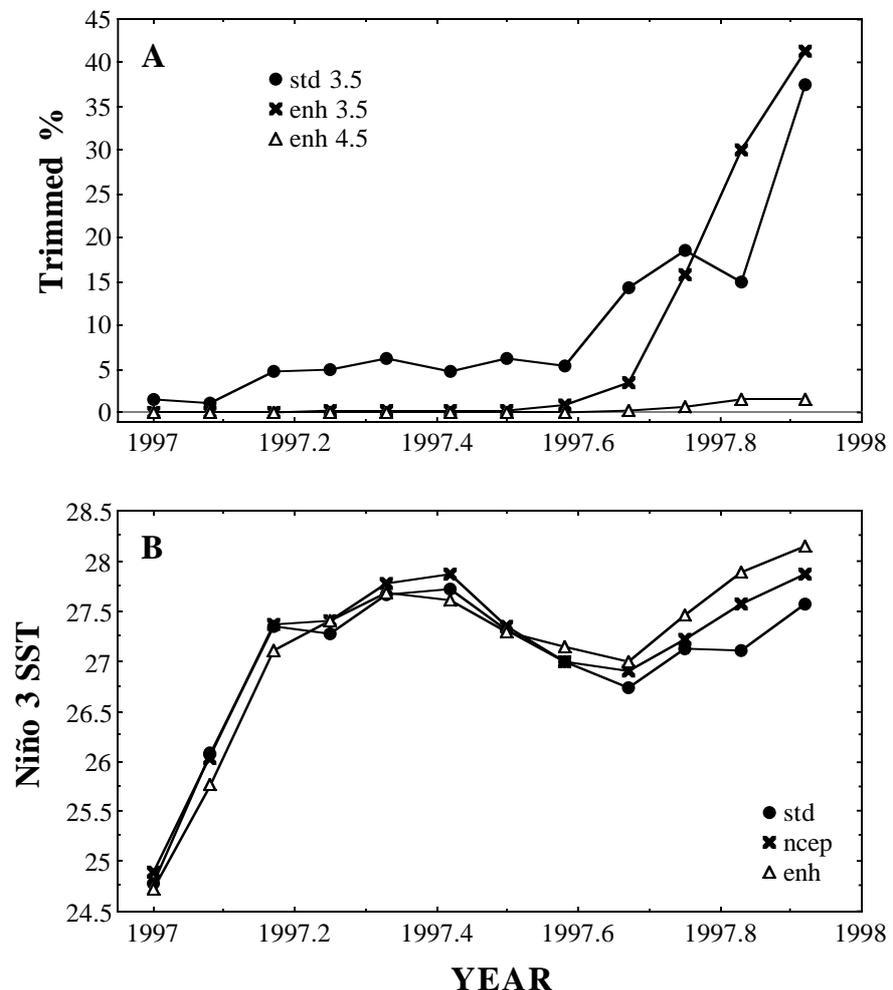
Most recently, *standard* and *enhanced* sets were created for COADS Release 1c (1784-1949) in a similar manner (Woodruff *et al.*, this publication), but utilizing original trimming information ($\sigma_{u/l}$) from 1854-1909 for the early part of the record (1784-1909), and 1910-49 trimming limits for the latter four decades. These trimming periods were originally designed to accommodate the historic shift from sailing ships and early steamers to 20th century observing platforms (Slutz *et al.*, 1985). There is a remaining concern that even *enhanced* trimming limits at $4.5\sigma_{u/l}$ may still lead to excessive Type I errors under extreme ENSO conditions (Figure 1; also Wolter, 1992). However, the addition of substantial new data (Woodruff *et al.*, this publication) without the resources to recompute trimming limits necessitated this compromise solution.

5. PENDING COADS QC ISSUES

A variety of near-real-time observational and statistical products are produced by CDC within the COADS framework, using marine data supplied by the National Center for Environmental Prediction (NCEP). The most recent El Niño 'event of the century' (1997-98) serves to remind us that extreme climate events and their associated trimming problems are not just a phenomenon of the past (Figure 2). In fact, both in terms of the measured SST anomalies in late 1997, and in terms of $3.5\sigma_{u/l}$ trimming losses, the most recent El Niño event even surpassed 1982-83 (compare current Figures 2(a) and (b) against Figures 2 and 5 in Wolter, 1992, respectively).

Near-real-time monitoring should be undertaken in a consistent manner with recent COADS Releases. In that context, current trimming limits at the 'inherited' $3.5\sigma_{u/l}$ level should be replaced with $4.5\sigma_{u/l}$ limits. Given the recent shift towards non-traditional marine observing platforms (Woodruff *et al.*, this

Figure 2—Impact of the 1997 El Niño event on the Niño 3 region [2°N - 10°S , 90° - 150°W] SST. The first panel (a) shows the monthly trimming losses during 1997 for the standard and enhanced COADS Release 1a data (labelled *std 3.5* and *enh 4.5*), as well as the hypothetical trimming losses for the enhanced platform mixture trimmed at $3.5\sigma_{u/l}$ (*enh3.5*). The latter most closely mimics the near-real-time trimming setting whose trimming rates are not available. The second panel (b) shows the resulting monthly SST for the same region and year, in the standard and enhanced COADS sets, as well as in the near-real-time (NCEP) product.



publication), a single *enhanced* near-real-time product that encompasses all platforms at the $4.5\sigma_{u/l}$ trimming level should be sufficient for most purposes.

However, the TOGA-TAO moored buoy array transmits over the GTS during local daytime only, for logistic and cost-saving reasons (cf. <http://www.pmel.noaa.gov/toga-tao/gts.html>). It contributes a large fraction of real-time SST observations in the tropical Pacific. This raises the possibility of a warm bias due to daytime heating of the sea surface. In the Release 1a (1980-97) update we were able to replace GTS receipts of TOGA-TAO data with delayed-mode data obtained from PMEL. During the latter half of 1997, the potential positive bias due to real-time TOGA-TAO data may have partially cancelled out the negative bias due to the conventional trimming limits ($3.5\sigma_{u/l}$) employed in COADS near-real-time processing (Figure 2(b)). In order to best adjust for the potential daytime bias by the TOGA-TAO array, historic differences should be assessed between daytime only and full 24-hour records for all TOGA-TAO moored buoys.

Trimming should not be considered in isolation. It is an important part of comprehensive QC procedures applied to COADS. Future COADS updates (such as Release 2) may include separate products for different times-of-day and types of observational platforms. These goals need to be balanced against the penalties of reduced sample sizes. Especially in the early instrumental record, observations were so few and far between that any attempt at more sophisticated QC measures is severely handicapped.

Nevertheless, collaborative efforts are under way (Woodruff *et al.*, this publication) to further improve COADS QC. As originally discussed in Wolter (1992), trimming should apply to the scatter of observations about the *individual* monthly median rather than about the *long-term* median \bar{g} . This type of trimming has been hampered in the past in regions of low observational density, but optimum interpolation techniques can be brought to bear at least on SST fields to capture the large-scale monthly mean (or median) fields correctly, as described in Reynolds and Smith (1994). Similar improvements in the trimming of wind observations are considered to be important as well. The proper balance of Type I and Type II error will remain an active research topic in COADS for many years to come.

ACKNOWLEDGEMENTS

The numerous discussions with Henry Diaz, Roy Jenne, Dick Reynolds and Steve Worley on COADS QC are gratefully acknowledged. We also thank the anonymous reviewers for their insightful comments that hopefully led to an improved manuscript. This research has been supported through EPOCS and the NOAA Climate and Global Change program.

REFERENCES

- IPCC (Intergovernmental Panel on Climate Change), 1996: *Climate Change 1995. The Science of Climate Change*. Cambridge University Press, Cambridge, xii + 572 pp.
- Kiladis, G.N. and H.F. Diaz, 1986: An analysis of the 1877-78 ENSO episode and comparison with 1982-83. *Mon. Wea. Rev.*, 114, 1035-1047.
- Reynolds, R.W. and T.M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, 7, 929-948.
- Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: Comprehensive Ocean-Atmosphere Data Set (COADS). Release 1. NOAA Environmental Research Laboratories, Boulder, CO, 268 pp.
- Wolter, K., 1992: Sifting out erroneous observations in COADS—the trimming problem. *Proceedings of the International COADS Workshop* (Boulder, Colorado, 13-15 January 1992). H.F. Diaz, K. Wolter and S.D. Woodruff, Eds., NOAA Environmental Research Laboratories, Boulder, CO, 91-101.
- Wolter, K., 1997: Trimming problems and remedies in COADS. *J. Climate*, 10, 1980-1997.
- Wolter, K., S.J. Lubker and S.D. Woodruff, 1989: Trimming — a potential error source in COADS. *Tropical Ocean-Atmosphere Newsletter*, 51 (Summer 1989), 4-7.
- Woodruff, S.D., S.J. Worley, J.A. Arnott, H.F. Diaz, J.D. Elms, M. Jackson, S.J. Lubker and D.E. Parker, 2001: COADS updates and the blend with the UK Met Office

- Marine Data Bank. *WMO Guide to the Applications of Marine Climatology*, WMO/TD-No. 1081, 2003, JCOMM Technical Report No. 13.
- Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, 68, 1239-1250.
- Woodruff, S.D., S.J. Lubker, K. Wolter, S.J. Worley and J.D. Elms, 1993: Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. *Earth System Monitor*, 4, No. 1, 6 pp. [on-line at: <http://www.cdc.noaa.gov/coads/coads1a.html>].
- Woodruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. earth*, 23, 517-527.

SECTION 4

DEVELOPMENT AND USE OF SATELLITE MARINE DATABASES

An intercomparison of voluntary observing, satellite data, and modelling wave climatologies	127
The joint calibration of altimeter and in situ wave heights	139
On the use of in situ and satellite wave measurements for evaluation of wave hindcasts	149
Scatterometry data sets: high quality winds over water	159

AN INTERCOMPARISON OF VOLUNTARY OBSERVING, SATELLITE DATA, AND MODELLING WAVE CLIMATOLOGIES

P. David Cotton¹; Peter G. Challenor, Lisa Redbourn-Marsh²; Sergei K. Gulev, P. Shirshov³; Andreas Sterl⁴; Roman S. Bortkovskii⁵

1. INTRODUCTION

This paper presents early results from an INTAS sponsored programme, the aim of which is to evaluate global scale wave climatologies compiled from visual observations, remotely-sensed data and global wave model output. Three large scale wave climatologies are compared, namely a climatology derived from the most recent release of COADS data (1979-96), a satellite altimeter-derived climatology based on measurements from Geosat, ERS-1 and TOPEX/Poseidon (1985-97), and finally a climatology based on output from the ECMWF Reanalysis project (ERA), whose homogeneous wind fields were used to drive a global scale third generation WAM model (1979-94). The INTAS programme is investigating 'static' differences in the climatologies (e.g. through point-by-point comparisons of co-located grid cells), and 'dynamic differences' by comparing how the three climatologies represent interannual climate variability. Separate climatologies of wind sea (i.e. waves generated by the local wind field) and swell (waves not generated locally) are available in the COADS and ERA analyses, but not from the altimeter data.

2. THE 'INTAS' PROJECT 2.1 DESCRIPTION

This study is supported by the INTAS foundation (project 96/2089), which exists to develop the scientific potential of the Newly Independent States (NIS) of the former Soviet Union by encouraging scientific cooperation between the INTAS partners (the NIS, a number of western European countries, and the European Community).

This programme — The intercomparison of the world ocean wind and wave climatology from in situ, voluntary observing, satellite data and modelling — started in February 1998 and continues until February 2001.

The main goal of the scientific programme is to evaluate three global scale ocean surface wind and wave climatologies through comprehensive intercomparisons and to eventually publish a climatology atlas. To achieve this goal a number of intermediary objectives have been set:

- An update of global scale sea-state parameters for the period 1979-1996 from the historical collection of merchant ships' observations and evaluation of the basic characteristics of the wind and waves.
- Analysis of accuracy and reliability of the 13-year (1985-1997) remotely-sensed global wave and wind data from different research satellites.
- Analysis of accuracy and reliability of the homogeneous surface wave hindcast from the WAM model, driving by the ERA (ECMWF Reanalysis) project winds.
- Cross calibration of voluntary observing fields, satellite wave and wind data and model wave hindcast using high quality instrumental measurements.
- Evaluation of the reliability of long-term changes in winds and waves during the 1980s and 1990s and the study of the relationships between interannual variability of the wave and wind characteristics and atmospheric circulation patterns, such as the North Atlantic Oscillation.

1 Satellite Observing Systems, 15, Church St, Godalming, Surrey, GU7 1EL, UK

2 Southampton Oceanography Centre, Southampton, UK

3 Institute of Oceanology, Moscow, Russia

4 Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

5 Main Geophysical Observatory, St. Petersburg, Russia

3. DATA SETS The main characteristics of the three data sets compared in this paper are summarized in Table 1, and discussed in more detail below:

Table 1

	<i>period</i>	<i>grid size</i>	<i>parameters</i>
COADS	1964–96	5° × 5° (na)	<i>U, hs, hw,</i>
ALT	1985–97	2° × 2° (g)	U, SWH
ERA/WAM	1979–94	1.5° × 1.5° (g)	<i>U,</i> wave spectra (12 dirs × 25 freqs)

The three data sets compared in this study are the following: COADS (visual observations), ALT (satellite altimeter data), and ERA/WAM (global wave model output). The COADS gridded data cover the North Atlantic (NA) only, ALT and ERA/WAM are global (g). Parameters in bold and italicised are available as vectors. U - ocean 10 m wind, SWH - significant wave height, hs - swell height, hw - wind sea height.

3.1 COADS DATA

The visual data used here were extracted from the Comprehensive Ocean-Atmosphere Data Set (COADS) for the 1964–1996 period. Data from Compressed Marine Reports (CMR-5) were used for the 1964–79 period, and from Long Marine Reports (LMR) for the 1980–96 period (Woodruff *et al.*, 1998). Maps of data density were produced and evaluated, so that ocean regions which provide high and low sampling could be identified. Data from areas of high sampling (North Atlantic, north-west Pacific, tropical Pacific) have been used for the cross-calibration of visual observations against instrumental measurements (Gulev, Proceedings of CLIMAR99).

These data have also been used to assess the algorithms used to combine the separate visual estimates of wind sea and swell into a single significant wave height (SWH) estimate (Gulev *et al.*, 1998). It was established that H30 (generated by taking the square root of the sum of squares of the wind sea and swell significant heights when their directions lay within 30° of each other, or otherwise by taking the higher of sea or swell height) provided the best fit to the instrumental data in regions where the wind sea and swell displayed directional steadiness. So this algorithm was selected for the comparisons presented in this paper.

To generate the gridded data (in this paper we consider only the North Atlantic) individual COADS reports were quality controlled (Gulev and Hasse, 1998) and then selected variables and derived products were extracted and averaged to provide monthly means on a 5° × 5° grid, for each month from 1964-1993.

Features

The major advantage of COADS data is their long term coverage (1964 onwards). Also, the availability of separate estimates of wind sea and swell with directions is a useful feature. To balance this however, the COADS data have the disadvantage of being based upon a subjective estimate and so their reliability may suffer. The COADS data are assessed against in situ data elsewhere in this publication (Gulev, Proceedings of CLIMAR99). There is also a possibility that the sampling of conditions may be self selective (few commercial vessels will choose to endure severe conditions unnecessarily) which could bias statistics. Whilst shipping lanes are well sampled, limited spatial coverage may also cause problems, if global fields are required.

3.2 SATELLITE ALTIMETER DATA (ALT)

The altimeter 2° × 2° monthly mean SWH climatology was generated at Southampton Oceanography Centre (SOC) from three Ku-band altimeters: on GEOSAT (1985-89), TOPEX/POSEIDON (1992-97) and ERS-1 (1991-96). There are some gaps in this altimeter data set, in 1986 and 1990-1991. Cotton and Carter (1994) compared altimeter monthly mean values on a 2° × 2° grid with data from 24 NDBC buoys. The linear regressions thus obtained were then applied to the data from each satellite to produce consistent and corrected SWH values. The 1 Hz altimeter geophysical data records provided by NOAA (Geosat), AVISO (TOPEX/Poseidon) and CERSAT (ERS-1) were quality controlled, calibrated and then averaged onto a 2° × 2° monthly mean grid. The gridding procedure at SOC took the average of the medians of each pass through a grid square in each month

(Cotton and Carter, 1994). For the comparison presented in Figure 1, the data were averaged onto a $5^\circ \times 5^\circ$ monthly mean grid.

Challenor and Cotton (this publication) have provided more recent calibrations which represent an improvement on previous work. Whilst these new calibrations have not been applied to the data set studied in this paper, it is not believed that they would materially effect the conclusions.

Features The companion paper by Challenor and Cotton contained in this publication, confirms the view of previous work (e.g. Gower, 1996; Cotton *et al.*, 1997), that individual altimeter measurements of significant wave height are highly accurate. Comparisons with co-located in situ buoy data show residual root mean square values of less than 0.5 m (0.3 m for TOPEX and Poseidon). Thus, the altimeter data can be regarded as being at least as accurate as the buoy measurements. However, whilst work is proceeding on the development and testing of a wave period algorithm (Davies *et al.*, 1998), the altimeter is not currently able to provide any directional or spectral information. A further problem is that global altimeter data are only available since 1985. The future is encouraging, though, with the launch of Geosat follow-on in 1998 (it became operational in November 2000) and the planned launches of two further altimeter satellites in 2001 (JASON and ENVISAT).

3.3 GLOBAL WAM/ERA MODEL OUTPUT (ERA)

The model output employed in this study were generated by the third generation WAM at the Royal Netherlands Meteorological Office (KNMI) which was forced by wind fields produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) under its ERA Reanalysis project. The ERA project ran a consistent version of the ECMWF atmospheric model for a 15-year period from

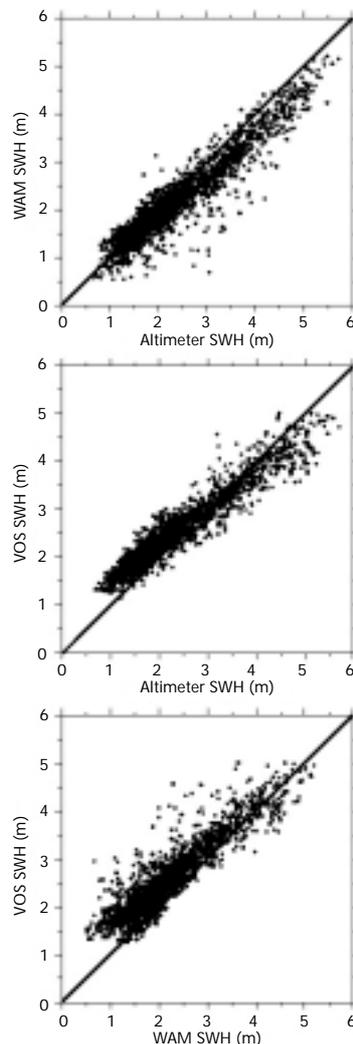


Figure 1—Scatter plots of monthly means on a $5^\circ \times 5^\circ$ North Atlantic grid, from the ALT, ERA/WAM and VOS data sets (see text), (top) WAM vs. ALT, (middle) VOS vs. ALT, (bottom) VOS vs. WAM. Copyright Physics and Chemistry of the Earth (Pergamon).

January 1979 to February 1994. The WAM model was run as part of an evaluation study for the ERA project (Sterl *et al.*, 1998). For the ERA study, the WAM model was run in both low resolution (LR, $3^\circ \times 3^\circ$ grid) and high resolution (HR, $1.5^\circ \times 1.5^\circ$ grid) versions. We only consider results from the HR version here, because these data showed less scatter and smaller biases than those from the LR version when compared with instrumental data. The HR version covers the globe from 81°S to 81°N and computes wave spectra in 12 directions and at 25 frequencies. Results are output every six hours, giving, among other quantities, heights and periods of sea, swell, and SWH. Again, for the comparison presented in Figure 2, the data were further averaged onto a $5^\circ \times 5^\circ$ monthly mean grid.

Note that ECMWF has recently commenced the ERA40 project, in which it will run a consistent version of its atmospheric model, coupled with the WAM wave model, over a period of 40 years.

4. 'STATIC' COMPARISON OF ALT, WAM AND VOS CLIMATOLOGIES

In this section we present results from two studies. In the first study (Sterl *et al.*, 1998), the ERA/WAM data were assessed through comparisons against the ALT data and in situ measurements. The ERA/WAM data were then studied for evidence of significant change in wave climate over the 15-year ERA period (1979-94). We subsequently refer to this paper as S98. In the second study (Gulev *et al.*, 1998), the VOS, ERA/WAM and ALT data were intercompared, and preliminary studies were carried out into representation of climate variability in the ERA/WAM and VOS data. This paper is subsequently referred to as G98.

4.1 WAM Vs. ALT

S98 compared ERA/WAM gridded wave fields with ALT data and in situ data. The altimeter data were co-located within 30 minutes and 50 km of the buoy measurements. The model parameters were extracted at the buoy location. In the comparison it was established that, even with the high resolution data (HR - $1.5^\circ \times 1.5^\circ$ grid), the ERA/WAM estimates of significant wave height were consistently lower than the ALT data for higher waves, but were higher for low waves. This tendency was seen in comparisons of climatological charts, and in time series of averaged data. Following further comparisons with in situ buoy data, S98 concluded that the WAM output appeared to be in error since the WAM significant wave height displayed the same tendencies when compared to buoy measurements (Figure 2). The consequence of these WAM underestimates of high waves and overestimates of low waves is that, although the averaged model and buoy values compare well, and the WAM estimates display little overall bias, the full extent of true short-term variability in wave height may not be recreated in the model. This in turn may have consequences in how well climate variability is represented in the model output.

S98 considers a number of possible sources for the mismatch between model output and altimeter and buoy data. It concludes the two most likely sources lie within WAM; namely model resolution and model error. A comparison between the results from the low and high resolution versions of WAM showed that increasing the resolution had a beneficial effect in situations of high SWH and highly variable SWH, while at low wave heights the model results actually became worse (i.e. the WAM SWH overestimated to a greater extent). This suggests that the higher resolution model runs improve the representation of variability, but then reveal an underlying tendency in WAM to overestimate the magnitude of the background wave height field. Maps of the relative strengths of swell and significant wave height revealed that the areas of overestimation of SWH coincided with areas of high swell to SWH ratio. This raised the possibility that WAM contains too high swell, and there has been some discussion as to whether the swell propagation within WAM could be improved. S98 concluded, therefore, that the WAM underestimation of high waves was due to limited resolution in the wind fields (meaning that the highest wind peaks are missed), whereas the overestimation of low waves may be due to internal WAM errors, possibly in the swell propagation terms.

4.2 VOS Vs. WAM Vs. ALT

G98 intercompared $5^\circ \times 5^\circ$ climatologies for the North Atlantic produced from the VOS, ALT and ERA/WAM North Atlantic data sets, and demonstrated that all three products have their strengths and weaknesses.

Figure 2—Time series of WAM HR (solid line), WAM LR (dashed line) and Buoy (NDBC buoy 46006 -dotted line) significant wave heights, with co-located, calibrated, Geosat altimeter data (asterisks) for January-March 1987 (a) and July-August 1987 (b). Copyright J. Geophys. Res. (AGU).

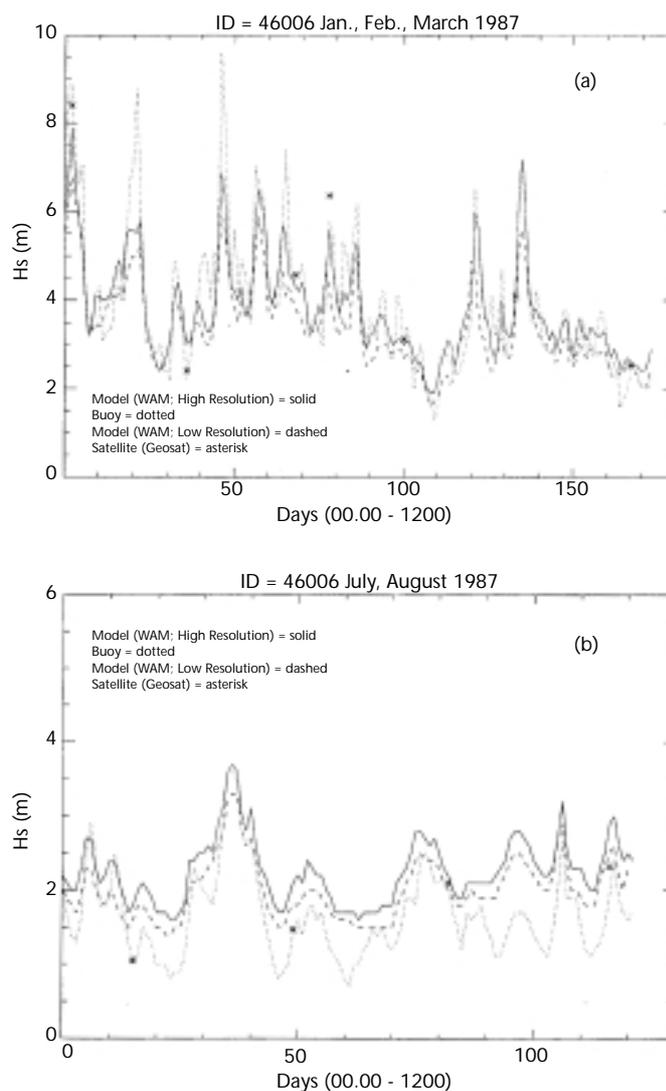


Figure 1 shows scatter plots of the monthly mean SWH on a $5^\circ \times 5^\circ$ grid over the North Atlantic from the ALT, WAM and VOS data. These data cover the 80 months during which the three data sets overlap. Note that the large spatial scale and monthly averaging in this comparison may mask some of the WAM tendencies to underestimate variability. Bearing this in mind, we see that the WAM data agree fairly well with the ALT data, but display a tendency to underestimate higher values (by about 0.5 m for Mean SWH > 2.5 m). The orthogonal regression slope of WAM against ALT SWH is 0.86.

When compared against ALT and WAM monthly means, the VOS data are seen to overestimate low waves, and also to underestimate high waves with respect to ALT. The orthogonal regression slopes of VOS against WAM and ALT were 0.89 and 0.77 respectively. The overall biases in VOS data were 0.32 m and 0.14 m, respectively, against WAM and ALT.

The major climatological spatial patterns and the seasonal cycle in all three products are at first glance comparable, and appear to depict the North Atlantic wave climatology quite realistically. In fact, previous comparisons of different VOS-based atlases have shown even higher biases with respect to each other than the biases which have been found between the three independent climatologies considered in this paper. At the same time, the differences between the VOS wave data, altimeter measurements and the model hindcast are not negligible, and the nature of biases must be carefully studied. Figure 3(a) presents a mean VOS SWH climatology for the 80 months of contemporaneous data. The general pattern of this figure is repeated in the ALT and WAM data sets, with the mid-latitude maxima occurring in the same location, as are the subtropical and equatorial

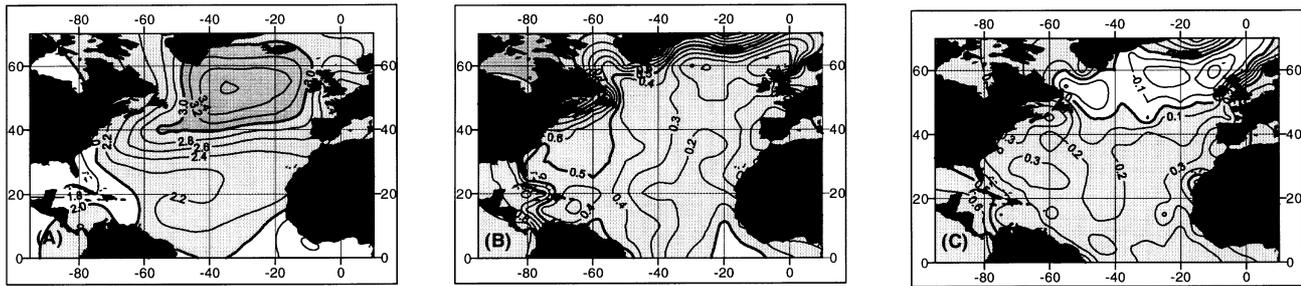


Figure 3—(a) Mean VOS significant wave height climatology (1985-94), and difference plots against WAM (b) and ALT (c). Contours in m. Copyright Physics and Chemistry of the Earth (Pergamon).

minima. Figures 3(b) and 3(c) illustrate the differences between this mean climatology and those from WAM and ALT. Note that the largest differences between the climatologies are seen at high latitudes, which may be a consequence of ice cover being dealt with in different ways by the three data sets. Thus, the reader is advised to focus on areas which remain ice-free throughout the year. From Figures 3(b) and 3(c) we see that VOS SWH is systematically higher than the WAM SWH over the whole North Atlantic, by 0.2-0.6 m. The largest differences between the VOS and WAM fields are found in the western subtropics, in regions close to the North American coast, and at high latitudes (possibly due to ice cover). The best agreement occurs in the north-east Atlantic, where differences are less than 0.2 m.

Differences between VOS and ALT are generally of lower magnitude, and VOS SWH are again higher than ALT SWH, except at mid- to high-latitudes (50°-70°). Note that in this region WAM gives 0.3-0.5 m lower waves than ALT.

Of the three sets of comparisons, the VOS and altimeter SWH show least scatter, whilst the largest scatter is obtained for the VOS-WAM comparison.

4.3 VOS Vs. WAM: WIND SEA AND SWELL

G98 also compared the separate wind sea and swell fields from VOS and WAM. Figure 4 shows the mean VOS climatology, and the VOS minus WAM difference fields. Again the spatial patterns are similar in nature, but some important differences can be identified.

Considering first the wind sea (Figures 4(a) and 4(c)), it is apparent that the VOS climatology is systematically higher than WAM apart from a small region centred on 60°N 30°W. The difference becomes largest at subtropical and equatorial latitudes. This may be a consequence of the use of the lowest COADS code '1', which in theory corresponds to a height of 0.5 m, though in practice is also used to represent all heights below this. Thus, very low wave height values may be over-represented in COADS, perhaps by tens of centimetres. In fact, a study of seasonal values demonstrated that the VOS overestimate was greatest during the summer months in the tropics and subtropics, and was 0.6-0.7 m. In the winter at mid-latitudes, the WAM wind sea was greater than the VOS wind sea.

As regards the swell fields (Figures 4(b) and 4(d)), one can see that the VOS swell climatology is higher than the WAM swell over the entire North Atlantic, apart from a small region off West Africa (0°-10°N, 0°-40°W). If the WAM climatology does indeed contain an overestimate of swell, then this would indicate an even greater error in the VOS swell climatology. This overestimation gets progressively higher towards the North West. Clearly, WAM generates much less swell than the VOS data show in the Labrador Sea.

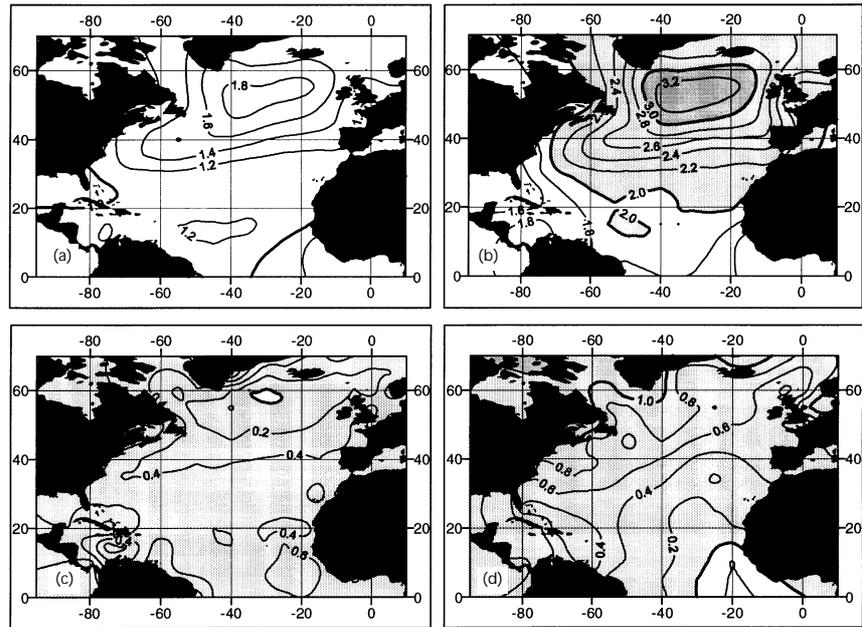
5. 'DYNAMIC' COMPARISON OF ALT, WAM AND VOS CLIMATOLOGIES

5.1 INTRODUCTION

A significant aim of this scientific study was to develop a better understanding of the nature of wave climate variability, by taking advantage of the individual merits of the three separate data sets. However, before this could be achieved it was clearly important to establish the major characteristics and differences in climate variability as represented in the different WAM, ALT and VOS climatologies.

Given the widely different characteristics of the three data sets, the definition of a suitable methodology for comparing the climate variability within them is not a trivial task, and a significant part of the INTAS project is given over to this problem. At a simple level one can compare trends over specified ocean regions, but it is now widely accepted that the true nature of climate variability is complex,

Figure 4—VOS wind sea (a) and swell (b) long-term mean and difference plots against WAM wind sea (c) and swell (d) (contours in m), for the period 1979-93. Copyright Physics and Chemistry of the Earth (Pergamon). Grey shading indicates where VOS > WAM.



since it occurs on a range of temporal and spatial scales. Therefore, early work has concentrated on evaluating potentially useful statistical procedures which could be employed in a global scale study of wind and wave climate variability in different data sets (e.g. EOF analysis, SVD methods).

Because we are at the early stages of this part of the scientific programme, the following section will mostly present independent studies of patterns of variability within the three data sets. However, some early results from a comparison between trends in the ERA/WAM and VOS data from G98 are also discussed.

5.2 ALT

Initial work on the ALT data investigated the increasing trend of mean winter wave heights in the north-east Atlantic, as observed by Bacon and Carter (1991), inter alia. Figure 5 presents mean wintertime (December, January, February, March) ALT SWH from the $2^\circ \times 2^\circ$ square covering Ocean Weather Station Lima (57°N, 20°W), appended to the ship-borne wave recorder data taken at this location. A linear trend fitted to these data gives an increase of 0.33 m per decade between 1975 and 1996. It is clear that the winter of 1995-96 was unusually calm (in the context of recent years), and that up until that year a steeper trend of about 0.75 m per decade was in evidence (achieved from fitting a trend to 1975-94 data). It is too early to say whether the winter of 1995/96 represents a turning point in the long-term trend, or is merely a short-term anomaly.

To investigate the spatial nature of this trend, the altimeter data were divided into two sets, the first containing data from Geosat (1985-89), and the second

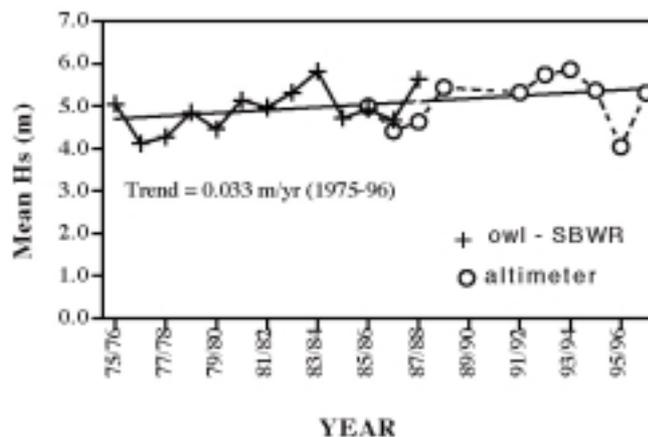


Figure 5—Mean winter significant wave heights at Ocean Weather Station Lima from ship-borne wave recorder (crosses) and altimeter (zeros) data.

containing data from ERS-1 and TOPEX/Poseidon (1991-95). Figure 6 shows the mean winter wave height over the north-eastern Atlantic for these two periods. An increase to north-east in the extent of the 5 m contour can be seen, as can a new region of 5.5 m mean winter wave height, centred on 55° 25'W.

However, this analysis does not provide any information on the variability of wave climate at different time scales. To investigate this, Cotton and Challenor (1999) used the technique of empirical orthogonal functions, employed by a number of other researchers in climate related studies (see e.g. Preisendorfer, 1988). They first fitted a simple sinusoidal model for the annual cycle from the ALT data set, then smoothed the residuals from the fit in time (five-month running mean) and in space (nine point, nearest neighbour, Gaussian filter), before extracting the highest orthogonal modes of variability (those which explain the most variance in the data) from the SWH residuals variance-covariance matrix. Whilst their study found interesting evidence of connections between the wave climates of the North Atlantic and North Pacific, we shall only consider their North Atlantic results here.

Figure 7 shows the most significant eigen mode, which accounted for over 42 per cent of the variance in the residual SWH ALT data. The North Atlantic Oscillation Index (smoothed with a five-month running mean) is also given. This figure clearly shows a bipolar structure in which the south-western North Atlantic is anti-correlated with the north-eastern North Atlantic, the dividing line running south-east from the southern tip of Greenland toward the west coast of the Iberian peninsula. This pattern matches well with the pattern identified by Kushnir *et al.* (1997) in a model wave height climatology. Through a canonical correlation analysis, coupled with sea level pressure fields, they connected this pattern to the two main phases of the NAO. When the NAO is in its negative phase (i.e. the pressure gradient across the North Atlantic is lower than normal), westerly winds over the Atlantic are weaker than usual and wave heights are lower than normal in the north-east Atlantic. In the converse case (the positive NAO phase, more common in recent years), westerly winds are stronger and hence wave heights are greater in the north-eastern Atlantic. The time series of the first

Figure 6—Mean winter significant wave heights in the north-eastern Atlantic from ALT data for the two periods 1985-89 (left) and 1991-95 (right). Contours in m.

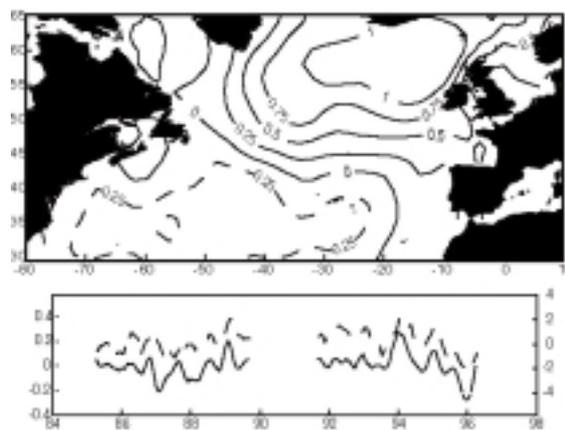
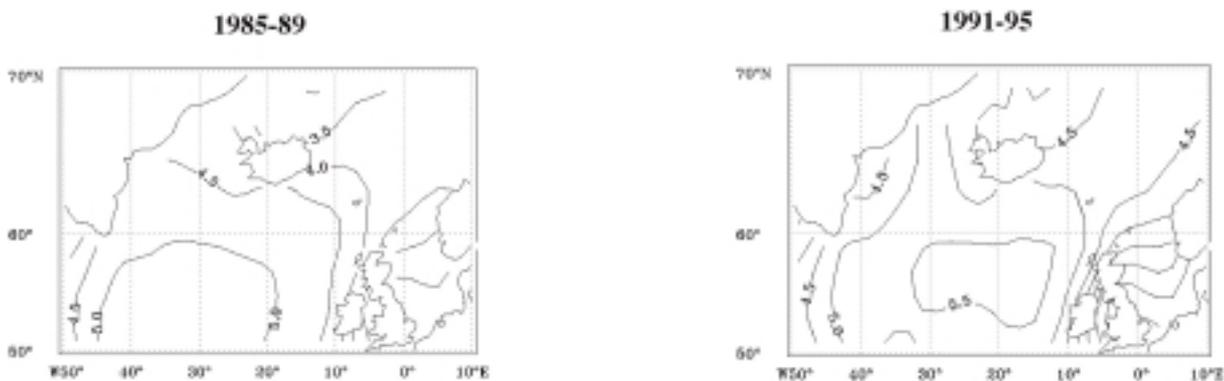


Figure 7—First EOF mode from ALT data (annual cycle removed), time series in bottom panel (solid line - LH scale), with North Atlantic Oscillation Index (dotted line - RH scale). Copyright ISOPE.

eigen mode of the altimeter data (bottom panel of Figure 7) shows the pattern was negative (i.e. lower than average wave heights in the north-east Atlantic) in the winters of 1986–87, 1987–88, and 1995–96, but positive (higher than average waves) in 1988–89, 1993–94 and 1994–95. The correlation between this time series of the first eigen and the smoothed NAO index is 0.78, confirming the strength of the connection.

The next three modes of variability (not illustrated) together explain a further 29 per cent of the variability in the data. Thus, the first four modes account for over 70 per cent of the interannual variance in the monthly mean SWH.

5.3. The 15-year WAM global wave climatology has been analysed in terms of annual
WAM cycle and trends. The largest trends in SWH were seen to occur in the North Atlantic with an increase of more than 12 cm/yr in January and south of Africa where the increasing trend exceeds 7 cm/yr in July. These trends, however, are only marginally significant. Furthermore, they exhibit a large month-to-month variability, so that on a seasonal basis the trends are significant only in small areas.

Figure 8 shows the trends in SWH and U for each calendar month averaged over the North Atlantic (40°–60°N, 10°–40°W). At the 95 per cent level, the SWH trends are only significant in April, September and October, and at the 90 per cent level also in January. The trend in SWH for the winter season (DJF) in this area reaches a maximum of 0.4 m per decade, which is not significant (but is close to the trend found in the measurements at the Ocean Weather Station Lima, 59°N 20°W). When one looks for long-term trends in the annual mean SWH (not shown here, but in S98), there remains little convincing evidence of any large-scale long-term increase. The increase in mean annual SWH in the north-east Atlantic rarely exceeds 0.1 m/decade and is significant (at the 95 per cent level) only in a very small region in the direct vicinity of Iceland. There is, however, a large area of significant negative trend in the western North Atlantic, of more than -0.15 m per decade.

Changes in wave statistics were also investigated. To this end, the 10 per cent and 90 per cent exceedance wave heights and their trends were computed for a number of regions (North Atlantic, 40°–60°N and 10°–40°W, North Pacific, 30°–60°N and 140°E–120°W, northern hemisphere, southern hemisphere, and tropics with latitudinal boundaries at 20° N and 20° S, respectively). These trends form a similar picture to those of annual mean SWH. Significant trends are only found for some months over the North Atlantic, while in the other regions the distribution of wave heights remained more or less the same over the ERA period. In the North Atlantic, the 10 per cent and 90 per cent exceedance wave heights are increasing in parallel to the annual mean SWH, the increase of which is thus accomplished by a shift of the whole wave height distribution towards higher waves.

Readers should also note the results of the WASA (1998) study, discussed in section 4.6 below.

5.4. Gulev and Hasse (1999a and 1999b) report on a study of changes in the North
VOS Atlantic VOS wave climatology between 1963 and 1993. Whilst they found signif-

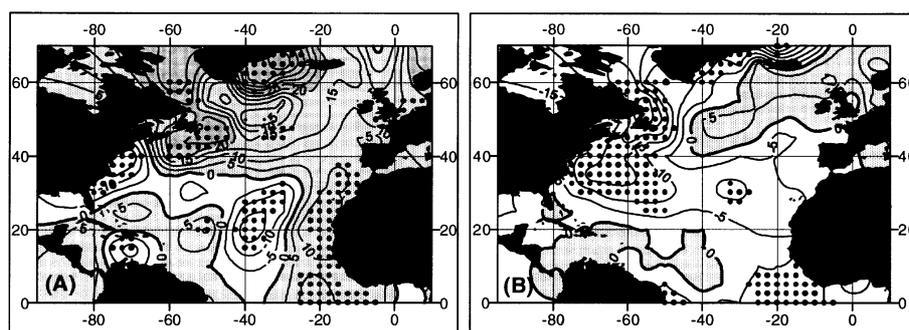


Figure 8—Linear trend of ERA/WAM SWH (solid line, m/yr) and U (dashed line, ms⁻¹/yr) averaged over a region of the North Atlantic (40°–60°N, 10°–40°W). Copyright J. Geophys Res. (AGU).

icant positive linear trends in the wind sea in the north-western Atlantic, and at mid-latitudes (0.1-0.18 m/decade), they did not find any significant trends in the north-eastern Atlantic. However, they did see strong increasing trends in swell height at mid-latitudes in the central and eastern Atlantic of 0.2-0.3 m /decade. They were able to use the directional information available from COADS data to further investigate the possible source of this increased swell, and establish its directional characteristics. When they studied these data for the region 10°-20°W and 50°-60°N, they found a negative trend in wind seas coming from the westerly directional sectors, but a positive trend in other directions. They also showed that the swell entering from the north had the largest and most significant increase.

5.5
VOS Vs. WAM

G98 compared the interannual variability in the WAM and VOS data. They separated out variability into the seasonal cycle, intraannual variability and long-term interannual variability, and then used the latter to derive estimates of long-term trends. Figure 9 compares the long-term trends from VOS and WAM. The two data sets clearly present quite different patterns. The VOS data show significant positive trends in the mid-latitude north-western Atlantic and at high-latitudes to the west of Iceland. They show a significant negative trend in the central subtropical North Atlantic. In contrast, the WAM data show a significant decreasing trend over a large part of the mid-latitude western Atlantic, but no significant positive trends anywhere. WAM and VOS show significant trends of opposite sign in the area to the south and south-west of Newfoundland. Furthermore, the long-term trend in swell is increasing in the VOS data, but decreasing in the WAM data.

The source of these major inconsistencies is not known. They may result from inconsistent partitions of the wave field into wind sea and swell in the VOS and WAM data, or may possibly be related to sampling within the VOS data (more important at high-latitudes). An equivalent analysis of altimeter data is hoped to shed some light on this.

5.6
OTHER STUDIES

Within the WASA study (WASA, 1998), a number of different analyses of past wave North Atlantic wave climate were carried out.

Buows *et al.* (1996) studied operational analyses based on ship routing charts from KNMI. In particular, they considered a box to the west of Ireland (50°-55°N, 10°-20°W), and estimated the trend between the years 1961-87. In this area they found an increasing trend of 3.8 cm per year in both the annual maximum and annual 90 per cent quantile (increases of 0.3 and 0.7 per cent respectively). Whilst it is not possible to compare these trends with those found above, we can note that this region showed an increasing trend in both the VOS and ERA/WAM climatologies.

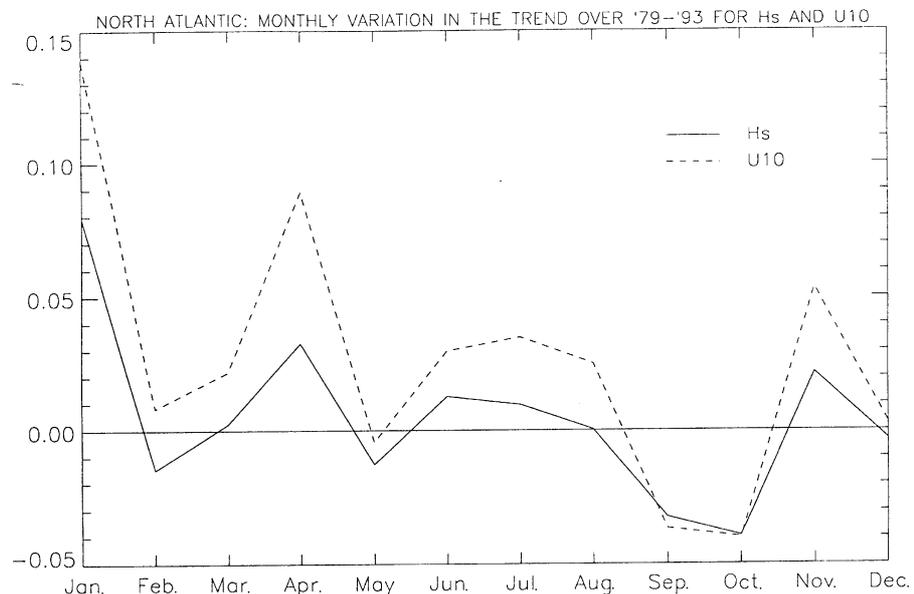


Figure 9—Estimates of the interannual linear trends (cm per decade) in VOS (a) and WAM (b) SWH climatologies in the North Atlantic. Areas marked with black dots indicate 95 per cent significance. Copyright Physics and Chemistry of the Earth (Pergamon).

In a modelling study, a regional version of fourth generation WAM (north-eastern North Atlantic, 0.5° lat. \times 0.75° long resolution) was forced by winds derived from operational pressure fields from the Norwegian Meteorological Office (DNMI), and boundary conditions provided from a coarse run with FNOC winds. This model was run for the 1955-94 period (Gunther *et al.*, 1998). The resultant chart of trends in the interannual 90 per cent quantile showed mostly negative trends in the region considered by Buows *et al.* (1996), but positive trends in the North Sea and Norwegian Sea.

6. CONCLUSIONS
6.1 'STATIC' COMPARISONS
WAM/ERA

The comparisons completed thus far indicate that the WAM/ERA data systematically overestimate low waves and underestimate high waves. This suggests an inability of the model to fully match the variability in the actual wave fields, due to resolution limitations, and a possible overestimate of swell.

ALT The ALT data have been carefully calibrated against buoy data, and no dependency of accuracy on sea state has been identified. Work continues to investigate this possibility.

VOS The VOS data are seen to overestimate low waves, possibly partly as a result of the problem with the COADS code '1'.

WIND SEA/SWELL In a comparison of these fields from ERA/WAM and VOS, there was no spatially consistent pattern of differences. VOS appeared to increasingly overestimate (with respect to ERA/WAM) towards the north-western North Atlantic. It is possible that the separation of the wave field into wind sea and swell is not consistent in the WAM and VOS climatologies.

6.2 'DYNAMIC' COMPARISONS

The patterns of climate variability are different in each data set. Interannual trends are of different magnitudes and sign. The ERA/WAM climatology shows a significant trend only in the western mid-latitude Atlantic (decreasing), whereas the VOS climatology shows significant increasing trends to the west of Iceland, and the south and south-west of Newfoundland. VOS also showed a significant decreasing trend in the central subtropical North Atlantic. ALT data indicate an increase (in the years 1985-95) centred on 55°N , 25°W .

6.3 FUTURE WORK

Future work will need to consider a number of problems highlighted in this paper:

- The consistent partition of wave fields into wind sea and swell within the VOS and WAM data.
- The carrying out of an equivalent analysis of interannual variability on the ALT data, and widening of the methodology to look at other time scales.
- Investigations into how VOS sampling may effect these comparisons.
- Consideration of whether satellite instrumentation could provide a wider range of wave parameters. e.g. synthetic aperture radar data, or the use of an altimeter-derived wave period parameter.

ACKNOWLEDGEMENTS This work has been supported by INTAS (project 96-2089), Deutsche Forschungsgemeinschaft, Sonderforschungsbereich SFB 133, the Ministry of Science and Technology of the Russian Federation under the 'world Ocean' National Programme, and the Dutch National Research Programme on Global Air Pollution and Climate Change (contract 951207).

Data have been made available by NOAA, AVISO, ESA, and Steve Worley of NCAR. Computer time was provided by ECMWF.

REFERENCES Bacon, S. and D.J.T. Carter, 1991: Wave climate changes in the North Atlantic and North Sea, *Int. J. of Climatology*, 11, 545-558.

- Buows E., D. Jannick and G.J. Komen, 1996: On increasing wave height in the North Atlantic Ocean. *Bull. Amer. Meteor. Soc.*, 77, 2275-2277.
- Challenor, P.G. and P.D. Cotton, 1999: The joint calibration of altimeter and in situ wave heights, JCOMM Technical Report No. 13, WMO/TD-No. 1081.
- Cotton P.D. and D.J.T. Carter, 1994: Cross calibration of TOPEX, ERS-1 and Geosat wave heights, *J. Geophys. Res.*, 99, 25025-25033. (correction *J. Geophys. Res.*, 100, 7095).
- Cotton, P.D. and P.G. Challenor, 1999: North Atlantic wave climate variability and the North Atlantic oscillation index, *proc ninth (1999) ISOPE conf* (Brest, France, 30 May-4 June 1999), 3, pp. 153-157.
- Cotton, P.D., P.G. Challenor and D.J.T. Carter, 1997: An assessment of the accuracy and reliability of Geosat, ERS-1, ERS-2, and TOPEX altimeter measurements of significant wave height and wind speed, *proc. CEOS Wind and Wave Validation Workshop* (3-5 June 1997, Noordwijk, The Netherlands), ESA wpp-147.
- Davies, C.G., P.G. Challenor, P.D. Cotton and D.J.T. Carter, 1998: *Proc. CEOS Wind and Wave Validation Workshop* (3-5 June 1997, Noordwijk, the Netherlands), ESA wpp-147. 75-80.
- Gower, J.F.R., 1996: Intercalibration of wave and wind data from TOPEX/POSEIDON and moored buoys off the west coast of Canada, *J. Geophys. Res.*, 101, 3817-3829.
- Gulev, S.K. 1999: Evaluation of ocean wind and wave fields from COADS. *Proceedings of CLIMAR99*, JCOMM Technical Report No. 13, WMO-TD 1062.
- Gulev, S.K., D. Cotton and A. Sterl, 1998: Intercomparison of the North Atlantic wave climatology from voluntary observing ships, satellite data and modelling. *Physics and Chemistry of the Earth*, 23, No. 5/6, pp. 587-592.
- Gulev, S.K. and L. Hasse, 1998: North Atlantic wind waves and wind stress from voluntary observing data. *J. Phys. Oceanogr.*, 28, 1107-1130.
- Gulev, S.K. and L. Hasse, 1999a: Changes of wind waves in the North Atlantic over the last 30 years. *Int. J. Climatol.*, in press.
- Gulev, S.K. and L. Hasse, 1999b: Climate changes of wind waves in the North Atlantic over the last several decades, *proc. ninth (1999) ISOPE conf*. (Brest, France, 30 May-4 June, 1999), 3, pp. 164-167.
- Gunther, H., W. Rosenthal, M. Stawarz, J.C. Carretero, M. Gomez, I. Lozano, O. Serano and M. Reistad, 1998: The wave climate of the north-east Atlantic over the period 1955-1994. The WASA wave hindcast. *Global Atmos. Ocean Sys.* In press.
- Kushnir, Y., J.G. Greenwood and M.A. Cane, 1997: The recent increase in North Atlantic wave heights. *J. Climate*, 10, 2107-2113.
- Preisendorfer, R., 1988: *Principal Component Analysis in Meteorology and Oceanography*, Elsevier Pub. Co., N. Y. 425 pp.
- Sterl, A., G.J. Komen and P.D. Cotton, 1998: 15 years of global wave hindcasts using ERA winds: Validating the reanalysed winds and assessing the wave climate. *J. Geophys. Res.*, 103, 5477-5492.
- WASA Group, 1998: Changing waves and storms in the Northeast Atlantic? *Bull. Amer. Meteor. Soc.* 79, 741-760.
- Woodruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Physics and Chemistry of the Earth*, 23, 517-527.

THE JOINT CALIBRATION OF ALTIMETER AND IN SITU WAVE HEIGHTS

P.G. Challenor¹; P.D. Cotton²

ABSTRACT The changes in the world's wave climate are subtle, and to investigate them we need long-term and well-calibrated global measurements. One source of good quality consistent measurements of significant wave height is the radar altimeter. Such instruments have flown on a number of satellites and, apart from a short gap in 1989-1991, we have continuous global data since 1985. However, this data set involves a number of different satellites and sensors, each of which has a slightly different calibration.

In this paper we look at the problem of producing a coherent well-calibrated set of buoy and satellite altimeter data. In the classical method of calibration, a well-known and more accurate standard is used to calibrate an instrument. If we try to calibrate a radar altimeter against a set of wave buoys we do not have such a standard. The buoys are no more accurate than the altimeter itself. Thus, we need to use more sophisticated statistical techniques than simple linear regression which can take into account errors in both variables. We present calibration results for all radar altimeters since Geosat and discuss the drift in the TOPEX measurement of wave height. We demonstrate that it is necessary to apply these calibration results to altimeter data if measurements from different satellites are to be used to assemble multi-year climate data sets.

In addition, we discuss the possible use of radar altimeters as 'standards' for the cross-calibration of buoys around the world. We compare results from four different buoy data sets (operated by the US NDBC, Canadian MEDS, the UK Met Office, and the Japan Meteorological Agency). We demonstrate that the biggest obstacle to generating a coherent blended buoy/in situ data set are different reporting standards. We will also discuss the comparison of altimeter data with wave information from Voluntary Observing Ships (VOSs) using comparisons between individual satellite and ship observations.

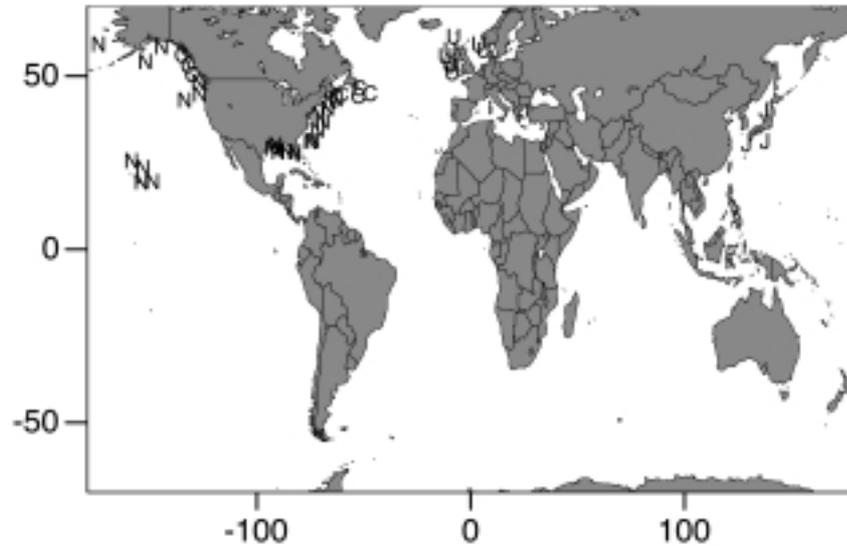
INTRODUCTION To produce a coherent, long-term wave climatology for the world's oceans we need to be able to combine data from a number of sources. In particular, we need to use data from buoy networks, satellites and VOSs. If well-maintained, buoys can produce good quality regular data, not only significant wave height but also other spectral parameters including directional information. However, the number of buoys deployed at any one time is limited, and buoy networks will never produce more than a very limited areal coverage. Over the last few years the Southampton Oceanography Centre (SOC) has been trying to discover as many buoy deployments in deep water as possible. So far we have only found four significant networks. These are deployed around the USA, Canada, Japan and the UK. Other data are available from oil companies, but these data are often of a short duration and are sometimes confidential. We have discovered no buoy data from the Southern Ocean. The positions of these buoys are shown in Figure 1.

Our second source of data is radar altimetry from satellites. Apart from a short gap in 1991, radar altimeters have been continuously flying since 1985. Altimeters produce good quality significant wave height information (see Carter *et al.*, 1992; Cotton and Carter, 1994 and below). Altimeters also measure wind speed, and recently it has been shown that it is possible to extract data on wave

1 James Rennell Division for Ocean Circulation and Climate, Southampton Oceanography Centre, Southampton, SO14 3ZH UK

2 Satellite Observing Systems Ltd.

Figure 1 — The positions of the NDBC (N), UK Met Office (U), Japan Meteorological Agency (J) and the Meteorological Service of Canada (C) buoys.



period as well (Davies *et al.*, 1997), but we will concentrate on significant wave height in this paper. In general, altimeters have been shown to produce significant wave height data with a similar accuracy to wave buoys, but with a bias (Carter *et al.*, 1992). This is discussed below.

Although altimeters deliver data over the entire globe, there are gaps. For instance, TOPEX/POSEIDON has an inclination of only 66° so no data are recovered poleward of a latitude of 66° . The altimeter also has no swath, so data are only collected directly beneath the satellite. This means that there are gaps between satellite tracks where no data are ever collected. The size of these gaps depends upon the repeat period of the satellite. The more often the ground track is repeated, the larger the gap between tracks. An example of TOPEX/POSEIDON tracks is shown in Figure 2. This pattern is repeated every ten days. Note that the track separation is not constant but varies with latitude.

An alternative to altimeter data which also has quasi-global coverage are wave observations from ships of opportunity. These are not instrumental as with the other data but consist of subjective estimates of wind sea and swell height, direction and period. Although in principle the visual ship data are global, in practice there are very large gaps away from the shipping lanes, particularly in the Southern Ocean.

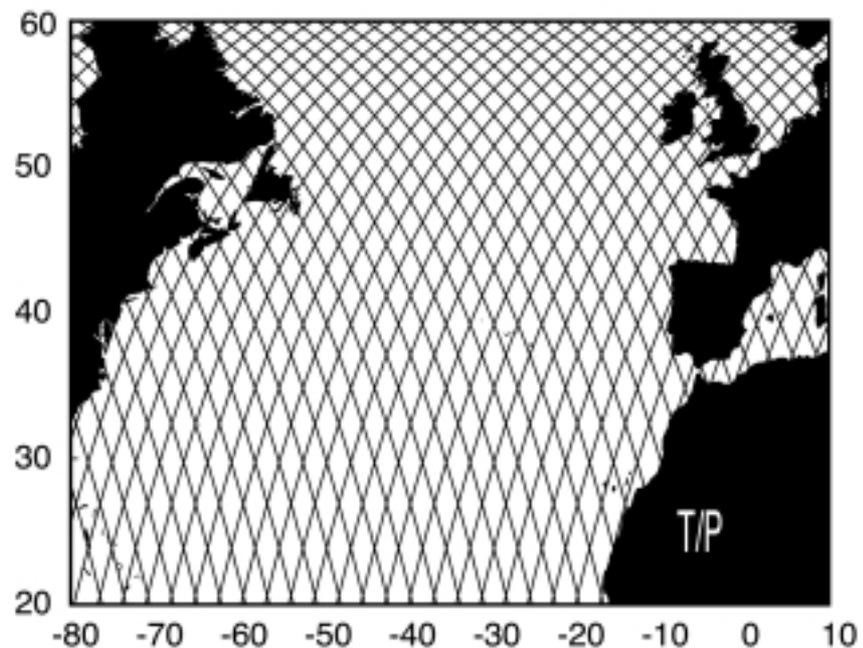


Figure 2 — The coverage from the TOPEX/POSEIDON altimeter over the North Atlantic. Data are only collected directly below the satellite.

In this paper we will look at these three sources of information about significant wave height (H_s) and how we can make them consistent. Our basic plan is to use the US NDBC buoys as a standard and calibrate everything relative to these using calibrated altimetry as a transfer standard. Unlike the standard calibration problem, however, we cannot assume that our standard, the NDBC buoys, have such a small error that it can be neglected. We know from previous work (e.g. Cotton and Carter, 1994) that the random error on buoys and altimeters is similar, and that the visual data from ships can be expected to have a larger error, but even then, we cannot assume that the altimeter error is zero. Standard regression techniques demand that the 'x' variable is without error, so we need to use a more sophisticated method which does not make this assumption. There are a number of such techniques available and we will use two of them. The simplest is principal component regression. Here we take the line which passes through the mean of the two data sets and has a slope equal to the geometric mean of the 'x on y' regression and the 'y on x' (this line is also the first principal component of the data, hence the name). This is appropriate when the variables being regressed have approximately the same errors. We use this technique for comparing altimeter and buoy data. A more complex method which can be used in situations where the errors cannot be assumed to be the same, or where more complicated linear (or even non-linear) models are required, is orthogonal distance regression (ODR) (Boggs and Rogers, 1990). This minimizes the orthogonal distance to be line from the 2-d dataset and provides error estimates for both the 'x' and 'y' variables. We will use this method to find trends in altimeter data and for altimeter/COADS comparisons. A comparison of the ODR and principal component regression for the buoy/altimeter comparisons showed negligible differences.

THE NDBC DATA SET

The 'reference' set of buoys we use consists of 24 buoys around the US coast which are run by the US NOAA Data Buoy Center (NDBC). We have selected these buoys because they are in deep water and are not too close to any coasts and, therefore, should be representative of deep water conditions. The buoy locations can be divided into four areas: the North Pacific (including buoys off Alaska), the North Atlantic, the Gulf of Mexico and off Hawaii. This data set therefore covers a range of conditions from areas off Hawaii where swell dominates, to places such as the coast of Alaska where the wave climate is dominated by large storms.

CO-LOCATION CRITERIA

Initially we followed other authors and used co-location criteria of 1 h and 100 km. However, after experimenting with varying the criteria, we decided that the optimum for use with the NDBC buoys (which report hourly) was to use 30 minutes and 50 km. The altimeter data were not averaged in any way and the nearest 1 Hz value, as provided by the space agencies, was used.

ALTIMETER CALIBRATIONS

GEOSAT

Geosat was a US Navy satellite that operated from 1985-1989. The early part of the mission was in a very long repeat orbit (168 days) to make geodetic measurements, while from 1987 the satellite went into a 17-day repeat. During 1989, the satellite started to degrade and the data became much less reliable.

The calibration equation for the years 1985-1988 is given by (standard errors in parenthesis):

$$H_s(\text{Geosat}) = -0.0943 + 0.9092H_s(\text{NDBC}) \quad \text{rrms}=0.28 \text{ m}$$

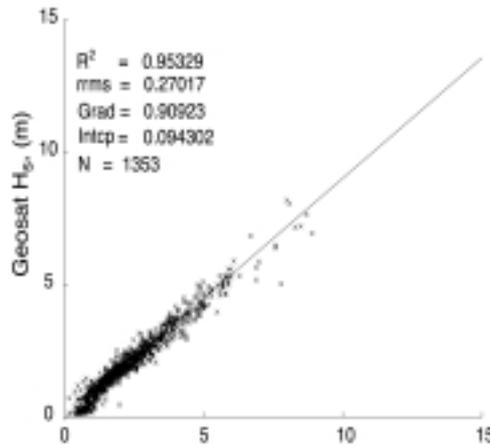
$$(0.0142) \quad (0.0054)$$

This is shown in Figure 3. The inclusion of the 1989 data gives a slightly different equation:

$$H_s(\text{Geosat}) = -0.0798 + 0.8976H_s(\text{NDBC}) \quad \text{rrms}=0.38 \text{ m}$$

$$(0.0190) \quad (0.0073)$$

Figure 3 — Geosat H_s plotted against NDBC Buoy H_s for the period 1985–1988. The line shows the best fit calibration.



ERS-1 The next altimeter to be launched was ERS-1. There are two main classes of data from the ERS series of satellites. There are fast delivery (FD) data that are produced within three hours of collection. One use of this sort of wave data is assimilation into wave forecast models. The other class of data is off-line products (OPR) data. These data are reprocessed on the ground and are, therefore, more accurate than the fast delivery data. For climatological purposes we are interested in this latter class of data, so we will only discuss the calibration of the off-line products. To further confuse matters two versions of the off-line data for ERS-1 are available. Data collected between the launch of the satellite in August 1991 and March 1995 form part of version 3. A new version of the processing software was then introduced (version 6), and this was used until May 1996 when the ERS-1 satellite was put into 'storage' and ERS-2 became the source of data.

For the version 3 data we obtain a calibration equation:

$$H_s(\text{ERS-1}) = -0.3025 + 0.9016H_s(\text{NDBC}) \quad \text{rrms}=0.45 \text{ m}$$

(0.0229) (0.0094)

This is shown in Figure 4.

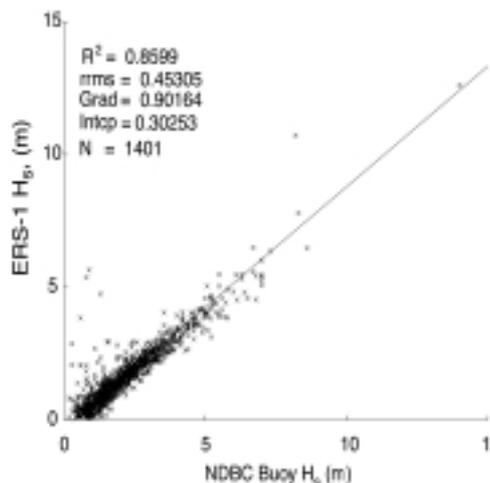


Figure 4 — ERS-1 OPR(v3) H_s plotted against NDBC buoy H_s . The line shows the best fit calibration.

Early forms of the version 6 software were faulty, resulting in lower quality H_s data for ERS-1 cycles 144-148, covering the period 04/95 to 07/95. Subsequent data are of better quality and gave the calibration equation:

$$H_s(\text{ERS-1}) = -0.1906 + 0.8871H_s(\text{NDBC}) \quad \text{rrms}=0.36 \text{ m}$$

(0.0444) (0.0181)

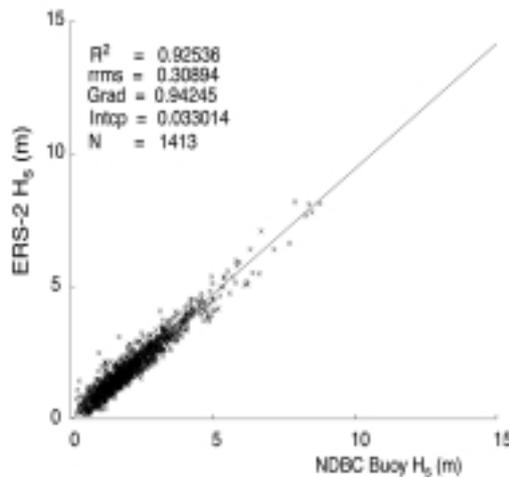
ERS-2 ERS-2 succeeded ERS-1 and was launched in April 1995. Although it is still producing data, we only use data to the end of 1997. All the data were processed with version 6 of the software but without any of the problems associated with ERS-1. The calibration is given by:

$$H_s(\text{ERS-2}) = -0.0330 + 0.9425H_s(\text{NDBC}) \quad \text{rrms}=0.30 \text{ m}$$

$$(0.0163) \quad (0.0070)$$

The plot is shown in Figure 5.

Figure 5 — ERS-2 H_s plotted against NDBC buoy H_s . The line shows the best fit calibration.



TOPEX/POSEIDON

The TOPEX/POSEIDON satellite is a US/French mission that was launched in 1992. The satellite has two on-board altimeters: the US TOPEX and the French POSEIDON. Because they share certain hardware components, in particular the antenna, both altimeters cannot operate at the same time. TOPEX operates for 90 per cent of the time, with POSEIDON providing data for the remaining 10 per cent.

The calibration of POSEIDON is shown in Figure 6. The calibration equation is:

$$H_s(\text{POSEIDON}) = -0.0340 + 1.0214H_s(\text{NDBC}) \quad \text{rrms}=0.28 \text{ m}$$

$$(0.0362) \quad (0.0154)$$

The final altimeter we shall consider in this section is TOPEX. Using all the data from 1992 to 1997, the equation:

$$H_s(\text{TOPEX}) = -0.0895 + 0.9503H_s(\text{NDBC}) \quad \text{rrms}=0.26 \text{ m}$$

$$(0.0113) \quad (0.0048)$$

is obtained. The fit is shown in Figure 7.

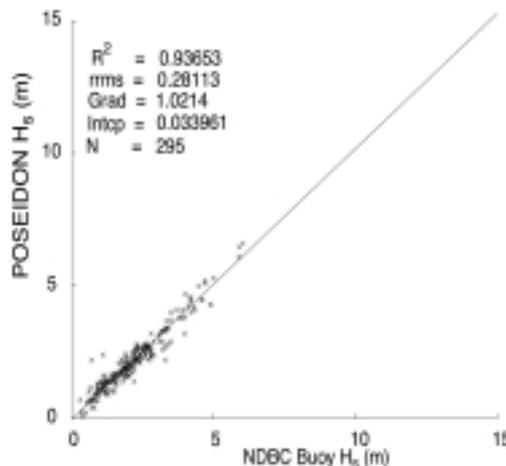


Figure 6 — POSEIDON H_s plotted against NDBC buoy H_s . The line shows the best fit calibration.

However, if we plot the daily mean difference between the buoys and the TOPEX altimeter against time, as shown in Figure 8, we find that there is a trend apparent in the latter part of the data. It appears that the instrument characteristics have been changing since the launch (Hayne, pers. comm.), but the effect on the estimated significant wave height only becomes apparent towards the end of 1996. We estimate that after day 1730 since the start of 1992 (26th September 1996), there is a trend of 0.4 mm day⁻¹ in the TOPEX significant wave height measurement. In January 1999, the TOPEX electronics were switched to the alternative 'B' side and since then there has been no discernible trend; however a new calibration of wave height is now required. This is given by:

$$H_s(\text{TOPEX} - \text{B}) = -0.0800 + 0.9676H_s(\text{NDBC}) \quad \text{rms}=0.19 \text{ m}$$

(0.0357) (0.0185)

Figure 7 — TOPEX H_s plotted against NDBC buoy H_s . The line shows the best fit calibration.

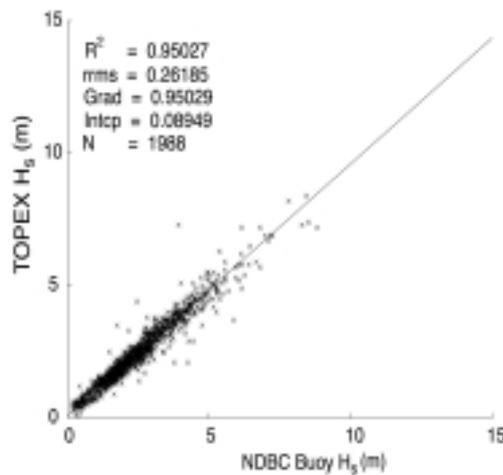
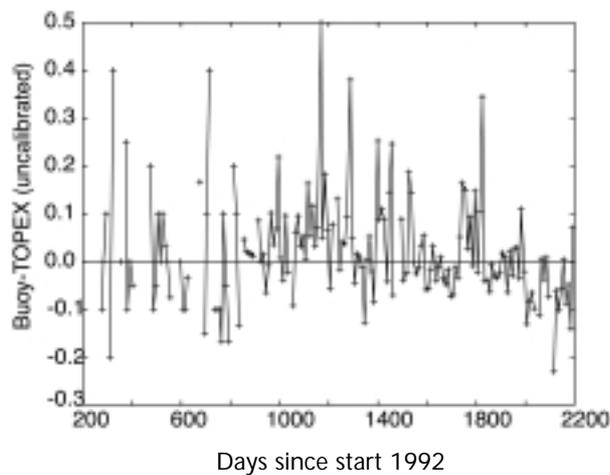


Figure 8 — Ten day averages of the difference between TOPEX and NDBC buoy significant wave heights. Note the trend in the latter part of the data.



COMPARISONS WITH OTHER BUOY NETWORKS

We have now calibrated the altimeter data to be internally consistent and consistent with the NDBC buoy network. The consistency of the calibration across missions is such that we can now use the calibrated altimeter set to look at and map wave climate change in the North Atlantic (Cotton and Challenor, 1999).

Once we have this 'standard' data set we can use it to check the calibrations on other instruments and measuring systems. In this section we look at the consistency of the buoy networks around the world. In the next section we do some initial work with visual wave observations from the COADS data set.

As stated in the introduction, we have data from three buoy networks in addition to the NDBC data used so far. These are the UK Met Office (UKMO), the Japan Meteorological Agency (JMA) and the Meteorological Services of Canada (MSC) (supplied to us by the Canadian Marine Environmental Data Service

(CMEDS)) buoy networks. Their positions are shown in Figure 1. If these data are calibrated to the same 'standard' as the NDBC buoys, then a comparison with the calibrated altimeter set will be the same for all buoy networks. Note that if we have a difference between calibrations in the buoys, we cannot say which is right with this method; we can only say whether they are consistent with the NDBC buoys. A summary of the buoy data sets is given in Table 1. One of the differences that must be taken into account is the reporting standards of each network. For example, the JMA buoys only report every three hours, whereas the other three networks report every hour. To obtain a meaningful number of co-locations with the satellites we have had to relax the time co-location criterion to 1 h rather than 30 minutes.

Figures 9 to 11 show scatterplots of the combined calibrated altimeter data against each of the buoy networks. Although it may seem circular, we have included the NDBC data in this analysis both as a check on our analysis and as a means of identifying possible 'rogue' buoys. See Fedor and Brown (1982) for an example where an apparently miscalibrated buoy is identified. The comparisons with the other buoy networks are more interesting.

It is immediately apparent that the UKMO buoys only report significant wave height to the nearest 0.5 m. This will clearly be reflected in the calculation of any accuracies. Table 2 gives the details of the buoy 'calibrations'. There are significant differences between the buoy networks in terms of their slopes (UKMO, MSC) or intercept (JMA). Thus, we expect UKMO buoys to read about 4 per cent high compared to NDBC, MSC to be 5 per cent low and the JMA buoys to have a bias of about 30 cm. We stress again that these are relative measures and we cannot say which calibration is correct. The residual rms values for the non-NDBC buoys are higher. Because we fitted the altimeter to the NDBC set, this rms is depressed relative to the other buoy networks, so any comparisons must be made with caution. To get a true measure of the NDBC rms we should hold back data from the fitting process and use these independent data to estimate the rms. However, it is clear from Figure 10 that there are a number of very poor comparisons between MSC and the altimeter data set for low buoy Hs values. Removing these has little effect on the regression line but does reduce the rms. Similarly, the rms for the UKMO data are increased by the 0.5 m resolution. Degrading the NDBC data to 0.5 m resolution increased the rms from 0.335-0.354 m. As regards the JMA data, the co-location criteria were relaxed since the buoys only report every three hours, and this will increase the rms.

FURTHER STUDIES

When all the calibrated altimeter-NDBC buoy co-located data are combined, they provide us with a large data set (about 5 500 data pairs) with which to study possi-

<i>Source</i>	<i>Coverage</i>	<i>Data Type</i>	<i>Dates</i>
US NOAA Data Buoy Center (NDBC)	24 selected buoys in N. Atlantic, N. Pacific, Caribbean Sea	Hourly wave spectra, met. data	1972-97
UK Met Office (UKMO)	7 open ocean buoys in N. E. Atlantic and North Sea	Hourly summary wind and wave data, met. data	1991-97
Japan Meteorological Agency (JMA)	3 open ocean buoys around Japanese Coast	3 hourly summary wind and wave data, met. data	1985-96
Meteorological Services of Canada (MSC)	7 open ocean buoys. N. Atlantic, N. Pacific	Hourly wave spectra, met. data	1988-96

Table 1 - Sources of Buoy Wave Data.

Figure 9 — NDBC buoys plotted against the combined, calibrated altimeter data set.

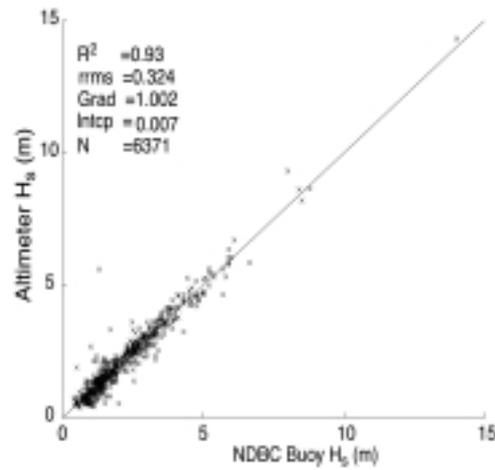


Figure 10 — MSC Buoy H_s plotted against the combined, calibrated altimeter data set.

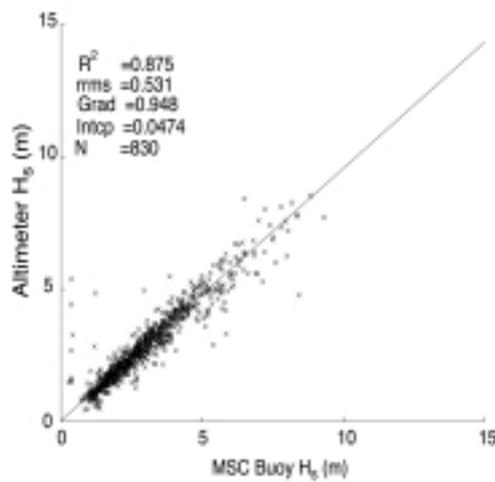


Figure 11 — JMA Buoy H_s plotted against the combined, calibrated altimeter data set.

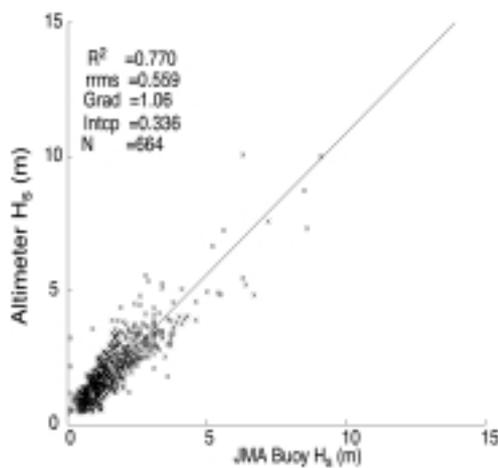


Table 2—Principal component regression parameters from comparisons of co-located altimeter and buoy significant wave height data. *Co-located data within nearest hour, rather than 30 minutes.

Data Source	No	Slope	Std. err.	Int. (m)	Std. err.	rms (m)
NDBC	6371	1.002	0.007	-0.007	0.016	0.325
UKMO	1228	1.041	0.021	-0.124	0.072	0.604
JMA*	664	1.062	0.041	0.337	0.080	0.559
MSC	830	0.948	0.024	0.047	0.079	0.531

ble dependencies of alt/buoy wave measurement. One such study tested for a possible dependency of the altimeter/buoy H_s relationship on buoy size. Although the buoy platforms range from 3-12 m in diameter, we found no significant change in the gradient (or intercept) of the ODR regression which might have indicated a change in the sensitivity of the buoy measurement.

Furthermore, it had been suggested (Janssen, pers. comm., 1998) that the altimeter measurement may be less accurate under certain sea conditions (e.g. steep young seas), when the assumption of Gaussian distributions of sea surface heights may not hold. To test this assertion the normalized altimeter-buoy error in H_s was plotted against wave age. No dependency on wave age was found. Further tests are planned against buoy data which contain more spectral information.

COMPARISONS WITH COADS DATA

The third source of global, or near-global, wave height information are visual observations from ships of opportunity. Such data are collected in the COADS data set. Unlike our other data, the visual estimates do not give a simple estimate of significant wave height. Instead, there are at least two estimates, one for the wind sea (or waves travelling in the same direction as the local wind) and one for swell. (Occasionally secondary swell trains are also identified but this is rare and we ignore such data). There are a number of formulae to compute an estimate of significant wave height from these two components. Hogben (1988) uses the formula:

$$H_s(\text{Hogben}) = \sqrt{h_w^2 + h_s^2}$$

where h_w and h_s are the wind sea and swell estimates, respectively. Wilkerson and Earle (1990) use the maximum of h_w and h_s , whereas Barratt (1991) uses a combination of the two: Hogben's formula when the direction of the wind sea and swell differs by less than a certain angle, and Wilkerson and Earle's when it is greater. Gulev and Hasse (1998) suggest using an angle of thirty degrees. Although we have analysed all three definitions, since the results were very similar, we will only report the results for Hogben's definition.

It is non-trivial to co-locate data from moving ships with altimeters which, because of orbital dynamics, have a complex sampling of the sea surface. So far we have co-located COADS for the three years 1993-1995 inclusive with altimeter measurements from TOPEX. Over the three years, this gives us 21 150 data points with a visual estimate of either sea or swell from the ships. Using orthogonal distance regression we obtain the following equation:

$$H_s(\text{Hogben}) = -0.5331 + 1.0274 H_s(\text{Alt})$$

which means that on average the individual visual estimates of significant wave height are fairly good although both the slope and intercept are significantly different from 1 and 0, respectively, at the 95 per cent level. The intercept is higher than for the buoys but may reflect our decision to place the visual estimates at the top of the range, so an estimate between 1 and 1.5 m was set to 1.5 m in calculating H_s . However, the residual root mean square is 1.04 m, showing that while good climatologies should result from averaging large quantities of visual data, individual observations should be used with caution.

CONCLUSIONS

We have shown that by calibrating against a buoy network, in our case the NDBC buoys, it is possible to produce a consistent long-term inter-mission altimeter data set for significant wave height. The accuracy of each individual data point in this data set is better than 0.5 m. However, the drift in the TOPEX altimeter has shown the need for continual monitoring of satellite systems throughout their lives, rather than simply relying on a three- or six-month 'calibration' phase at the start of the mission. This implies that we need well-maintained and calibrated buoy networks to provide such calibrations. The altimeter systems are a valuable addition to the buoy networks and not a substitute for them. Ad hoc deployments of buoys for special purposes (including satellite calibration!) are of much less use. It is difficult to validate the altimeters at both very high and very low wave heights. Since high

sea states are rare, we have little data to work with, while all measurements appear to have difficulty in measuring waves lower than about 0.5 m H_s . A further problem is the lack of calibration data in the southern hemisphere. We can expect wave conditions to be different here with larger fetches and more swell. We would therefore like to confirm our calibrations and the altimeter algorithms in these regions.

Once created, we can use the combined altimeter set as a 'standard' to check the calibration of other wave measuring systems, both long-term networks and ad hoc deployments. Our work on this so far shows that there are differences between the calibrations of the different buoy networks and a large proportion of these are probably due to different reporting standards and quality control. If all buoy operators worked to the same standards, we believe that most of these differences would disappear. So far, the work carried out on COADS has been very limited but would appear to show that on average the data are of good quality (with a possible bias). The quality is very variable though, and the data should only be used in averages.

ACKNOWLEDGEMENTS

Simon Keogh carried out the co-location of the altimeter and COADS data. We would like to thank the US NDBC, the UKMO, JMA, MSC and the Canadian Marine Environmental Data Service for permission to use their buoy data. Altimeter data were supplied by NOAA (Geosat), ESA (ERS) and AVISO (TOPEX/POSEIDON). The COADS data were supplied by NOAA. Some of the calculations were carried out using ODRPACK written by P.T. Boggs, R.H. Byrd, J.E. Rogers, and R.B. Schnabel and available from <http://netlib.bell-labs.com/netlib/odrpac>. Finally, a proportion of this work was funded by the British National Space Centre as part of its ENVISAT Exploitation Initiative.

REFERENCES

- Barratt, M. J., 1991: *Waves in the North-East Atlantic*. Department of Energy Report OTI 90545, HMSO.
- Boggs, P.T. and J.E. Rogers, 1990: Orthogonal distance regression. *Contemporary Mathematics*, 112, 183-194.
- Carter, D.J.T., P.G. Challenor and M.A. Srokosz, 1992: An assessment of Geosat wave height and wind-speed measurements. *Journal of Geophysical Research-Oceans*, 97(C7), 11383-11392.
- Cotton, P.D. and D.J.T. Carter, 1994: Cross calibration of TOPEX, ERS-1 and GEOSAT wave heights. *Journal of Geophysical Research*, 99(C12), 25025-25033.
- Cotton, P.D. and P.G. Challenor, 1999: North Atlantic wave climate variability and the North Atlantic Oscillation Index. *Proceedings of the Ninth International Offshore and Polar Engineering Conference* (Brest, France).
- Davies, C.G., P.G. Challenor, P.D. Cotton and D.J.T. Carter, 1997: Validation of wave period measurements from radar altimeter data. CEOS Wind and Wave Validation Workshop, ESTEC, Noordwijk, The Netherlands, 75-80.
- Fedor, L.S. and G.S. Brown, 1982: Waveheight and wind speed measurements from the SEASAT radar altimeter. *Journal of Geophysical Research*, 87(C5).
- Gulev, S.K. and L. Hasse, 1998: North Atlantic wind waves and wind stress fields from Voluntary Observing Ship data. *Journal of Physical Oceanography*, 28, 1107-1129.
- Hogben, N., 1988: Experience from compilation of global wave statistics. *Ocean Engineering*, 15, 1-31.
- Wilkerson, J.C. and M.D. Earle, 1990: A study of differences between environmental reports by ships in the voluntary observing program and measurements from NOAA buoys. *Journal of Geophysical Research*, 95, 3373-3385.

ON THE USE OF IN SITU AND SATELLITE WAVE MEASUREMENTS FOR EVALUATION OF WAVE HINDCASTS

Andrew T. Cox and Vincent J. Cardone¹; Val R. Swail²

ABSTRACT Two long-term wave hindcasts based on the NCEP/NCAR Reanalysis (NRA) products have recently been completed. The Global Reanalysis of Ocean Waves (GROW) project ran 40 years of unmodified NRA winds on a global 1.25×2.5 degree wave grid. The AES40 project ran 40 years of re-hindcast wind fields based on the NRA products on a high resolution North Atlantic grid.

This paper discusses the use of in situ and satellite wave measurements in evaluating and understanding the bias and skill in these wave hindcasts. Direct time-series, quantile-quantile, and other statistical properties of the wave hindcasts are presented. Comparisons of the change in wave height bias at buoy locations over the 1975-1997 period are evaluated to assess the homogeneity of the wave hindcasts over this period. Finally, regional statistical comparisons and spatial plots of wave height bias and scatter derived from satellite measurements are also included.

1. INTRODUCTION

The ocean wave climate has long been of interest to the ocean engineering community because of the need for accurate extreme and operational wave data for applications such as vessel design, specification of peak loads of coastal and offshore structures, and planning of naval and marine operations. In recent years, there has been a major resurgence of interest in wave climate within the scientific community as a result of indications of worsening storm wave regimes in some areas (e.g. Bacon and Carter, 1991; Athanassoulis and Stefanakos, 1995) and evidence that trends and variability in wave climate on a regional basis may be linked to more familiar modes of atmospheric climate trend and variability such as the North Atlantic Oscillation (NAO). Even the response of the global wave climate to a possible global warming scenario has been studied using a GCM model (WASA, 1998).

Recently, two long-term wave hindcasts based on the NCEP/NCAR Reanalysis project (henceforth NRA, Kalnay *et al.*, 1996) have been completed. In order to access their use for operational use and climate trend analysis, the skill and bias of the hindcasts over time must be validated. This study describes the use of in situ and satellite measurements in validating each of the wave hindcasts.

This paper is organized as follows. Section 2 briefly describes the two wave hindcasts being evaluated, while section 3 discusses the in situ and satellite data sets. Sections 4 and 5 present the validation results, while section 6 gives our conclusions.

2. WAVE HINDCASTS A. GLOBAL REANALYSIS OF OCEAN WAVES (GROW)

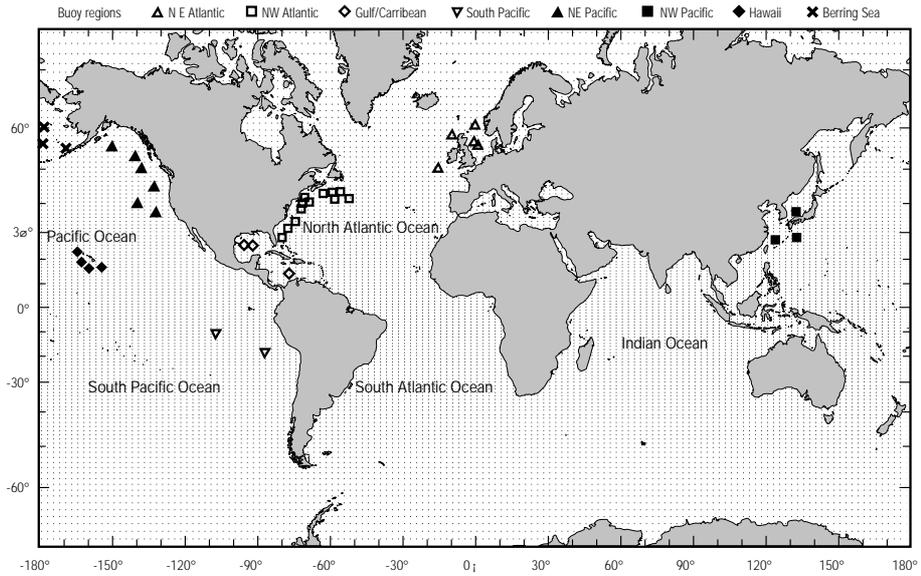
GROW was carried out by Oceanweather Inc. using a deep-water version of its proven spectral ocean wave model ODGP2, as described in Khandekar *et al.* (1994). The model was adapted to a grid spacing of 1.25° in latitude by 2.5° in longitude on a global projection as shown in Figure 1. The model was run in deep mode with first generation (1G) formulation. Ice tables were provided from long-term mean monthly historical ice concentration data.

The wind fields driving the GROW model were the NRA 10 m wind fields. The NRA winds were adjusted to neutral stability using the technique described

1 Oceanweather Inc. - Cos Cob, CT

2 Environment Canada - Toronto, Ontario

Figure 1 — GROW wave model grid with buoy locations.



by Cardone *et al.* (1990) and interpolated onto the wave model grid. No other adjustments were made to the input wind fields. Additional information on GROW can be found in Cox and Swail (1999).

B. NORTH ATLANTIC 40-YEAR REFERENCE WIND AND WAVE CLIMATOLOGY (AES40)

AES40 used the same ODGP2 wave model as GROW; however, it was run with third generation (3G) physics (see Khandekar *et al.* (1994) for description) on a higher resolution grid. The wave model grid (Figure 2) is 0.625° in latitude and 0.833° in longitude on a projection covering the North Atlantic. The southern boundary along the equator was updated with interpolated wave spectra from GROW to preserve any South Atlantic swells. Ice tables were updated monthly, rather than the long-term monthly averages used in GROW.

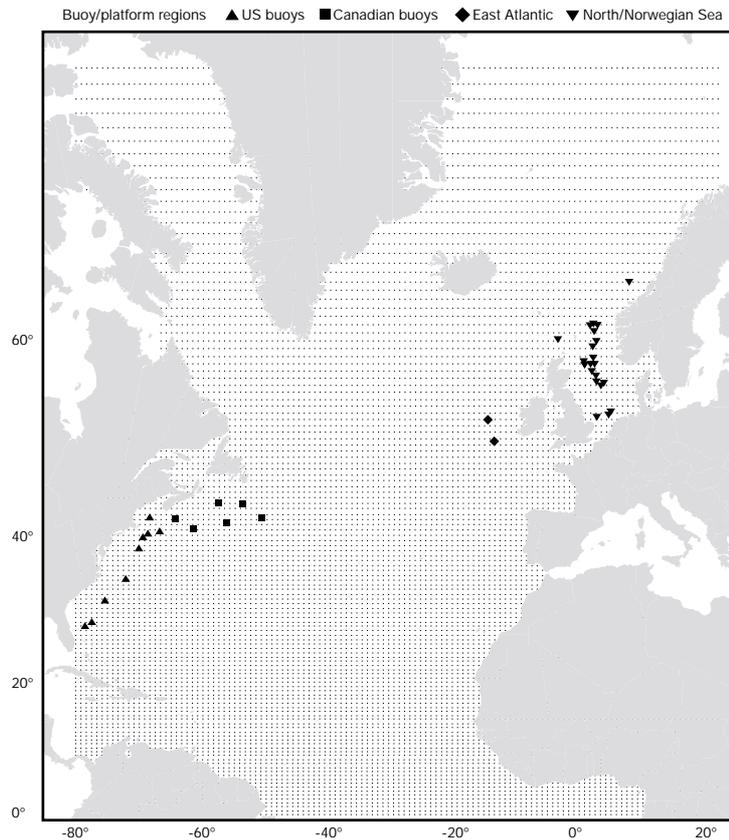


Figure 2 — AES40 wave model grid with buoy/platform locations.

The most striking difference between GROW and AES40 is in the generation of the input wind fields. NRA adjusted winds were used as primary wind inputs; however, modifications of intense extratropical storms were performed using interactive kinematic techniques. Furthermore, all tropical systems in the 40-year period were hindcast using a proven tropical boundary layer model and were included in the final wind fields. Finally, ships, buoys and satellite wind measurements were assimilated after adjusting each to a reference level of 10 m. Further information on AES40 can be found in Swail and Cox (1999).

3. VALIDATION DATA SETS

A. BUOYS AND PLATFORMS

The in situ validation data set included buoys and measurement platforms mainly located in the northern hemisphere along the continental margins (Figures 1 and 2). The in situ measured wind and wave data came from a variety of sources. US buoy data came from the NOAA Marine Environmental Buoy Database on CD-ROM; the Canadian buoy data came from the Marine Environmental Data Service marine CD-ROM; the remaining buoy and platform data (notably the north-east Atlantic and north-west Pacific) came from the Comprehensive Ocean-Atmosphere Data Set (COADS) data set described by Slutz *et al.* (1985). Comparisons were restricted to well-exposed deep-water sites with the longest records. The wave measurements are comprised of 20-minute samples (except for Canadian buoys which were 40 minutes) once per hour. The wind measurements were taken as ten-minute samples, scalar averaged, except vector averaged at the Canadian buoys, also once per hour. The wind and wave values selected for comparison with the hindcast were three-hour mean values centered on each six-hour synoptic time with equal (1,1,1) weighting. All wind speeds were adjusted to 10-m neutral winds following the approach described in Cardone *et al.* (1990).

B. DATA FROM OCEAN WEATHER STATION (OWS) BRAVO

Data from OWS Bravo were obtained from the US National Climatic Data Center. A large number of vessels occupied OWS Bravo; however, they tended to be one of two classes, with anemometer heights of 24 m. All ship wind speeds were also adjusted to 10-m neutral winds using the technique described by Cardone *et al.* (1990).

C. SATELLITE DATA

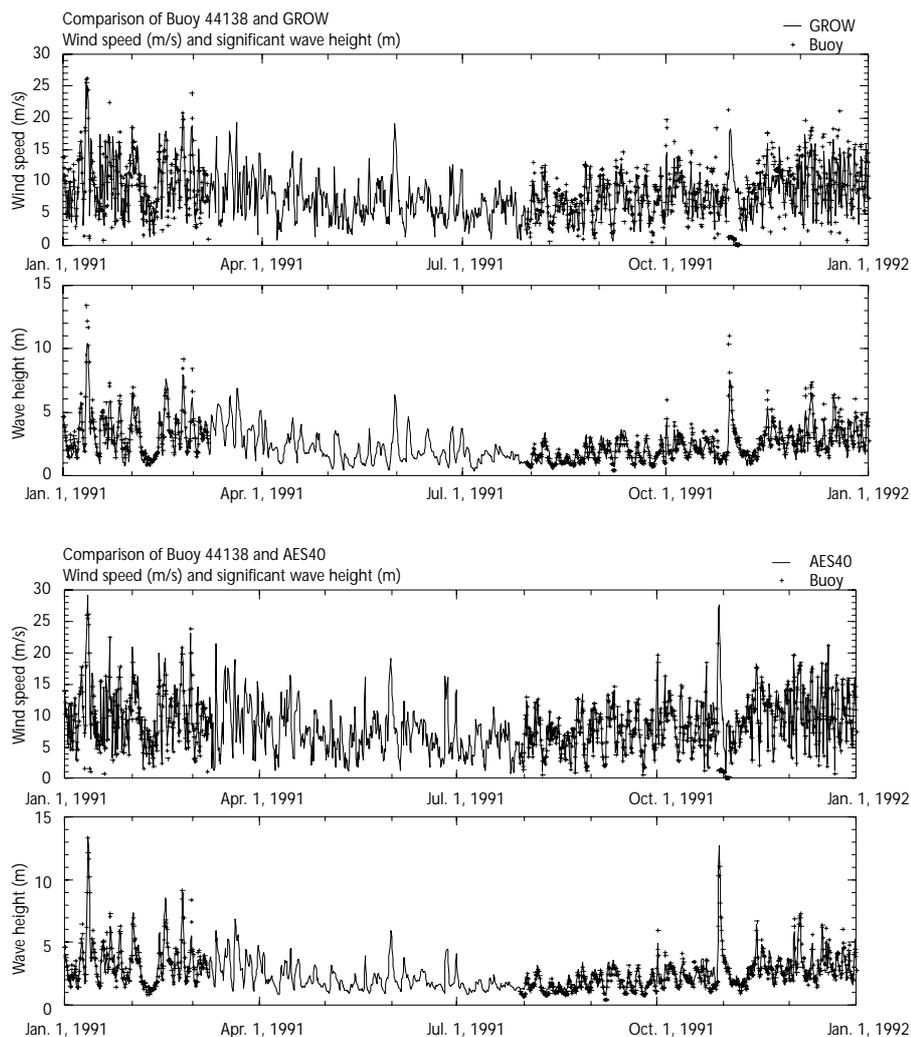
Altimeters from the ERS-1, ERS-2 and TOPEX/Poseidon instruments were used for global wind and wave comparisons. The ERS-1/2 altimeter data sets were obtained from the Ifremer CD-ROM data set, while TOPEX data (GDR Generation-B CD-ROM set) were obtained from the NASA Physical Oceanography Distributed Active Archive Center at the Jet Propulsion Laboratory/California Institute of Technology. Both data sets were decoded using the recommended quality controls described in the respective documentation. Further adjustments and quality control measures were used as recommended by Cotton and Carter (1994). Individual data points were then spatially binned onto the wave model grid, and output onto six-hour synoptic times using a ± 3 -hour window. Additional quality control was performed for measurements along land and ice edges where some contamination of the altimeter wave measurements was encountered despite rigorous checking of ice/quality control flags available with each data set.

4. IN SITU COMPARISONS

A. VALIDATION AGAINST BUOY AND PLATFORM MEASUREMENTS

Figure 3 shows a typical time series of wind speed and significant wave height for buoy 44138 for both GROW and AES40. The buoy time record is not continuous and has periods where wind and/or wave observations were not available. In general, both the GROW winds and waves track the buoy observations. The largest discrepancies occurred when strong extratropical systems passed close to a measurement site. The highest winds and waves in each type of event tended to be underpredicted; typically the lowest winds and waves tended to be somewhat overpredicted. The AES40 winds track very closely, as a result of the wind assimilation. The waves also track very well, and tend to better resolve the highest wave heights. This is partially a result of the local wind assimilation, but mainly due to the kinematic reanalysis of the storms that concentrate on following the major 'jet streaks' of wind maxima associated with storms.

Figure 3 — Comparison of buoy 44138 wind speed (m/s) and wave height (m) vs. GROW (top) and AES40 (bottom).



Individual buoys and platforms were then grouped by region (Figure 1) for comparison against GROW. Table 1 shows regional grouped statistics and represents more than 500,000 wind and wave observations. Highest scatter indices (SI, RMS/Mean Measurement) are from the north-west Pacific and north-east Atlantic regions, which were made up exclusively of COADS data. The COADS data lack both the time resolution (3/6 hours versus 1 hour) and coding accuracy (winds nearest 1 knot, waves 0.5 m) that pertain to the other regions obtained from the CD-ROM marine data sets, which may explain some of the differences in SI. The Canadian and US buoys were grouped into one data set since they represented the best science quality validation data set. These statistics show very good agreement with a mean bias of 0.12 m/s for winds and 0.10 m for waves and SI of 0.31 and 0.27, respectively.

Table 2 shows the same statistics for AES40, although a different number of buoys/platforms were selected for this comparison (Figure 2). The North/Norwegian Sea observations show higher SI in waves in comparison to AES40, the same finding as GROW. Wind speed scatter at the Canadian buoys is high, 0.31, mainly due to questionable data from one buoy which was left out of the wind assimilation but left in the comparisons shown here. Overall, AES40 has similar bias with lower SI and higher correlation coefficients when compared to GROW at the buoys/platforms.

While overall statistics are useful for evaluating the skill of a hindcast, they do not indicate how the hindcast has changed over time relative to the in situ data. A comparison of seasonal wave height bias and scatter over the 1975-1997 period (Figures 4 to 7) shows any trends that may exist in the hindcasts. Of course, trends may also occur in the measurements themselves (number of observations

Table 1—Regional statistical comparison of GROW vs. in situ buoy and platform observations.

	<i>Number of points</i>	<i>Mean meas.</i>	<i>Mean hind.</i>	<i>Diff. (H-M)</i>	<i>RMS error</i>	<i>Std. dev.</i>	<i>Scatter index</i>	<i>Corr. coeff.</i>
<i>North-east Atlantic</i>								
Ws (m/s)	30026	8.40	8.73	0.33	2.73	2.71	0.32	0.80
Wd (°)	30032	243.06	238.06	-4.81	N/A	29.78	0.08	N/A
Hs (m)	24530	2.58	2.84	0.26	1.29	1.27	0.49	0.76
<i>North-west Atlantic</i>								
Ws (m/s)	179938	7.14	7.54	0.40	2.57	2.54	0.36	0.78
Wd (°)	179940	248.55	270.12	4.40	N/A	36.00	0.10	N/A
Hs (m)	175256	1.98	2.04	0.06	0.57	0.56	0.28	0.89
<i>Gulf of Mexico/Caribbean</i>								
Ws (m/s)	59104	6.20	6.47	0.27	2.02	2.01	0.32	0.76
Wd (°)	59104	101.09	90.47	-5.78	N/A	31.87	0.09	N/A
Hs (m)	55642	1.17	1.49	0.33	0.49	0.36	0.31	0.88
<i>South Pacific</i>								
Ws (m/s)	12727	6.48	6.77	0.29	1.42	1.39	0.21	0.77
Wd (°)	12727	122.71	125.27	2.60	N/A	19.21	0.05	N/A
Hs (m)	12607	2.14	1.82	-0.32	0.48	0.36	0.17	0.77
<i>North-east Pacific</i>								
Ws (m/s)	121323	7.99	8.04	0.05	2.26	2.26	0.28	0.82
Wd (°)	121323	252.01	250.03	1.40	N/A	32.32	0.09	N/A
Hs (m)	121793	2.75	3.01	0.26	0.67	0.62	0.23	0.92
<i>North-west Pacific</i>								
Ws (m/s)	37893	7.44	6.72	-0.71	2.79	2.70	0.36	0.73
Wd (°)	37896	357.96	4.58	-3.40	N/A	43.07	0.12	N/A
Hs (m)	29555	1.40	1.88	0.48	0.97	0.85	0.60	0.67
<i>Hawaii</i>								
Ws (m/s)	70304	7.17	6.53	-0.64	1.85	1.74	0.24	0.74
Wd (°)	70304	73.68	75.62	1.12	N/A	23.01	0.06	N/A
Hs (m)	69289	2.38	2.10	-0.29	0.50	0.42	0.17	0.82
<i>Bering Sea</i>								
Ws (m/s)	19600	8.60	8.79	0.19	2.49	2.49	0.29	0.81
Wd (°)	19601	34.99	42.84	-1.44	N/A	33.61	0.09	N/A
Hs (m)	16271	2.68	3.08	0.40	0.75	0.64	0.24	0.93
<i>US and Canadian data combined</i>								
Ws (m/s)	466252	7.30	7.42	0.12	2.30	2.30	0.31	0.79
Wd (°)	466258	107.88	94.02	1.41	N/A	32.40	0.09	N/A
Hs (m)	453750	2.18	2.28	0.10	0.59	0.58	0.27	0.90

available, differing instrumentation, etc.), and the measured data must be evaluated carefully. These plots were produced by computing bias and SI for each region for every three months and plotting the resulting time series. All four figures show good agreement between the buoy observations and GROW/AES40 over time. The plots show nearly linear bias and SI over time indicating that both GROW and AES40 have remained consistent over the 22 years for which the buoy measurements are available. Highest SI occur in the data from COADS, while the US and Canadian comparisons are more consistent. Early US buoy comparisons show more bias and slightly higher SI, which may be due to the relatively few experimental buoys available in the late 1970s/early 1980s.

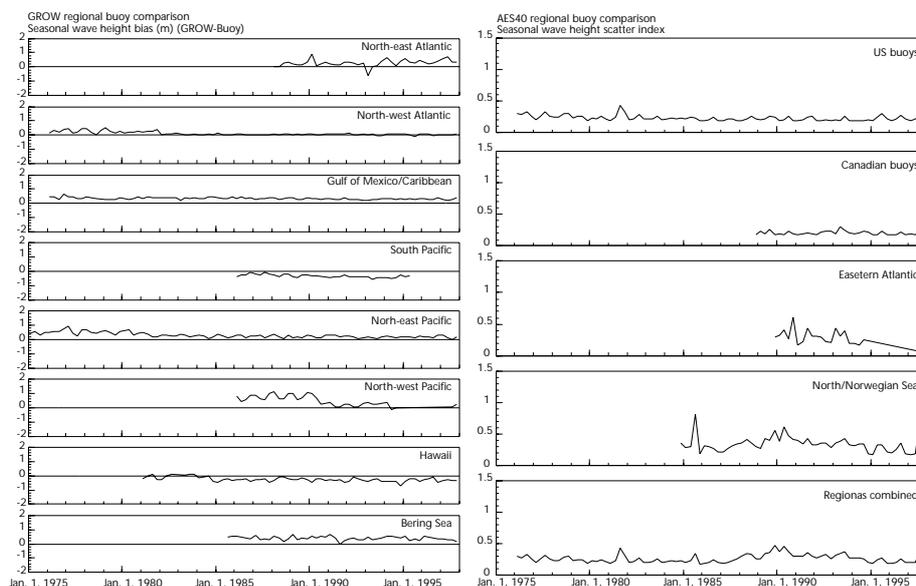
B. VALIDATION AGAINST OCEAN WEATHER STATION (OWS) BRAVO

OWS Bravo, located in the North Atlantic, gives an opportunity to evaluate the hindcasts well away from the coast and for the time period 1958-1974 where the buoy observations are not available.

Table 2—Regional statistical comparison of AES40 vs. in situ buoy and platform observations.

	Number of points	Mean meas.	Mean hind.	Diff. (H-M)	RMS error	Std. dev.	Scatter index	Corr. coeff.
US buoys								
Ws (m/s)	169927	6.92	7.18	0.26	1.31	1.28	0.19	0.94
Wd (°)	169925	240.47	251.65	0.99	N/A	16.65	0.05	N/A
Hs (m)	164834	1.83	1.94	0.12	0.43	0.42	0.23	0.93
Canadian buoys								
Ws (m/s)	49272	7.94	8.41	0.46	2.54	2.50	0.31	0.84
Wd (°)	49272	263.46	268.87	1.58	N/A	29.48	0.08	N/A
Hs (m)	48890	2.51	2.53	0.03	0.53	0.53	0.21	0.93
East Atlantic buoys								
Ws (m/s)	11019	9.75	9.71	-0.04	1.64	1.64	0.17	0.93
Wd (°)	11027	245.40	244.27	-0.44	N/A	17.98	0.05	N/A
Hs (m)	8071	3.73	3.47	-0.27	1.68	1.65	0.44	0.74
North/Norwegian Sea platforms and buoys								
Ws (m/s)	117198	8.58	9.14	0.56	2.24	2.17	0.25	0.88
Wd (°)	117204	240.17	239.27	-1.09	N/A	22.64	0.06	N/A
Hs (m)	107301	2.47	2.67	0.20	0.96	0.94	0.38	0.83
US and Canadian data combined								
Ws (m/s)	219199	7.15	7.45	0.31	1.67	1.64	0.23	0.91
Wd (°)	219197	247.72	257.67	1.11	N/A	20.14	0.06	N/A
Hs (m)	213724	1.98	2.08	0.10	0.46	0.45	0.23	0.93

Figure 4/5 — Seasonal wave height bias (m) (left) and SI (right) comparison of GROW vs. buoys by region.



Time-series comparisons (not shown) show similar characteristics to the buoy time-series figures. Storms tend to be underpredicted in GROW and better resolved in AES40. Figure 8 shows the seasonal bias and SI for Bravo wave heights. Both bias and SI comparisons show less bias/SI in the 1960s than the 1970s which results in an apparent trend. Whether this is due to changing measurement instrumentation/platform or a trend in the hindcasts is not known, although the ‘step-up’ nature of the comparison around 1968-69 suggests changes in the Bravo measurements.

5. SATELLITE COMPARISONS

Altimeter wind and wave measurements provide the best spatial coverage to evaluate wave hindcasts. Statistics and plots from the individual instruments (ERS-1, ERS-2 and TOPEX) showed very good agreement between each other, so the data

Figure 6/7 — Seasonal wave height bias (m) (left) and SI (right) comparison of AES40 vs. buoys by region.

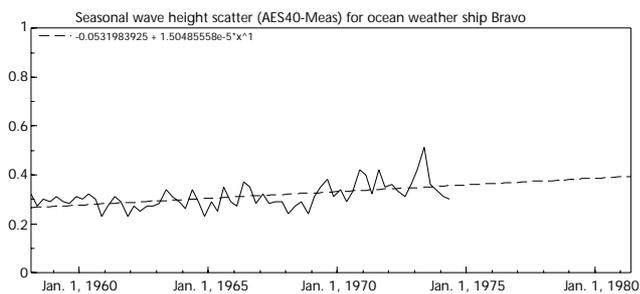
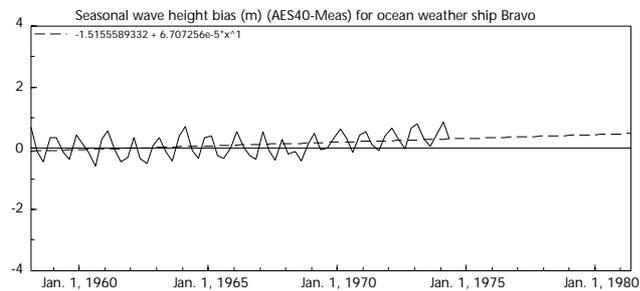
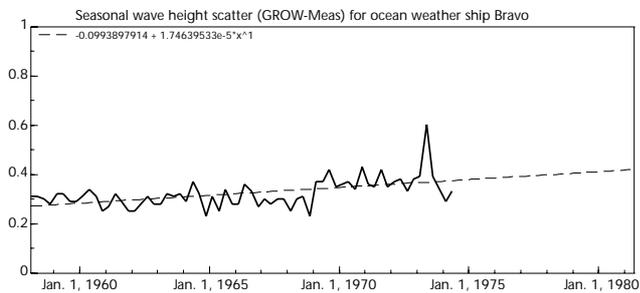
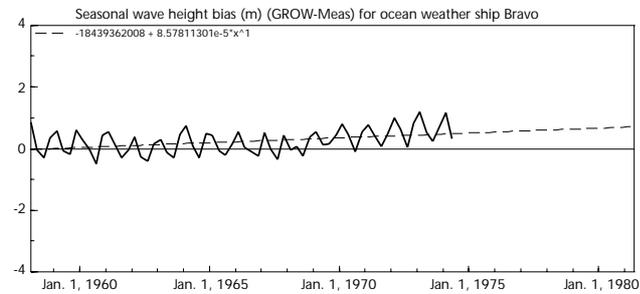
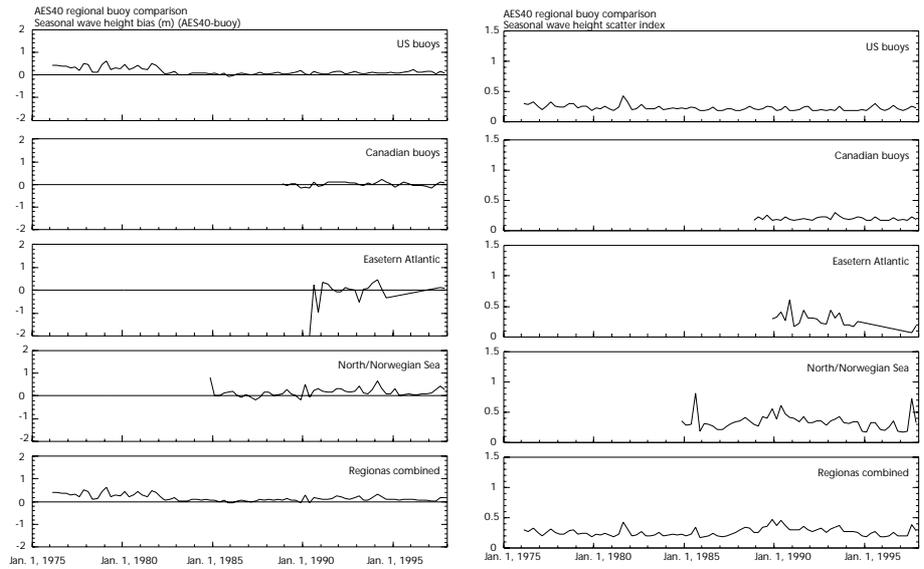


Figure 8 — Seasonal wave height bias (m) and SI for OWS Bravo vs. GROW (top) and AES40 (bottom).

sets were combined for these comparisons. The GROW model comparison was broken up into four regions: southern hemisphere (SH) (65S to 20S), tropical (TROP) (20S to 20N), northern hemisphere (NH) (20N to 70N), and all regions combined (65S to 70N). The AES40 comparisons were done for the full basin only. Statistics are summarized in Table 3.

Table 3—Regional statistical comparison of GROW and AES40 vs. altimeter measurements.

	Number of points	Mean meas.	Mean hind.	Diff. (H-M)	RMS error	Std. dev.	Scatter index	Corr. coeff.
<i>GROW: Southern hemisphere (65S to 20S)</i>								
Ws (m/s)	4004211	8.68	8.62	-0.06	2.40	2.40	0.28	0.79
Hs (m)	4001377	3.39	3.34	-0.05	0.79	0.79	0.23	0.85
<i>GROW: Tropics (20S to 20N)</i>								
Ws (m/s)	2608601	6.02	5.99	-0.03	1.86	1.86	0.31	0.71
Hs (m)	2593660	1.96	1.87	-0.08	0.46	0.45	0.23	0.77
<i>GROW: Northern hemisphere (20N to 70N)</i>								
Ws (m/s)	2086601	7.43	7.60	0.18	2.09	2.08	0.28	0.84
Hs (m)	2067467	2.54	2.56	0.02	0.65	0.65	0.26	0.91
<i>GROW: Global comparison</i>								
Ws (m/s)	8699413	7.60	7.60	0.00	2.18	2.18	0.29	0.81
Hs (m)	8662504	2.73	2.73	-0.04	0.68	0.67	0.24	0.89
<i>AES40: North Atlantic comparison</i>								
Ws (m/s)	3471109	7.66	7.81	0.15	1.94	1.94	0.25	0.86
Hs (m)	3523575	2.52	2.51	-0.01	0.56	0.56	0.22	0.93

Quantile-quantile (Q-Q) plots of the combined altimeter versus GROW (Figure 9) show excellent agreement for both wind speed and wave height. At the highest percentiles, winds appear to be overpredicted, while waves are underpredicted. This is suspected to be a wind speed saturation problem with the altimeter in wind speeds above 15 m/s. The wave underestimation appears to be a property of the GROW wave hindcast. A Q-Q comparison of AES40 (Figure 10) shows the same overestimation of wind speed, but tracks the waves up to the 99th percentile. This is a result of AES40's intensive reanalysis of the strongest storms.

The global coverage of the altimeter measurements makes it possible to plot contours of wave bias on a global projection. Figure 11 shows the global wave height bias which indicates spatially coherent regions of GROW overestimating and underestimating the measured waves. Many of the regions, such as the Caribbean Sea, the Aleutian Island Chain, and North Sea, are suspected to be resolution effects of the GROW wave model since the grid spacing is too coarse to resolve the coastline. The large region of bias off Antarctica is suspected to be the effects of using mean-monthly ice tables for the entire hindcast. There appears to be a large area of underestimation of wave height in the southern hemisphere along 30S with the strongest bias in the south-east Pacific.

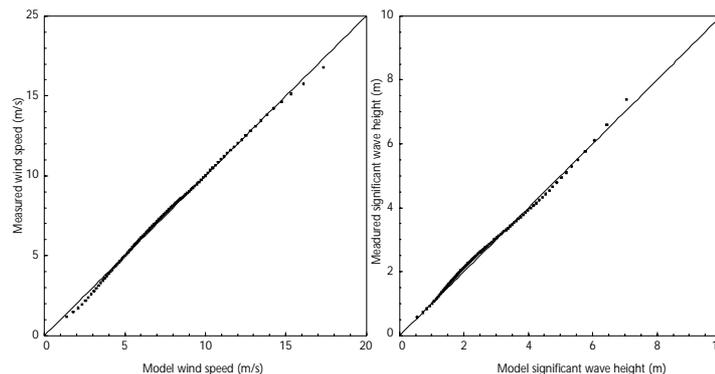


Figure 9—Q-Q wind speed (m/s) and wave height (m) comparisons of GROW and altimeter measurements.

Figure 10—Q-Q wind speed (m/s) and wave height (m) comparisons of AES40 and altimeter measurements.

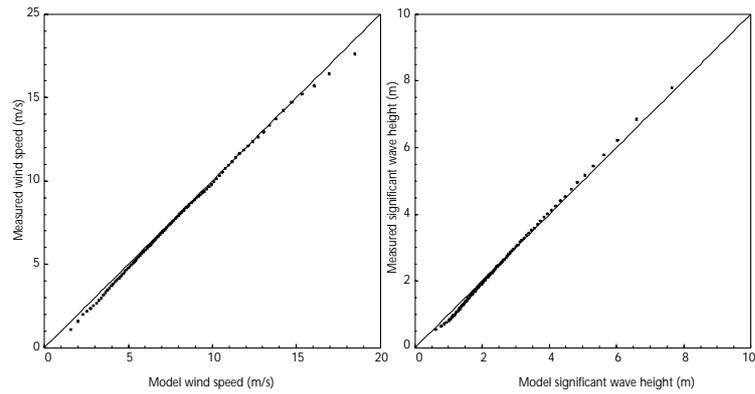
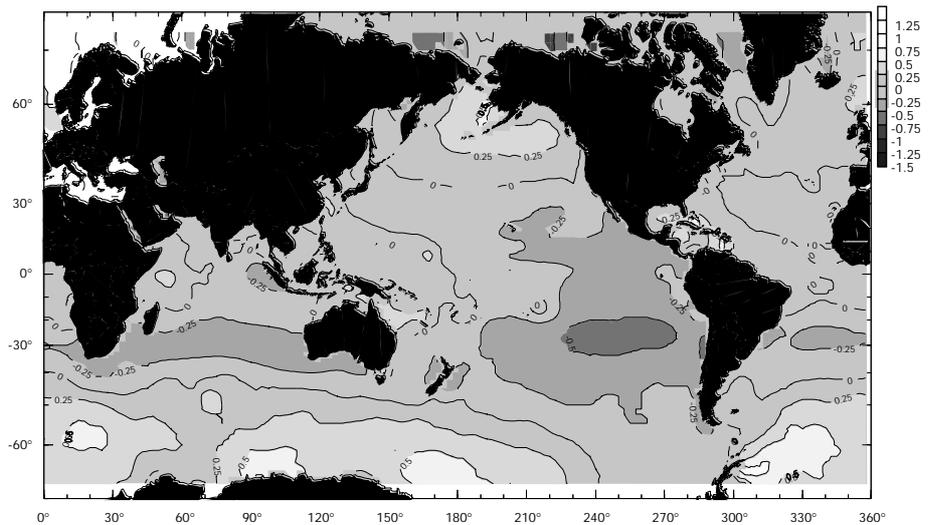


Figure 11—Mean difference of wave height (m) between GROW and altimeter measurements (GROW-Altimeter).



A spatial wave bias plot of AES40 (Figure 12) shows that over most of the North Atlantic, AES40 has very little bias. The largest feature is the underestimation in the Baffin Sea and in the Denmark Strait. This is suspected to be a result of ice edge effects, and, to some degree, an underestimation of the wind speed in the NRA winds. While the AES40 winds were kinematically enhanced, the lack of data in these areas made it difficult to track all significant systems. When sufficient data were available, large discrepancies of the wind speed were found and corrected in the NRA winds. Grid scale effects explain most other areas of bias near island chains or in the shallow southern North Sea.

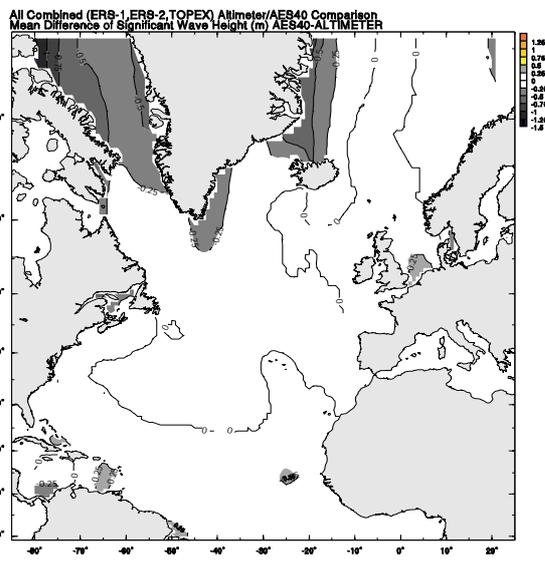


Figure 12—Mean difference of wave height (m) between AES40 and altimeter measurements (AES40-Altimeter).

6.
CONCLUSIONS

This paper has presented the use of in situ and satellite observations to evaluate long-term wave hindcasts. Both the GROW and AES40 validation show that each hindcast compares well against the available buoy, platform, OWS and satellite measurements. Comparisons of in situ data over the full 1958-1997 period show that both hindcasts have remained consistent with the observations.

In the top percentiles, GROW shows a tendency to underpredict the highest sea states, while AES40 better resolves the peak storms. Spatial comparison of AES40 shows very little bias across most of the North Atlantic, while GROW appears to show some coherent areas of under- and overestimation that cannot be explained by grid/ice edge effects.

In summary, it has been shown that in situ and satellite data serve powerful and complementary roles in the evaluation of global and basin scale long-term hindcasts. However, we caution that due consideration must be given to the limitations of each measurement data set before biases and trends that appear in the comparison statistics are attributed to either nature or model error.

REFERENCES

- Athanassoulis, G.A. and Ch. N. Stefanakos, 1995: A nonstationary stochastic model for long-term time series of significant wave height. *J. of Geophys. Res. (Oceans)*, Vol 100, No. C8, pp. 16,149-16,162.
- Bacon, S. and D.J.T. Carter, 1991: Wave climate changes in the North Atlantic and North Sea. *Int. Journal of Climate*, Vol. 11, 545-558.
- Cardone, V.J., J.G. Greenwood and M.A. Cane, 1990: On trends in historical marine wind data. *J. of Climate*, 3, 113-127.
- Cotton, P.D. and D.J.T. Carter, 1994: Cross calibration of TOPEX, ERS-1, and Geosat wave heights. *J. of Geophys. Res.*, Vol 99, No. C12, pp. 25,025-25,033.
- Cox, A.T., V.J. Cardone and V.R. Swail, 1997: Evaluation of NCEP/NCAR reanalysis project marine surface wind products for a long term North Atlantic wave hindcast. *Proc. 1st WCRP International Conference on Reanalyses* (Silver Spring, MD, 27-31 October 1997). WCRP-104, WMO/TD-NO. 876, p. 73-76.
- Cox, A.T., V.J. Cardone and V.R. Swail, 1998: Evaluation of NCEP/NCAR reanalysis project marine surface wind products for a long-term North Atlantic wave hindcast. *Proc. 5th International Workshop on Wave Hindcasting and Forecasting* (Melbourne, FL).
- Cox, A.T. and V.R. Swail, 1999: A global wave hindcast over the period 1958-1997: validation and climate assessment. Submitted to *JGR (Oceans)*.
- Kalnay, E. *et al.*, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. AMS*, 77(3), 437-471.
- Khandekar, M.L., R. Lalbeharry and V.J. Cardone, 1994: The performance of the Canadian Spectral Ocean Wave Model (CSOWM) during the Grand Banks ERS-1 SAR wave spectra validation experiment. *Atmosphere-Ocean* 31 (1), pp. 31-60.
- Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate*, 7, 141-157.
- Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set Release 1*. NOAA Environmental Research Laboratories, Boulder, Colo., 268 pp. (NTIS PB86-105723).
- Swail, V.R. and A.T. Cox, 1999: On the use of NCEP/NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. Accepted in *J. Atmos. Ocean. Technology*.
- WASA Group, 1998: Changing waves and storms in the Northeast Atlantic? *AMS Bul.*, Vol. 79, No. 5, May 1998 pp. 741-760.

SCATTEROMETRY DATA SETS: HIGH QUALITY WINDS OVER WATER

Mark A. Bourassa*, David M. Legler and James J. O'Brien, Center for Ocean-Atmospheric Prediction Studies (COAPS), Florida State University, USA

1. INTRODUCTION

In the late 1990s microwave scatterometry is finally catching up to other radiometric instruments of the SeaSat era: altimeters (ocean height and wave height), radiometers (temperatures and humidity), and scatterometers (wind speed and direction), all designed to provide the previously unattainable quantity and quality of data regarding variability of the ocean and adjoining atmospheric boundary-layer (Katsaros and Brown, 1991). Europeans have been working with two successive scatterometers beginning in 1991. These scatterometers on the European Remote Sensing Satellite Systems (ERS-1 and ERS-2) provided the first scatterometer data that could be used for climatological studies. Operational constraints have prevented continuous scatterometer observations over water; however, the scatterometer was usually operating away from land and ice.

The Japanese satellite, ADEOS, which was launched in August 1996, had the first dedicated microwave scatterometer since SeaSat: the NASA Scatterometer (NSCAT). This scatterometer determined wind speed and direction over 90 per cent of the ice-free global water surface in two days with 25-km in-swath resolution. It functioned until a catastrophic failure of the satellite platform on 29 June 1997. Despite this loss, the unprecedented coverage and resolution of global wind data gave light to profound impacts on oceanographic and meteorological applications.

The unprecedented accuracy and coverage of NSCAT winds led to the rapid deployment of a new type of scatterometer (SeaWinds) to fill the void in NSCAT-like observations. SeaWinds instruments are on QuikSCAT (launched on 19 June, 1999) and ADEOS-2 (planned). SeaWinds scatterometers have approximately double the coverage of NSCAT, covering 90 per cent of the world's oceans in one day. The NSCAT and SeaWinds periods may be the only times to date when ocean modellers could not reasonably argue that errors in model output were due mainly to shortcomings in wind observations. Owing to the relatively recent development of these wind products, few researchers are aware of their nature and quality. This report describes the wealth of current products, as well as providing a brief discussion of their strengths and weaknesses.

Several types of data sets, appropriate to different applications, are publicly available. The swath winds (i.e. gridded relative to the satellite track) are available for those who need near-instantaneous winds that are not further processed (e.g. Jones *et al.*, 1999). For example, comparisons of these winds to research vessels (R/Vs) (Bourassa *et al.*, 1997) and National Data Buoy Center (NDBC) buoys (Freilich and Dunbar, 1999) have shown that these winds could be used to quality control ship and buoy observations. However, these data sets are not regularly gridded in a latitude-longitude grid, and have gaps in daily coverage. Most ocean modelling applications require winds (or stresses) to be regularly gridded in space and time, with no missing data over water. Many such gridded daily products are also available. Such products are also useful in constructing wind vector climatologies that include synoptic-scale and some mesoscale variations (e.g. Bourassa *et al.*, 1999b; Chelton *et al.*, 2000; Milliff *et al.*, 1999a). Animations of gridded products (Bourassa *et al.*, 1998, 1999b) have also been developed for data visualization. Animations are of great use in examining the vast quantity of scatterometer data to find features of

*Corresponding author's address: Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida 32306-2840, USA; e-mail: bourassa@coaps.fsu.edu; Telephone: (850) 644-6923, FAX: (850) 644-4841

interest. They clearly show frontogenesis, cyclogenesis, and larger scale circulation patterns. These scatterometer data products will be discussed.

2. SCATTEROMETRY BACKGROUND

Scatterometers are unique among satellite remote sensors because of their ability to accurately determine wind speed and direction. Microwaves are Bragg-scattered by short water waves, which respond quickly to changes in winds. This backscatter (the fraction of transmitted energy that returns to the satellite) is a function of wind speed and wind direction. The wind direction is found by determining the angle that is most likely to match the observed backscatters. A digital filtering technique (Naderi *et al.*, 1991) is used to sample locations from multiple angles in less than five minutes. There are substantial design differences for ERS scatterometers, NSCAT, and SeaWinds (Table 1). For example, the ERS backscatter is spatially smoothed, thereby reducing the resolution to ~70 km (M.H. Freilich and D.G. Long, 1998, personal communication). On NSCAT, there were three fixed antennas on each side, allowing swaths on each side of the satellite track to be sampled by fore, mid, and aft beams. Wind speeds and directions were calculated when radar observations were available from all three of these antennas. In contrast, ERS scatterometers are sometimes forced to use observations from only two antennas (Zeccheto *et al.*, 1999). For fixed-antenna scatterometers (SeaSat, ERS-1/2 and NSCAT), the use of three or more antennas is essential for accurate determination of the wind direction (Naderi *et al.*, 1991). The beam arrangement on SeaWinds instruments is a new design, with two conically rotating beams at fixed incidence angles. This design allows a single, very wide observational swath. This scanning geometry has four substantially different angles over an area similar to the NSCAT swaths. However, near nadir and near the edges of the swath, the angles are similar, resulting in decreased accuracy in these parts of the swath. Furthermore, only one of the two beams reaches the outer 75 km of the swath. These problems are somewhat compensated by a much greater observation density. NSCAT had three or four observations within a 25 × 25 km cell, whereas SeaWinds typically has between eight and 25 observations within its 25 × 25 km cells.

The functions describing the wind direction are sinusoidal. Combining these functions to minimize the misfit usually results in multiple minima (ambiguous solutions often called ambiguities). Ideally, for fixed-antenna scatterometers, the best fit corresponds to the correct direction, the next best fit is in approximately the opposite direction, and the next two minima are in directions roughly perpendicular to the wind direction. For SeaWinds scatterometers, the solution geometry varies across the swath. The solutions are similar to fixed-antenna scatterometer solutions in the part of the swath similar to NSCAT coverage, but differ greatly near nadir and

Table 1—Scatterometer characteristics. Note that for ERS-1/2 scatterometers, the three cells closest to nadir do not meet all the desired retrieval requirements; wind vectors from these cells are often ignored.

Scatterometer	Period in service	Scan characteristics	Swath width (km)	Nadir gap (km)	In-swath grid Spacing (km)	Cell size (km)	Scan characteristics	Operational frequency
ERS-1 Scatterometer	1991/7 to 1997/5/21	one-sided (single swath)	475	NA	25 × 25	50 × 50	one-sided (single swath)	C band (5.3 GHz)
ERS-2 Scatterometer	1997/5/21 to current	one-sided (single swath)	475	NA	25 × 25	50 × 50	one-sided (single swath)	C band (5.3 GHz)
NSCAT	1996/9/15 to 1997/6/30	two-sided (double swath)	600	329	25 × 25 50 × 50	25 × 25 50 × 50	two-sided (double swath)	Ku band (13.995 GHz)
SeaWinds on QuikSCAT	~1999/7/19 to current	conical scan, one-wide swath	1900	NA	25 × 25	25 × 25	conical-scan, one wide swath	Ku band (13.995 GHz)
SeaWinds on ADEOS II	TBA	conical scan, one-wide swath	1900	NA	25 × 25	25 × 25	conical-scan, one wide swath	Ku band (13.995 GHz)
ASCAT	TBA	two-sided (double swath)	550	660	25 × 25	50 × 50	two-sided (double swath)	C-band (-5.3 GHz)

near the swath edges. The process of choosing the direction is called ambiguity selection. Noise and spatial/temporal variability can change the quality of fit and thereby cause incorrect directions (also known as aliases) to be chosen. NSCAT's ambiguity removal was further improved by using two polarizations with one antenna, whereas SeaWinds ambiguity selection is improved by greater observation density. For NSCAT and QuikSCAT winds, a median filter (applied to ambiguity selection rather than wind direction) is also used to improve ambiguity selection.

Rain influences radar returns through three processes: backscatter from rain drops, attenuation of the signal passing through the rain (Moore *et al.*, 1999), and modification of the sea surface shape by raindrop impacts (Bliven *et al.*, 1993; Sobieski and Bliven, 1995; Sobieski *et al.*, 1999). The influence of these considerations on the accuracy of winds is a function of scatterometer design. Rain has a greater influence at large incidence angles (the signal passes through more rain) and for Ku-band (NSCAT and SeaWinds) rather than C-band (ERS-1/2). Rain is not considered a serious problem for the ERS scatterometers. For NSCAT, rain contributed to substantial errors in the outer parts of the swaths; however, rain can have a substantial influence on SeaWinds observations throughout the swath. Modelling these problems is a concern of ongoing research (Weissman *et al.*, 2000). In the meantime, several rain flags are being developed. On ADEOS-2, rain-related contamination will be identified through co-located rain observations from sensors aboard the satellite.

3. SWATH DATA

Scatterometers are carried onboard polar orbiting satellites. QuikSCAT and NSCAT have been in sun synchronous orbits, with approximately 15 orbits per day. Polar orbits, in contrast to the geostationary orbits of more routinely used weather satellites, have areas of coverage that change with time. One great advantage of using polar orbits is obtaining observations at latitudes much farther from the equator than can be achieved with geostationary orbits. Polar-orbiting satellites take observations in swaths below and/or to the sides of the satellite (usually described relative to its forward motion). The ERS-1/2 scatterometer observations come from a 500 km wide swath on only one side of the satellite. NSCAT more than doubled this coverage by measuring the return signals from 600 km wide swaths on both sides of the satellite, with a 400 km wide nadir gap below the satellite (Figure 1). The spatial coverage of SeaWinds scatterometers is doubled again by filling the nadir gap and extending the far edges of the swath another 75 km. The observation rate is staggering, with the number of daily observations provided by SeaWinds approximately equal to the number of annual wind observations routinely provided by all buoys and ships available through the Global Telecommunication System (GTS) data stream.

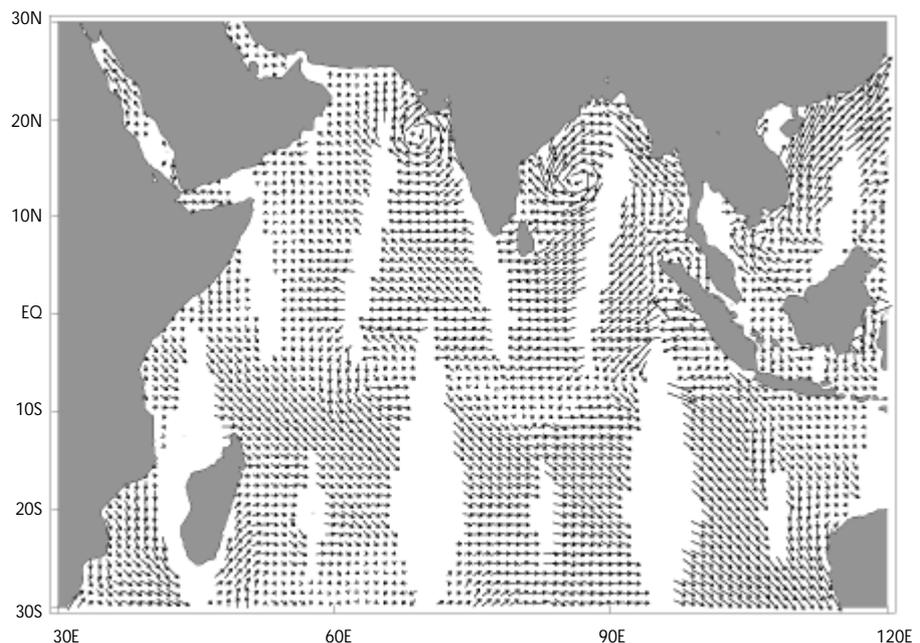


Figure 1—NSCAT observations from 26 October 1996, vector averaged in $1 \times 1^\circ$ bins. The single day of observations captures cyclones in the Arabian Sea and the Bay of Bengal.

Winds determined with the NSCAT-1 geophysical model function (Wentz and Smith, 1999) have been validated against a wide range of in situ and remotely-sensed winds. NSCAT and SeaWinds wind speeds have been calibrated to 10 m 'equivalent neutral wind speeds' (Liu and Tang, 1996; Verschell *et al.*, 1999), which differ from wind speed in a manner believed to be consistent with the physics to which the scatterometer responds. The differences can easily be explained with the equation for the modified log-wind profile:

$$U(z) - U_{sfc} = (u_* / k) [\ln (z/z_o) - \phi(z, z_o, L)] \quad (1)$$

where U is the vector wind, U_{sfc} is the velocity frame of reference (the surface current), u_* is the friction velocity, k is von Karman's constant, z is the height above the local mean surface (10 m in this case), z_o is the roughness length, ϕ is a function of atmospheric stability, and measure of atmospheric stability is the Monin-Obukhov scale length (L). Scatterometers respond to the sea surface ($z=0$), and the stability term (ϕ) is largely a function of z/L . Therefore, the concept is to eliminate the stability term in the height adjustment. Equivalent neutral wind speed (Cardone *et al.*, 1969; Ross *et al.*, 1985; Cardone *et al.*, 1996; Liu and Tang, 1996, Verschell *et al.*, 1999) is parametrized similarly to (1) and uses the same non-neutral values of u_* and z_o ; however, the stability term (ϕ) is set to zero. Hence, the differences between U_{10EN} and U_{10} are stability dependent.

$$U_{10EN} - U_{10} = u_* \phi (z, z_o, L) / k \quad (2)$$

It considers that the scatterometer probably responds to stress rather than wind speed. The kinematic stress is equal to the square of the friction velocity; therefore, the friction velocity used in the calibration of scatterometer winds should be the non-neutral value rather than the neutral value. Relatively large values of $|U_{10EN} - U_{10}|$ tend to be associated with very stable stratification. Values of $|U_{10EN} - U_{10}|$ are usually $<0.5 \text{ m s}^{-1}$ (hereafter U_{10EN} will be referred to as winds).

Studies comparing scatterometer winds to in situ observations have been made with buoys, Voluntary Observing Ships (VOSs), and R/Vs (Table 4). These studies investigate the accuracy of wind speed, wind direction (usually for correctly selected ambiguities) and vector winds (Freilich and Dunbar, 1999), as well as the fraction of correct ambiguity selection. These studies usually determined the rms difference between scatterometer and in situ winds, which provides an upper limit on uncertainty in scatterometer winds (since a substantial fraction of the differences is probably due to uncertainty in the comparison data set). The findings are summarized in Table 4.

Underestimation of NSCAT-1 model function wind speeds for $U_{10} > 20 \text{ ms}^{-1}$ (R. Brown and R. Foster, 1997, personal communication, Jones *et al.*, 1999), as well as biases in selection of ambiguities (Ebuchi, 1999), led to the development of the NSCAT-2 and NSCAT-2p geophysical model functions. For most applications, the differences in NSCAT-1 and NSCAT-2 winds are very small; however, for high wind speed problems these changes are systematic and could be significant. The differences between the NSCAT-2 and NSCAT-2p model functions are also small (F. Wentz and M. H. Freilich, 2000, personal communications); however, they led to improved impact in the weather model forecasts (R. Atlas, 2000, personal communication). The NSCAT-2p model function is consistent with the JPL model function (QSCAT-1) currently used for the SeaWinds scatterometer. An alternative SeaWinds model function (Ku-2000) has been developed by Remote Sensing Systems (F. Wentz and D. Smith, 2000, personal communication). The following is the first published evaluation of the NSCAT-2, QSCAT-1, and Ku-2000 model functions.

3.1 VALIDATION OF NSCAT-2 AND QSCAT-1 WINDS

The differences between scatterometer winds and the comparison data sets were often expressed in terms of rms differences (for correctly selected ambiguities) due to programmatic requirements on accuracy. When there is no uncertainty in the comparison data set, no biases in either data set (or equal biases), and no complications due to geophysical inconsistencies (e.g. space and time scales or inexact

SECTION 4 — DEVELOPMENT AND USE OF SATELLITE MARINE DATABASES

Scatterometer	Comparison data set(s)		Comparison statistics		Qualifiers	Model function or product	Reference
			(ms^{-1})	(deg.)			
ERS-1/2	Buoys	NDBC	2.0	41° at 0-50 ms^{-1}	Speeds: r, SA Directions: s, SA	JPL CMOD4-FD	Graber <i>et al.</i> (1996)
		ODBS	1.9	23° at 5-50 ms^{-1}			
		TAO	1.8	21° at 5-50 ms^{-1} * * Excludes inner 3 cells			
		NDBC	1.6	55° at 0-50 ms^{-1}	Speeds: r, SA Directions: s, SA	CMOD4	Graber <i>et al.</i> (1996)
		ODBS	1.8	40° at 5-50 ms^{-1}			
		TAO	1.8	37° at 5-50 ms^{-1} * * Excludes inner 3 cells			
		NDBC	1.3	25° at 3-50 ms^{-1}	Speeds: r, SA Directions: s, SA	IFREMER	Graber <i>et al.</i> (1996)
		ODBS	1.3	22° at 5-50 ms^{-1}			
		TAO	1.3	21° at 5-50 ms^{-1} * * Excludes inner 3 cells			
			NDBC	2.5	—	c, s, SA	Operational CMOD4
		NDBC	1.7	zonal component	c, r, SA	COMD4	Stoffelen (1998)
		NCEP	1.4	meridional comp.			
ERS-2	Ships	VOS	3.3	24	r, CA	Operational CMOD-4	Atlas <i>et al.</i> (1999)
NSCAT	Ships	Research	1.4	12 for 2-20 ms^{-1}	r, CSA	NSCAT-1	Bourassa <i>et al.</i> (1997)
		Research	<1.3	<10 for 2-20 ms^{-1}	u, CSA	NSCAT-2	This article
		VOS	2.7	21	r, CA	NSCAT-1	Atlas <i>et al.</i> (1999)
	Buoys	NDBC	1.3	30° at 3 ms^{-1} 17° at 5 ms^{-1} 14° at >10 ms^{-1}	S, CSA	NSCAT-1	Freilich and Dunbar (1999)
		NDBC	2.0	18.8	r, CA	NSCAT-1	Atlas <i>et al.</i> (1999)
		NDBC	0.6	—	c, s, SA	NSCAT-2	Freilich and Vanhoff (2000)
		TAO	1.14	33 20	r, SA r, CA	NSCAT-2	Dickenson <i>et al.</i> (2000)
		TAO	1.6 1.2 1.1 1.4 2.9	54° at 0-5 ms^{-1} 25° at 5-7.5 ms^{-1} 17° at 7.5-10 ms^{-1} 20° at 10-12.5 ms^{-1} 20° at 12.5-50 ms^{-1}	Speeds: r, SA Directions: s, SA	NSCAT-2 (25 km)	Caruso <i>et al.</i> (1999)
		WHOI	1.6 0.68 0.79 3.9	54° at 0-5 ms^{-1} 18° at 5-7.5 ms^{-1} 15° at 7.5-10 ms^{-1} 6° at 10-12.5 ms^{-1}	Speeds: r, SA Directions: s, SA	NSCAT-2 (25 km)	Caruso <i>et al.</i> (1999)
		Model	GEOS-1	2.8	22	r, CA	NSCAT-1
winds	NCEP	2.0	19	r, CA	NSCAT-1	Atlas <i>et al.</i> (1999)	
QuikSCAT	Ships	Research vessels	<0.45	<5° for 2-20 ms^{-1}	u, CSA	QSCAT-1	Bourassa <i>et al.</i> (2001)
		Research vessels	<0.3	<3° for 2-20 ms^{-1}	u, CSA	Ku-2000	Bourassa <i>et al.</i> (2001)

Table 4—Uncertainties in scatterometer observations. Many different assumptions (listed in the column labelled ‘Qualifiers’) have been used to determine these statistics: closest ambiguities (CA), correctly selected ambiguities (CSA), selected ambiguities (SA), vector wind component rather than speed (c), rms difference (r), standard deviation (s), uncertainty (u).

co-location), rms differences (and standard deviations) are essentially identical to traditional estimates of uncertainty. The scatterometer rms differences (and standard deviations) in Table 4 included contributions from the problems listed above, as well as geophysical differences due to in situ wind observations being earth relative, and the scatterometer winds being surface relative. Differences between random uncertainty and rms differences are highlighted in the comparison of ERS winds from various model functions to buoy winds (Table 4). The rms differences are substantially different for each model function; however, these differences are due more to biases than to differences in random uncertainty (Graber *et al.*, 1996). An accurate assessment of uncertainty, which is far more useful than an rms difference, requires that these additional factors be considered.

The uncertainty in the comparison data set is difficult to assess in this case since there is no absolute standard of truth for ocean winds. Techniques for estimating uncertainty in observations and comparison data sets have been developed (Stoffelen, 1998) using a third set of co-located observations. This approach uses the estimated uncertainties in the calculation of systematic gains and offsets. A similar approach, modified to consider a random component error (Freilich and Vanhoff, 2000), efficiently deals with the low wind speed problems identified by Freilich (1997). Unfortunately, these techniques require at least thousands of collocated observations from three sources. Such a large quantity of co-locations is not readily available from R/V data. An elegant alternative to these approaches is Principal Component Analysis (PCA; Pearson, 1901; Preisendorfer and Mobley, 1988), which assumes that the uncertainty in the comparison data set is equal to the uncertainty in the observations. This approach finds the variance perpendicular from a best-fit line. In the case of our comparisons between quality controlled ship observations and correctly selected ambiguities of NSCAT and SeaWinds observations, this assumption is good; otherwise a more complicated technique would be needed.

The uncertainty in the comparison data set (at the location of the satellite observation) is reduced by restricting this analysis to coincident satellite and R/V observations. For the calculation of rms differences, the central differences in observation times are less than twenty minutes (usually <30s), and the differences in locations were <25 for the NSCAT 25 km product, <50 km for the NSCAT 50 km product, and <12.5 km for SeaWinds. The co-location distance requires greater consideration, because the rms differences and estimated uncertainties are highly dependent on co-location distance. This point will be demonstrated in section 3.4 on QuikSCAT directional uncertainty.

The observations come from many ocean and atmospheric conditions (Tables 2 and 3); consequently, net biases in these findings due to location are unlikely, a specific sea state or atmospheric stability. There were 135 co-locations for the 25 km NSCAT product, 424 co-locations with SeaWinds QSCAT-1 product, and 425 co-locations with the SeaWinds Ku-2000 product. In all cases, wind speeds ranged from 2 to 20 m s⁻¹. A boundary-layer model (Bourassa *et al.*, 1999a) is used to adjust the R/V wind speeds to neutral equivalent winds at a height of 10 m, the height for which scatterometer winds are calibrated.

3.2 SHIP OBSERVATIONS

Wind directions from quality controlled R/V observations have proven to be the most consistently accurate source of in situ surface comparison (Table 4). True winds (i.e. speeds relative to the fixed earth and directions relative to true north) from ships that are correctly calculated (Smith *et al.*, 1999) do not suffer from either the directional shortcomings of typical buoys (in light winds or heavy seas) or the large uncertainties in VOS observations (Pierson, 1990). Preliminary

Table 2—Research vessels used in NSCAT validation.

Ship	Location	Time
RSV Aurora Australis	Southern Ocean	Sept., Nov. 1996
R/V Knorr	North Atlantic	Oct. 1996 to March 1997
R/V Thompson	North and tropical Pacific	July-Sept. 1996

Table 3—Vessels used in QuikSCAT validation.

Ship	Location	Time
R/V Atlantis	Gulf of Alaska	July, Aug. 1999
RSV Aurora Australis	Southern Ocean	July–Sept. 1999
R/V Knorr	North and Eq. Atlantic	Jan.–June 2000
R/V Melville	Tropical Pacific	July–Nov. 1999
R/V Meteor	North Atlantic	July 1999 to Aug. 2000
R/V Oceanus	North Atlantic	July–Dec. 1999, April 2000
R/V Polarstern	North Atlantic	July 1999 to June 2000

comparisons between winds from VOSs and NSCAT found that the rms differences between NSCAT and VOS wind speeds were roughly three times as large as the differences with our quality-controlled R/V winds (V. Zlotnicki and R. Atlas, 1997, personal communications). Another advantage of ship observations over buoy observations is that the observation height is above the regime where wave motions modify the log-wind profile (Large *et al.*, 1995), which is not the case for buoys in heavy seas. Nevertheless, for most open ocean conditions, there is little difference between the quality of R/V and buoy winds.

The major shortcoming of ship observations is the impact of flow distortion on wind vectors. Directional errors due to flow distortion are reduced by eliminating winds from ship-relative angles that passed through or near the superstructure. Nevertheless, flow distortion does cause wind speed biases. Observational and model-based studies (Yelland *et al.*, 1998; Thiebaux, 1990) applied to different ships indicate that biases due to flow distortion vary from ship to ship. Much of the wind observation record from the R/V Ronald Brown was discarded during our quality control of the ship data (prior to comparison with the scatterometer); most cruises during this time period suffered from severe flow distortion (Chris Fairall, 2000, personal communication). The bias in QuikSCAT, relative to the ships used in this study, ranges from -0.4 to $+0.7$ ms^{-1} , with most speed biases being within ± 0.2 ms^{-1} . One fascinating potential use of high quality scatterometer data is the estimation of biases due to flow distortion. In less than one year of open-ocean operations, there would be sufficient observations (an average of two per day for QuikSCAT) to examine the problem as a function of wind speed and ship-relative wind direction.

Another minor shortcoming of ship data is that one-minute observation intervals are insufficient to remove averaging errors associated with ship acceleration (i.e. changes in speed or direction; Smith *et al.*, 1999). In 1999, the processing of wind data on the R/V Polarstern was changed to calculate true winds every 5 s and average them every minute. The acceleration-related errors are not evident in the winds recorded by this system. Ship winds associated with excessive acceleration are filtered out through the restriction that magnitude of the sum of variance in the component velocities be less than 1.0 m^2 s^{-2} .

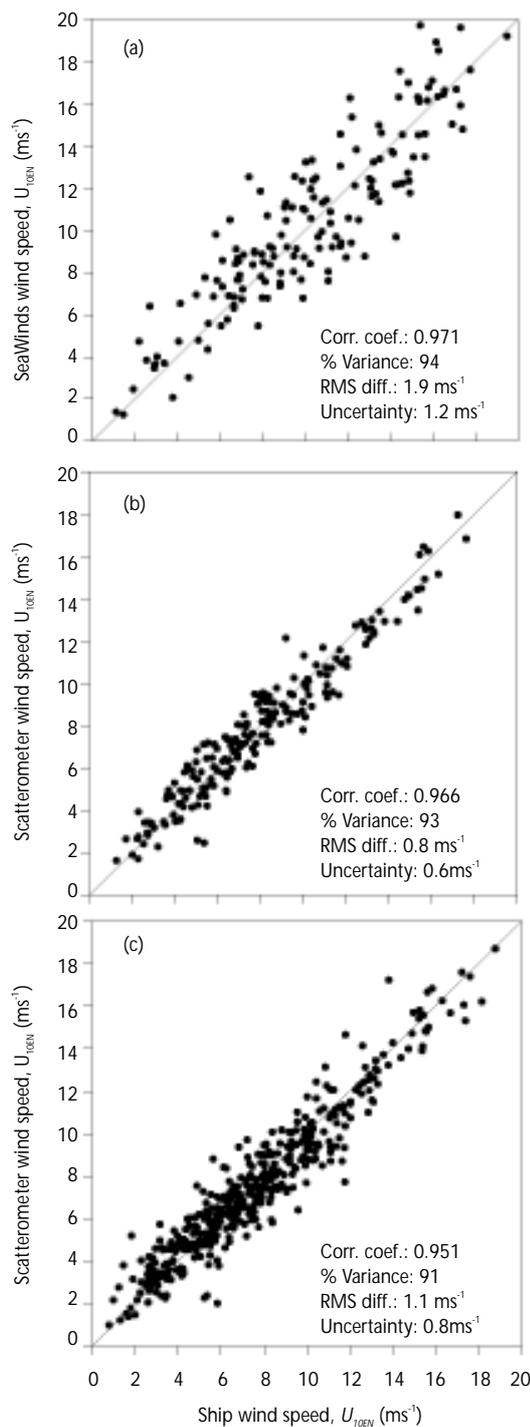
3.3 WIND SPEED CALIBRATION

The collocated pairs of winds are also quality controlled to remove gross errors in wind speeds (possibly related to rain) following the criteria of Freilich and Dunbar (1999). Scatterometer winds are compared to ship true winds (Figure 2). The collocation criteria are the closest match within 25 km and usually within 30 seconds. The close fit to the ideal line shows that there is an extremely good match. The apparent bias towards overestimation at low wind speeds is an expected consequence of comparing two quantities that must be positive, each of which has error characteristics expressed as vectors (Freilich, 1997).

3.4 WIND DIRECTION CALIBRATION

Scatterometers are unique among satellite-based wind sensors, in that they determine the wind direction as well as the wind speed. A scatterplot of scatterometer wind direction versus ship wind directions (Figure 3) shows that there is usually a close match. The solid lines indicate an ideal fit, the dotted lines indicate reversed wind directions, and the dashed lines indicate a 90° difference. The tight cluster around the ideal line indicates that in most cases the correct ambiguity is selected. The percentage of correctly selected ambiguities

Figure 2—Collocated ship and scatterometer winds: (a) 25 km NSCAT-2; (b) Ku-2000; and (c) QSCAT-1 products. The solid line is the ideal fit. The differences in correlation coefficient and variance explained are related to the accuracy, co-location constraints, and data distribution.



(Figure 4) is 90 per cent for the NSCAT 50 km product and 87 per cent for the 25 km product. QSCAT-1 ambiguity selection skill is 91 per cent, and Ku-2000 skill is 93 per cent. The chance that an incorrect alias is selected is dependent on wind speeds. For $U_{10} > 10 \text{ m s}^{-1}$, the chance of an incorrect alias is small, except for the (near nadir) QSCAT-1 winds. For $U_{10} < 10 \text{ m s}^{-1}$, the chance of an incorrect alias increases as the wind speed decreases. QuikSCAT shows a remarkable improvement in low wind speed ($<4 \text{ m s}^{-1}$) ambiguity selection, with the percentage of correct selection being almost double that of NSCAT.

3.5
DEPENDENCE OF
UNCERTAINTY ON SPATIAL
DIFFERENCE IN CO-LOCATION

For QuikSCAT speeds and directions, variance (uncertainty squared) was examined as a function of co-location distance, and a strong dependence on differences in spatial co-location (Figure 5) was found. The rms differences and random uncertainties (one standard deviation) increase as the spatial co-location

Figure 3—Collocated ship and SeaWinds wind directions. The solid line is the ideal fit, dashed lines are 180° errors, and dotted lines are 90° errors.

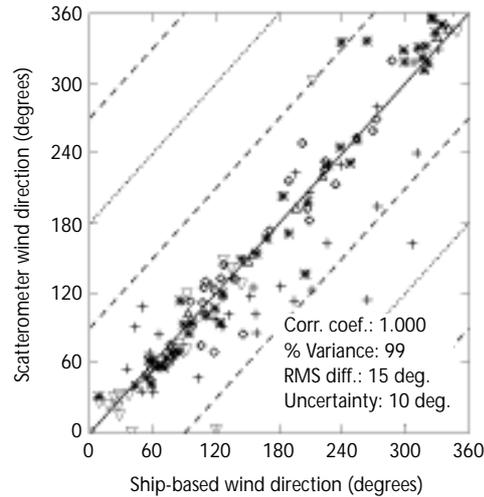
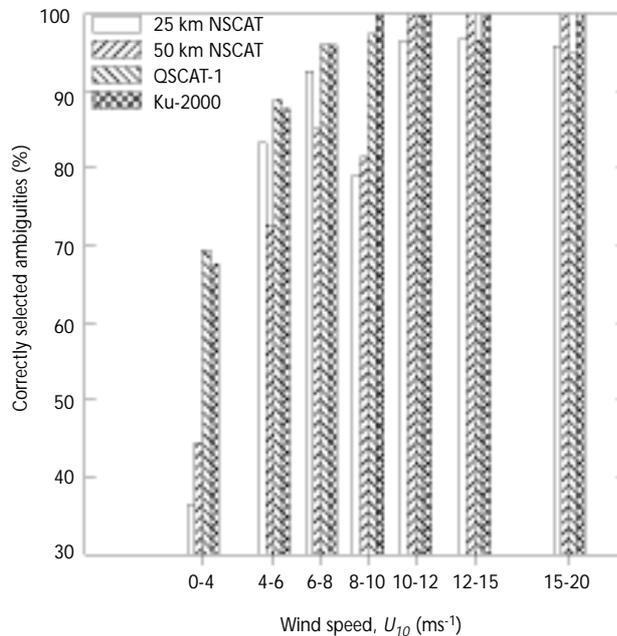


Figure 4—Fraction of correct ambiguity selections for various wind speed bins, for both NSCAT products and SeaWinds on QuikSCAT.

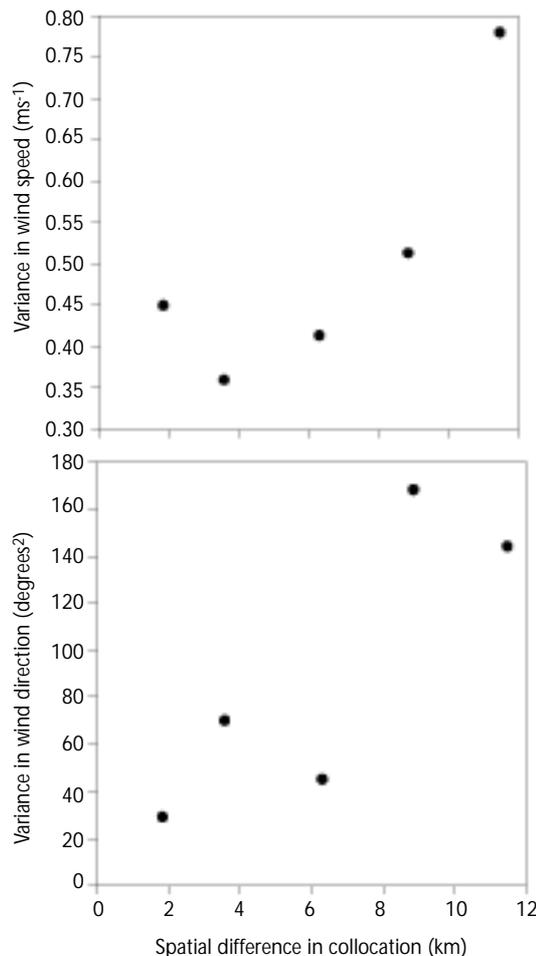


criteria increases. The dependence on spatial differences in co-location can be examined by binning observations in terms of these spatial differences (Figure 5(a)), and then reanalyzing the data in each bin. Variances are determined from the data in each bin, and then extrapolated to zero spatial difference in co-location. The extrapolated variance provides an estimation of observational uncertainty (Bourassa *et al.*, 2001). For co-location distances less than 6 km, the QSCAT-1 uncertainty drops to 0.45 m s^{-1} (0.3 m s^{-1} for Ku-2000; not shown). With this preliminary data set, there is no indication of improvement for smaller co-location distances. Similarly, the variance in direction drops from $\sim 160^{\circ 2}$ for co-location differences of $\sim 10 \text{ km}$, to $\sim 30^{\circ 2}$ for co-location differences of $\sim 2 \text{ km}$. Extrapolation with a parabolic best fit estimates an uncertainty of 4° . However, this result is heavily dependent on the point in the 0-2.5 km bin. Additional tests involving the magnitude of vector differences ($|U_{10\text{EN, scat}} - U_{10\text{EN, ship}}|$) support an uncertainty of 5° (Bourassa *et al.*, 2001).

3.6 CALIBRATION DEPENDENCE ON WIND SPEED

The accuracy of wind speed and direction for correctly selected ambiguities is not a function of wind speed; however, the accuracy of ambiguity selection is a function of wind speed (Figure 4). Ambiguity selection has little impact ($<0.1 \text{ ms}^{-1}$) on wind speed accuracy, but can lead to considerable additional uncertainty in direction. For ERS scatterometers, ambiguity selection is also a function of position in

Figure 5—Variance between ship and SeaWinds wind (a) speeds, and (b) directions as functions of co-location distance. Bin size is 2.5 km.



the observational swath (due to shortcomings in the design of the satellite rather than the scatterometer). For other scatterometers, ambiguity selection is largely a problem for low wind speeds. For very low wind speeds ($<2 \text{ m s}^{-1}$), these directional uncertainties are easily modelled as random component errors (Freilich, 1997; Freilich and Vanhoff, 2000), and this approach works very well for $U_{10} < 8 \text{ ms}^{-1}$. For stronger winds, the random component error model underestimates directional uncertainty (Bourassa *et al.*, 2001). A constant directional uncertainty is a very good model for $U_{10} > 8 \text{ ms}^{-1}$, where ambiguity selection is negligible for the NSCAT-2 and Ku-2000 model functions.

4. GRIDDED WIND PRODUCTS

Most oceanographic applications require winds (or stresses) that are on a regular latitude-longitude grid, and that are regular in time. The difficulties in creating such products are twofold: filling the gaps in observations, and the removal of spurious curl and divergence at swath edges and intersections. A large number of regularly gridded scatterometer wind products have become available in the few years since NSCAT winds were validated. For example, Tang and Liu (1996) filled the gaps in daily fields with ECMWF winds, and then applied an objective interpolation. The need for non-scatterometer data was eliminated through interpolation (Polito *et al.*, 1997; IFREMER/CERSAT, 1998; B. Cheng, 1998, personal communication).

Alternatively, wind fields were generated through spatial and/or temporal averages (Bourassa *et al.*, 1998, 1999b; Kutsuwada, 1998; Kelley *et al.*, 1999). Usually, these interpolation techniques did not adequately remove the observational pattern, and the averaging or smoothing techniques had too great a reduction in kinetic energy.

An approach designed to remove the observational pattern (Chin *et al.*, 1998) has applied wavelet-based resolution analysis to a combination of NSCAT, ERS-2, and NCEP winds. This approach explicitly preserves wind component energy

spectra seen in longer-term wind field averages (Freilich and Chelton, 1986; Milliff *et al.*, 1999b; Wikle *et al.*, 1999). A new approach (Pegion *et al.*, 2000) used a variational method to minimize the misfit to observed pseudostress (the product of scalar and vector winds, which is similar to the surface stress) and minimize the presence of orbital pattern. This approach used cross validation (Wahba and Wendelberger, 1980) to objectively determine the weighting of the constraints used in the variational method. This product has a daily average pseudostress similar to the scatterometer swath winds, and has very little appearance of the observational pattern. The wind component energy spectra are not constrained in this approach; however, these spectra also match the findings of the previous studies. These wind fields clearly show frontogenesis, cyclogenesis, and large-scale wind patterns.

Some caution should be utilized when applying any gridded wind product derived from a polar-orbiting satellite. For example, some gridded products do not adequately deal with the orbital pattern, causing areas with spurious curl and divergence. These spurious features can have considerable negative impact on ocean models. Furthermore, the sampling pattern results in non-uniform error characteristics (Schlax *et al.*, 2000), which can cause features to fluctuate in intensity. Consequently, the accuracy of long-wave or low-frequency signals can be much better than short-wave or high-frequency signals. Local frequency characteristics of the fields can be examined through comparison with buoy winds. For example, the gridded fields of Pegion *et al.* (2000) have been shown to reproduce most of the frequency characteristics of winds from several TAO buoys. Despite these potential problems, scatterometer wind fields are currently the most accurate and highest resolution winds available at this time.

Characteristics of the existing gridded scatterometer products are summarized in Table 5. The listed characteristics are spatial and temporal coverage, and spatial and temporal grid spacing. Links and/or contact information for all the scatterometry products can be found on the COAPS scatterometry web site (<http://coaps.fsu.edu/scatterometry/>). As new products become available, they will be linked to these web pages.

5. WIND ANIMATIONS

Wind fields (Bourassa *et al.*, 1999a; Pegion *et al.*, 2000) were used to produce animations of the winds and vorticity fields. Animations allow the vast quantity of scatterometer data to be easily examined for features of interest. The winds are shown with moving vectors. The motion of the vectors is Lagrangian, and the vector length indicates the wind speed. The changes in vector positions (i.e. motion) are calculated by interpolating the daily wind fields to one hour time steps, and integrating with a fourth order Runge-Kutta method (Kutta, 1901). The Runge-Kutta technique uses an adaptive time step, with a first guess equal to the time interval between frames. This time interval is dependent on the highest wind speeds and tightest circulations; however, a time step of two or three hours was found to be effective for most atmospheric conditions. The animations were designed for easy access. They are available on our web site (<http://coaps.fsu.edu/scatterometry/>) and are split into weekly animations.

These animations have proven to be extremely useful for visualizing the surface winds. For example, they show previously unsuspected directional variability in winds (Tehuantepecers) flowing from the Gulf of Mexico, through Chevela Pass and into the Gulf of Tehuantepec. These winds typically turn to the right; however, when Hurricane Marco was in the Caribbean Sea, they weakened, turned to the left, and moved through mountain passes in Nicaragua and into the Caribbean Sea (Bourassa *et al.*, 1999b). Animations based on the improved fields of Pegion *et al.* (2000) also reveal eddies-associated westerly bursts during the onset of the 1997/98 El Niño.

6. CONCLUSIONS

Quality-controlled and high temporal resolution wind observations from R/Vs have proven to be effective in providing surface comparison data to evaluate the accuracy of scatterometer winds. The SeaWinds design (Ku-band, with large incidence angles) is more sensitive to rain than NSCAT (Ku-band, with smaller

ADVANCES IN THE APPLICATIONS OF MARINE CLIMATOLOGY

<i>Scatterometer gridded product</i>	<i>Spatial coverage</i>	<i>Spatial grid</i>	<i>Temporal coverage</i>	<i>Temporal grid</i>	<i>Data fields</i>	<i>Input data</i>	<i>Processing technique</i>
NSCAT project fast look	Global (in swaths)	$0.5 \times 0.5^\circ$	9/15/96 to 6/29/97	Daily	u, v	NSCAT	Vector average within swaths
QuikSCAT project fast look	Global (in swaths)	$0.25 \times 0.25^\circ$	7/19/99 ongoing	Daily	u, v	QSCAT	Vector average within swaths
Cheng, Chao and Liu	Global	$1 \times 1^\circ$	9/15/96 to 6/29/97	2 Days	u, v	NSCAT, ECMWF	Gaussian-weighted
COAPS/FSU objectively analysed	Indian Ocean 34.5S - 28.5N, 25.5E - 124.5E	$1 \times 1^\circ$	9/15/96 to 6/29/97	Daily	UW, VW	NSCAT	Variational method, with objectively determined weights
	Pacific Ocean 34.5S - 28.5N, 25.5E - 124.5E	$1 \times 1^\circ$	9/15/96 to 6/29/97	Daily	UW, VW	NSCAT	
	Global	$1 \times 1^\circ$	9/15/96 to 6/29/97	Daily	UW, VW	NSCAT	
	Global		07/20/99 ongoing	4×daily		QSCAT	
COAPS/FSU temporal averaged	Global	$1 \times 1^\circ$	9/15/96 to 6/29/97	Daily	u, v	NSCAT	Centered, temporally weighted, mean
COAPS/FSU monthly stresses	Global	$0.5 \times 0.5^\circ$	10/96 to 6/97	Monthly	τ_x, τ_y	NSCAT	Temporal mean
Ifremer/Cersat	Global	$1 \times 1^\circ$	8/5/91 to 5/1/98	Bi-weekly and Bi-monthly	u, v, τ_x, τ_y , wind div., wind curl	NSCAT and ERS-1/2 Winds	Objective interpolation with a minimum variance method
Kelly, Caruso and Dickinson	Tropical Pacific	$1 \times 1^\circ$	10/1/96 to 6/26/97	Daily	UW, VW	NSCAT	Objective average
Kutsuada	30E to 90W	$1 \times 1^\circ$	9/15/96 to 6/29/97	Daily	yu, v, τ_x, τ_y	NSCAT	Weighted mean vectors
Chin, Milliff and Large	Gobal	$0.5 \times 0.5^\circ$	8/1/96 to 7/31/97	6 hours	u, v	NSCAT, ERS-2, NCEP	Wavelet-based multi-resolution analysis
Polito, Liu and Tang	Global	$1 \times 1^\circ$	9/15/96 to 6/29/97	Daily	u, v, τ_x, τ_y , Div. of stress, Ekman pumping	NSCAT	Correlation- based interpolation
Tang and Liu	Global	$0.5 \times 0.5^\circ$	9/15/96 to 6/29/97	12 hours	u, v	NSCAT, ECMWF QSCAT, ECMWF	Successive correction
		$0.25 \times 0.25^\circ$	9/03/99 ongoing	12 hours	u, v		

Table 5—Regularly gridded products using scatterometer data. The symbols in the data fields column are zonal wind (u), meridional wind (v), wind speed (w), zonal pseudostress (UW), meridional pseudostress (VW), zonal stress (τ_x), and meridional stress (τ_y). More information on these data sets, including format and access, is available through our web site (<http://coaps.fsu.edu/scatterometry/>).

incidence angles) or ERS-1/2 (C-band, with smaller incidence angles). Therefore, a rain flag based on QuikSCAT observations is used to remove rain from a set of co-located observations. The uncertainty in the comparison data set and differences in co-location were shown to be essential to the accurate estimate of uncertainty in the satellite winds. Principal component analysis (PCA) was used with co-location differences of less than 12.5 km and 10 minutes to estimate uncertainties for correctly selected ambiguities. The impact of co-location distance was shown by binning variance (uncertainty squared) as a function of co-location distance. Consideration of co-location differences and uncertainty in the comparison data set resulted in uncertainty estimates (for correctly selected ambiguities) of 0.45 m s^{-1} and 5° the SeaWinds QSCAT-1 model function, and 0.3 m s^{-1} and 3° the SeaWinds Ku-2000 model function. The excellent coverage and great accuracy of modern scatterometers will lead to greatly improved wind climatologies, as well as improved wave climatologies based on these winds.

Ambiguity selection was shown to be good for the NSCAT-2 model function (88 per cent), and excellent for SeaWinds observations (91 per cent). For $U_{10} > 10 \text{ m s}^{-1}$, the chance of an incorrect ambiguity selection is extremely small. Most of the ambiguity selection problems occurred in the $0 < U_{10} < 6 \text{ m s}^{-1}$ range. The greatest difference between NSCAT and SeaWinds ambiguity selection is for $U_{10} < 4 \text{ m s}^{-1}$, where SeaWinds is almost twice as effective. Many of the ambiguity errors associated with low wind speeds are likely to be associated with uncertainties in both the scatterometer observations and the comparison data set (Freilich, 1997).

The NSCAT and SeaWinds on QuikSCAT winds are more than sufficiently accurate for oceanographic studies on space/time scales greater than 50 km and three days. For regularly gridded products, the winds must be processed in a manner that removes errors related to the observational pattern and retains the observed pseudostress. The gridded products to date (Table 5) have varying degrees of success in meeting these goals. The fully objective technique of Pegion *et al.* (2000) does an excellent job of retaining the observed pseudostress without the appearance of the observational pattern. This consideration is essential for forcing ocean models, as the appearance of the observational pattern is synonymous with spurious wind forcing.

An excellent tool for visualizing the evolution of the wind fields is moving vector animation. The large-scale animations provide a good example of synoptic scale motion and general circulation patterns. The smaller scale animations reveal the larger mesoscale variations. These animations have proven to be useful for finding previously unexpected wind motions and vorticity patterns.

There is a wealth of user-friendly and publicly available scatterometry products. These include swath observations, gridded products, graphics, and animations as well as background information. An updated listing of all these products is available at <http://coaps.fsu.edu/scatterometry/>.

ACKNOWLEDGEMENTS

We thank the many people who provided observations from the RSV Aurora Australis, R/V Atlantis, R/V Knorr, R/V Melville, R/V Meteor, R/V Oceanus, R/V Polarstern, R/V Ronald Brown, R/V Thompson, and those who quality controlled the observations. The scatterometer data were obtained from the NASA Physical Oceanography Distributed Active Archive Center at the Jet Propulsion Laboratory/California Institute of Technology, and Remote Sensing Systems. Support for the scatterometer research came from the NASA/OSU SeaWinds project and the NASA OVWST project. NSF support of the WOCE DAC/SAC for surface meteorology funded the quality control of research vessel data. COAPS receives its base funding from the Secretary of Navy Grant from ONR to James J. O'Brien.

REFERENCES

Atlas, R., S.C. Bloom, R.N. Hoffman, E. Brin, J. Ardizzone, J. Terry, T.D. Bungato, J.C. Jusem, 1999: Geophysical validation of NSCAT winds using atmospheric data and analyses. *J. Geophys. Res.*, 104, 11,405-11,424.

- Bliven, L.F., H. Branger, P.C. Sobieski and J-P. Giovanangeli, 1993: An analysis of scatterometer returns from a water surface agitated by artificial rain. *International Journal of Remote Sensing*, 14, 2315-2329.
- Bourassa, M.A., M.H. Freilich, D.M. Legler, W.T. Liu and J.J. O'Brien, 1997: Wind observations from new satellite and research vessels agree. *EOS Trans. of Amer. Geophys. Union*, 78, 597 & 602.
- Bourassa, M.A., D.G. Vincent and W.L. Wood, 1999a: A flux parameterization including the effects of capillary waves and sea state. *J. Atmos. Sci.*, 56, 1123-1139.
- Bourassa, M.A., L. Zamudio and J.J. O'Brien, 1999b: Non-inertial flow in NSCAT observations of Tehuantepec winds. *J. Geophys. Res.*, 104, 11,311-11,320.
- Bourassa, M.A., D.M. Legler, J.J. O'Brien and S.R. Smith, 2001: SeaWinds validation with research vessels, *J. Geophys. Res.*, submitted.
- Bourassa, M.A., D.M. Legler, J.J. O'Brien, J.N. Stricherz and J. Whalley, 1998: High temporal and spatial resolution animations of winds observed with the NASA scatterometer, 14th International conference on IIPS (January, Phoenix, AZ), American Meteorological Society, 556-559.
- Cardone, V.J. (1969): Specification of the wind distribution in the marine boundary layer for wave forecasting. *Geophys. Sci. Lab. N.Y. Univ. Report TR-69-1* (NTIS AD 702-490).
- Cardone, V.J., R.E. Jensen, D.T. Resio, V.R. Swail and A.T. Cox, 1996: Evaluation of contemporary ocean wave models in rare extreme events: The "Halloween Storm" of October 1991 and the "Storm of the Century" of March 1993. *J. Atmos. Oceanic Technol.*, 13, 198-230.
- Caruso, M.J., K.A. Kelly, M. McPhaden and S. Dickinson, 1997: Evaluation of scatterometer winds using Equatorial Pacific buoy observations. *Proceedings of the NASA Scatterometer Symposium* (10-14 November 1997 Maui, Hawaii) [Available from the Scatterometer Projects Office, Jet Propulsion Lab., Pasadena, California].
- Caruso, M.J., S. Dickinson, K.A. Kelly, M. Spillane, L.J. Mangum, M. McPhaden and L.D. Stratton, 1999: Evaluation of NSCAT scatterometer winds using Equatorial Pacific buoy observations. *WHOI Technical Report 99-10* [Available from Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA].
- Chelton, D.B. and F.J. Wentz, 1986: Further development of an improved altimeter wind speed algorithm, *J. Geophys. Res.*, 91, 14250-14260.
- Chelton, D.B., M.H. Freilich and S.K. Esbensen, 2000: Satellite observations of the wind jets off Central America, Part II: Regional relationships and dynamical considerations. *Mon. Wea. Rev.*, 128, 1993-2018.
- Chin, T.M., R.F. Milliff and W.G. Large, 1998: Basin scale, high-wavenumber sea surface wind fields from a multi-resolution analysis of scatterometer data. *J. Atmos. Ocec. Tech.*, 15, 741-763.
- Dickenson, S., K.A. Kelly, M.J. Caruso and M.J. McPhaden, 2000: A note on comparisons between the TAO buoy and NASA scatterometer wind vectors. *J. Atmos. Ocec. Tech.*, accepted.
- Ebuchi, N., 1999: Statistical distribution of wind speeds and directions globally observed by NSCAT. *J. Geophys. Res.*, 104, 11393-11404.
- Freilich, M.H., 1997: Validation of vector magnitude data sets: effects of random component errors. *J. Atmos. Oceanic Technol.*, 14, 695-703.
- Freilich, M.H. and D.B. Chelton, 1986: Wavenumber spectra of Pacific winds measured by the SeaSat scatterometer. *J. Phys. Oceanogr.*, 16, 741-757, 1986.
- Freilich, M.H. and R.S. Dunbar, 1999: The accuracy of the NSCAT 1 vector winds: Comparison with National Data Buoy Center buoys. *J. Geophys. Res.* 104, 11231-11246.
- Freilich M.H. and B.A. Vanhoff, 2000: The accuracy of remotely sensed surface wind speed measurements. *J. Atmos. Oceanic Technol.*, in review.
- Gonzales, A.E. and D.G. Long, 1999: An assessment of NSCAT ambiguity removal. *J. Geophys. Res.*, 104, 11449-11458.

- Graber, H.C., N. Ebuchi, R. Vakkayil, 1996: Evaluation of ERS-1 scatterometer winds with wind and wave ocean buoy observations. *RSMAS Technical Report 96-003* [available from the Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL 33149-1098, USA].
- Graber, H.C., A. Bentamy and N. Ebuchi, 1997: Evaluation of scatterometer winds with ocean buoy observations. *Proceedings of the NASA Scatterometer Symposium* (10-14 November 1997, Maui, Hawaii) [available from the Scatterometer Projects Office, Jet Propulsion Lab., Pasadena, California].
- IFREMER/CERSAT, 1998: Mean surface wind fields from the ERS-AMI and ADEOS-NSCAT microwave scatterometers, 91/08/05 to 98/03/01. A contribution to WOCE, a CD-ROM published for the WOCE conference (Halifax, Canada).
- Jones, W.L., V.J. Cardone, W.J. Pierson, J. Zec, L.P. Rice, A. Cox and W.B. Sylvester, 1999: NSCAT high-resolution surface wind measurements in Typhoon Violet. *J. Geophys. Res.*, 104, 11247-11260.
- Katsaros, K.B. and R.A. Brown, 1991: Legacy of SeaSat mission for studies of the atmosphere and air-sea-ice interactions. *Bull. Amer. Meteor. Soc.*, 72, 967-981.
- Kelly, K.A., S. Dickinson and Z.-J. Yu, 1999: NSCAT Tropical wind stress maps: Implications for improving ocean modeling. *J. Geophys. Res.*, 104, 11291-11310.
- Kutta, W. 1901: Beitrag zur näherungsweise integration totaler differentialgleichungen. *Z Math. Phys.* 46, 435-453.
- Kutsuwada, K., 1998: Impact of wind/wind-stress field in the North Pacific constructed by ADEOS/NSCAT data., *J. Oceanogr.*, 54, 443-456.
- Large, W.G., J. Morzel and G.B. Crawford, 1995: Accounting for surface wave distortion of the marine wind profile in low-level ocean storm wind measurements, *J. Phys. Oceanogr.*, 25, 2959 - 2971.
- Liu, W.T. and W. Tang, 1996: Equivalent Neutral Wind, *JPL Publication 96-17*, Jet Propulsion Laboratory, California Institute of Technology, CA.
- Milliff, R.F., T.J. Hoar, H. van Loon and M. Rapheal, 1999a: Quasi-stationary wave variability in NSCAT winds. *J. Geophys. Res.*, 104, 11,425–11,436.
- Milliff, R.F., W.G. Large, J. Morzel, G. Danabasoglu and T.M. Chin, 1999b: Ocean general circulation model sensitivity to forcing from scatterometer winds. *J. Geophys. Res.*, 104, 11,337-11,358.
- Moore, R.K., D. Chatterjee and S. Taherion, 1999: Algorithm for correcting SeaWinds/ADEOS-II for rain attenuation. *Proceedings of the QuikSCAT Cal/Val – Early Science Meeting* (Arcadia, CA.) [Available from Scatterometer Projects Office, Jet Propulsion Lab., Pasadena, CA 91109].
- Naderi, F.M., M.H. Freilich and D.G. Long, 1991: Spaceborne radar measurements of wind velocity over the ocean — an overview of the NSCAT scatterometer system. *Proceedings of the IEEE*, 79, 850-866.
- Pearson, K., 1901: On lines and planes of closest fit to systems of points in space. *Phil. Mag.*, 2, 59-572.
- Pegion, P.J., M.A. Bourassa, D.M. Legler and J.J. O'Brien, 2000: Objectively derived daily pseudostress fields from NSCAT data created through direct-minimization, cross validation, and multigridding. *Mon. Wea. Rev.*, 128, 3150–3168.
- Pierson, W.J., Jr, 1990: Examples of, reasons for and consequences of the poor quality of wind data from ships for the marine boundary layer: implications for remote sensing, *J. Geophys. Res.*, 13,313 - 13,340.
- Polito, P.S., W.T. Liu and W. Tang, 1997: Correlation-based interpolation of NSCAT wind data. NSCAT Science Working Team Meeting (Maui, Hawaii, November, 1997).
- Preisendorfer, R.W. and C.D. Mobley, 1988: *Algebraic Foundations of PCA, Principal Component Analysis in Meteorology and Oceanography*. Elsevier, 11 - 24.
- Ross, D.B., V.J. Cardone, J. Overland, R.D. McPherson, W.J. Pierson Jr. and T. Yu, 1985: Oceanic surface winds. *Adv. Geophys.*, 27, 101-138.
- Schlx, M.G., D.B. Chelton and M.H. Freilich, 2000: The effects of sampling errors in wind fields constructed from single and tandem scatterometer datasets. *J. Atmos. Oceanic Technol.*, submitted.

- Smith, S.R., M.A. Bourassa and R.J. Sharp, 1999: True winds from vessels at sea: Requirements and implications for automated maritime observing systems, *J. Atmos. Oceanic Technol.*, 16, 939–952.
- Sobieski, P.W, C. Craeye and L.F Bliven, 1999: Scatterometric signatures of multi-variate drop impacts on fresh and salt water surfaces. *International Journal of Remote Sensing*, 20: 2149-2166.
- Sobieski, P. and L.F. Bliven, 1995: Analysis of high speed images of raindrop splash products and ku-band scatterometer returns. *International Journal of Remote Sensing*, 16, 2721-2726.
- Stoffelen A., 1998: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *J. Geophys. Res.*, 103, 7755 - 7766.
- Tang, W. and W.T. Liu, 1996: *Objective Interpolation of Scatterometer Winds*. Technical Report 96-19, Jet Propulsion Laboratory, California Institute of Technology, CA.
- Thiebaux, M.L., 1990: Wind tunnel experiments to determine correction functions for shipborne anemometers. *Canadian Contractor Rep. Hydrog. & Ocean Sci.* 36, BIO, Dartmouth, Nova Scotia, 57 pp.
- Verschell, M.A., M.A. Bourassa, D.E. Weissman and J.J. O'Brien, 1999: Model validation of the NASA scatterometer winds, *J. Geophys. Res.*, 104, 11,359–11,374.
- Wahba, G. and J. Wendelberger, 1980: Some new mathematical methods for variational objective analysis using splines and cross-validation. *Mon. Wea. Rev.*, 108, 1122-1143.
- Weissman, D.E., M.A. Bourassa and J. Tongue, 2000: Effects of rain-rate and wind magnitude on SeaWinds scatterometer wind speed errors. *J. Atmos. Oceanic Technol.*, submitted.
- Wentz, F.J. and D.K. Smith, 1999: A model function for the ocean-normalized radar cross section at 14 GHz derived from NSCAT observations. *J. Geophys. Res.*, 104, 11499-11514.
- Wikle, C.K., R.F. Millif and W.G. Large, 1999: Surface wind variability on spatial scales from 1 to 1000 km observed during TOGA COARE. *J. Atmos. Sci.*, 56, 2222-2231.
- Yelland, M.J., B.I. Moat, P.K. Taylor, R.W. Pascal, J. Hutchings and V.C. Cornell, 1998: Wind stress measurements from the open ocean corrected for airflow distortion by the ship. *J. Phys. Oceanogr.*, 28, 1511–1526.
- Zecchetto, S, I. Tando and Y. Quilfen, 1999: The CERSAT dealiasing of ERS-2 scatterometer winds in the Mediterranean Sea. *CERSAT NEWS*, Scientific Topic number 2, 1-4.

SECTION 5

ANALYSIS OF CLIMATE VARIABILITY AND CHANGE

Outlier detection in gridded ship data sets	177
A methodology for integrating wave data from different sources permitting a multiscale description of wave climate variability	187
Reduced space approach to the optimal analysis of historical marine observations: accomplishments, difficulties, and prospects	199
Analysis of wave climate trends and variability	217

OUTLIER DETECTION IN GRIDDED SHIP DATA SETS

Pascal Terray, Laboratoire d'Océanographie Dynamique et de Climatologie, and Université Paris 7, Paris, France

1. INTRODUCTION

This is the second of two papers attempting to develop robust statistical methods to deal with gridded ship data sets. The earlier study (Terray, 1999) focused on an extension of the traditional empirical orthogonal function (EOF) analysis which allows arbitrary positive weights to be assigned to each entry of the data matrix. If these weights are constructed in a responsible manner (for example, as a smooth function of the number of ship reports used to compute a particular raw monthly mean in the data set), it was demonstrated that this method allows us to analyse the natural variability exhibited by gridded ship data sets by directly taking into account the irregular space-time sampling of marine observations. In particular, the method takes care of missing values by assigning zero weights to such data entries.

In the current study, we discuss another robust statistical method to detect 'local errors' in gridded ship data sets. More precisely, we tackle the problem of outlying areal averages in gridded ship data sets such as the Comprehensive Ocean-Atmosphere Data Set (COADS; Woodruff *et al.*, 1987) 2° lat \times 2° long monthly summaries and how to test their statistical significance. Since the majority of climate researchers use gridded ship data sets instead of individual ship reports, we suggest that these data sets must be checked for the presence of doubtful raw monthly means in the same manner as individual ship reports are quality controlled before being integrated in ship reports databases. Moreover, it should be noted that such an approach may be a solution to the trimming problems which are apparent in COADS monthly summaries (Wolter, 1997).

The rest of this paper is organized as follows. First, we present some elements of outlier detection theory and the basic statistical tests we have used. Next, we discuss how these statistical tests may be adapted to ship data sets and integrated as building blocks in a fully computerized procedure for detecting many outliers in such data sets. Finally, this new approach has been experimented on a marine product in order to show how it works in practice. As a conclusion, we suggest that the two procedures, namely outlier detection and weighted EOF analysis, may be combined to obtain a truly robust statistical method particularly well suited to gridded ship data sets.

2. STATISTICAL THEORY OF OUTLIER DETECTION

In the context of gridded ship data sets, an outlying observation, or 'outlier', is a raw monthly mean in a 2° lat \times 2° long box (depending on the resolution of the data set) that appears to deviate markedly from adjacent or neighbouring grid-points in area or/and in time. Outliers in gridded ship data sets may be generated by three basic mechanisms (Wolter, 1997):

- An outlying raw monthly mean may be merely an extreme manifestation of the sampling inherent in the data, since some raw monthly means in 2° lat \times 2° long boxes are computed with very few marine observations for a given date while adjacent boxes may be well sampled.
- Outlying raw monthly means in some 2° lat \times 2° long boxes may also be the results of potential biases due to the origin of the 'source-decks' merged into the gridded ship data set or processing errors. For example, biases in sea surface temperature (SST) associated with different methods of measurements (bucket or intake) may well introduce errors in gridded ship data sets in particular atmospheric conditions and along some ship tracks.

- Finally, an outlying areal average may be the result of errors relating to instrumental readings or coding mistakes. But, most of these types of outliers must be discovered during basic quality controls which are automatically applied to individual ship reports merged into any reasonable marine product.

The problem of detecting outliers in a random sample has been extensively researched by statisticians in recent years and a number of test statistics are available for both the single outlier case and the many outlier case for testing a specified number k of outliers (Barnett and Lewis, 1978). In particular, the detection of outliers in a normal sample has received considerable attention. It is far beyond the scope of this paper to give a review of this vast subject. Suffice to say here, that the problem of outlier detection is generally treated as the statistical testing of a hypothesis. The null hypothesis, as usually stated, is that all the observations are drawn from the same (normal) population; the alternative hypothesis is that at least one of the observations has been drawn from another distribution. To discriminate between these two hypotheses, a sample criterion T which uses the doubtful observation(s) is calculated. This statistic is then compared with a critical value λ_α based on the theory of random sampling to determine whether the doubtful observation is to be retained or rejected. This critical value is the value of the chosen sample criterion which would be exceeded by chance with some specified and small probability α (say 0.01 or 0.05), which is the so-called significance level of the test, if the null hypothesis is true. Intuitively, this significance level is the risk of erroneously rejecting a good observation (statistical type I error). More precisely, statistical tests for outliers are the following:

- (1) Find λ_α such that $\Pr(T > \lambda_\alpha) = \alpha$ if the null hypothesis is true for some statistic T ;
- (2) Reject the null hypothesis and declare an outlier present if $T > \lambda_\alpha$, or accept the null hypothesis and declare the sample is clean if $T \leq \lambda_\alpha$.

In this statistical framework, outlier detection procedures differ by:

- The form of the underlying parent population (normal, gamma, etc.);
- The form of the test criterion T which has to be computed on the sample: among these test criteria, we can distinguish those which clearly identify particular observations as possible outliers from those which test the hypothesis that the random sample as a whole did indeed come from the specified parent distribution;
- The number of suspected outliers in the sample;
- The fact that the doubtful observations may be to one side of the bulk of the data or that some are too large and some are too small.

Several hundreds of statistical tests of this type are described in the book written by Barnett and Lewis (1978) which is a kind of 'bible' on the subject. In the context of gridded ship data sets, the problem is then to decide which tests to apply, and how to use them in order to obtain a fully computerized procedure for detecting outliers which may be applied to any ship data set. In this way, one can hope to trap anomalous cases and so ensure the integrity of most of the ship data sets currently in use.

We have used here a simple model, which is well documented in the statistical literature: when the data with the possible exception of any outlier form a sample from a normal distribution with unknown mean μ and unknown variance σ^2 . We recognize that this model is certainly not perfect in the context of samples of adjacent raw monthly means in 2° lat \times 2° long boxes extracted from gridded ship data sets. However, as we will show below, this model works 'reasonably' well as implemented in our computerized procedure on the basis of the spatial coherence of neighbouring 2° lat \times 2° long area values for many meteorological parameters. Several reasonably powerful statistical tests exist to detect one outlier in a normal sample, and our approach involves the following classical statistical criteria:

Let x_1, x_2, \dots, x_n be the observations of a random sample. Order the observations according to increasing magnitude and denote the i_{th} largest by y_i ; thus, $y_1 \leq y_2 \leq \dots \leq y_n$ is the ordered set of observations. Suppose the largest observation y_n is suspect. To test for discordancy in this single upper observation in a normal sample, a reasonable test statistic is:

$$T = \frac{y_n - \bar{x}}{s}$$

where:

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$$

is the sample mean, and

$$s^2 = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{n} \right)^2$$

is the sample variance calculated with n degrees of freedom.

If y_1 , the lower observation, rather than y_n , is the doubtful value, the criterion is as follows:

$$T' = \frac{\bar{x} - y_1}{s}$$

and the rest of the statistical procedure will be unchanged on the basis of the symmetry of the normal distribution. Finally, when it is not known a priori whether the contaminant is the lower or the upper observation in the sample, we should compute:

$$T^* = \max(T, T')$$

But, in this last case, we must use a critical value corresponding to the $\alpha/2$ significance level if we want the true significance level to be 0.05.

The rationale behind these tests may be found in Hawkins (1980) or Barnett and Lewis (1978). The null hypothesis that we are testing in every case is that all the observations in the sample come from the same normal population. It may be shown that these statistics are optimal in the sense of maximizing the probability of correct identification of an outlier when one is present. It should be noted, however, that these statistics may produce quite misleading results in the presence of many outliers, especially when suspected values are closer to each other than to the bulk of the other observations. This inability of a testing procedure to identify even a single outlier in the presence of several suspected values is called the masking effect. We will discuss this point further in the next section when we describe our computerized procedure for detecting outliers.

Before using these test statistics in outlier checks, we must know the significance probability attached to an observed value t of the statistic T (or T', T^*). That is to say, the probability that, on the null hypothesis of no contamination, T takes values more discordant than t . For this purpose, we need to find the null distribution of T or at least some fractiles λ_α of this distribution corresponding to specified significance levels α , say 0.01 or 0.05. The null distribution of T is available as a recursion relationship (Barnett and Lewis, 1978) or as a complicated multiple integral (Grubbs, 1950), and tables containing critical values for some standard significance levels have been published (Grubbs and Beck, 1972; Hawkins, 1980). However, we will show how approximate critical values for a given significance level α can be computed since our computerized procedure may involve a number of observations outside the range of these published tables.

Without loss of generality, we consider only the case of an upper outlier; approximate fractiles for T' or T^* may be derived similarly. We may compute some fractiles of the test distribution of T as follows:

Under the null hypothesis of no contamination, x_1, x_2, \dots, x_n are observations of random variables X_1, X_2, \dots, X_n which are independent and identically distributed as $N(\mu, \sigma^2)$. In this case, if x_i is an observation selected arbitrarily from the random sample of n items, it may be shown that if:

$$T_i = \frac{x_i - \bar{x}}{s}$$

then the probability density function of:

$$t_i = \frac{T_i \sqrt{n-2}}{\sqrt{n-1-T_i^2}}$$

is given by the 'student's' t-distribution with n-2 degrees of freedom. This is easily verified because t_i is the test statistic of the classical student's two sample t-test, where one sample consists of x_i and the second sample of the n-1 other observations. From this result, we are able to find the probability that an arbitrary observation i will be outlying since:

$$\Pr[T_i > \lambda] = \Pr \left[t_{(n-2)} > \frac{\lambda \sqrt{n-2}}{\sqrt{n-1-\lambda^2}} \right]$$

where λ is an arbitrary value in the range $]-\sqrt{n-1}, \sqrt{n-1}[$ and $t_{(n-2)}$ follows a student's t-distribution with n-2 degrees of freedom. However, this result does not yet give us an exact test for one outlier, because this probability is different from the probability that a particular observation (the lowest or the largest) will be greater than λ . More precisely, we need the distribution not of an arbitrary T_i , but of T, the greatest of the quantities T_i for $i=1$ to n.

Now, note that the event $(T > \lambda)$ is the union of the n events $(T_i > \lambda)$. Thus:

$$\Pr[T > \lambda] = \Pr \left[\bigcup_{i=1}^n (T_i > \lambda) \right]$$

In other words, the probability of the event $(T > \lambda)$ is the probability that at least one of the n events $(T_i > \lambda)$ is true. Bounds on $\Pr[T > \lambda]$ may then be obtained in terms of the component events $(T_i > \lambda)$ through the use of the so-called Bonferroni inequality (Feller, 1968):

$$\sum_i \Pr[T_i > \lambda] - \sum_{i < j} \Pr[(T_i > \lambda) \cap (T_j > \lambda)] \leq \Pr[T > \lambda] \leq \sum_i \Pr[T_i > \lambda]$$

Since the events $(T_i > \lambda)$ are equiprobable, and likewise the events $(T_i > \lambda) \cap (T_j > \lambda)$, we have the following inequality for arbitrary i and j:

$$n \Pr[T_i > \lambda] - \frac{n(n-1)}{2} \Pr[(T_i > \lambda) \cap (T_j > \lambda)] \leq \Pr[T > \lambda] \leq n \Pr[T_i > \lambda]$$

Now, by using the fact that for arbitrary i and j (Doornbos, 1966):

$$\Pr[(T_i > \lambda) \cap (T_j > \lambda)] < (\Pr[T_i > \lambda])^2$$

we finally obtain:

$$n \Pr[T_i > \lambda] - \frac{n-1}{2n} (n \Pr[T_i > \lambda])^2 < \Pr[T > \lambda] \leq n \Pr[T_i > \lambda]$$

for an arbitrary i. Thus, if:

$$\frac{\lambda \sqrt{n-2}}{\sqrt{n-1-\lambda^2}}$$

is the $1 - (\alpha/n)$ fractile of the student's t-distribution with n-2 degrees of freedom, the last equation shows that:

$$\alpha - \frac{n-1}{2n} (\alpha^2) < \Pr[T > \lambda] \leq \alpha$$

A result indicating that λ is a good and conservative approximation of the true critical value λ_α of the distribution of T under the null hypothesis of no contamination for any reasonable significance level α , say 0.01 or 0.05. Moreover, it can be shown that this method gives the exact critical value λ_α of T if:

$$\lambda \geq \sqrt{\frac{n-1}{2}}$$

(for example, the 0.05 critical value for any $n < 15$) since in this case we have:

$$\Pr[(T_i > \lambda) \cap (T_j > \lambda)] = 0$$

for arbitrary i and j . Following the same procedure, we may approximate the true critical value λ_α of T^* on the null hypothesis of no contamination by λ^* , if:

$$\frac{\lambda^* \sqrt{n-2}}{\sqrt{n-1-(\lambda^*)^2}}$$

is the $1 - \alpha/(2n)$ fractile of the student's t -distribution with $n-2$ degrees of freedom.

3. OUTLIER DETECTION IN GRIDDED SHIP DATA SETS

- Suppose now that we want to check the 'local' consistency of a given ship data set with, say, a 2° lat \times 2° long resolution. This data set may contain raw data or anomaly fields after removal of the annual cycle with a climatology. In both cases, the same algorithm is used and the preceding theoretical results are then used as follows:
- (1) First, we specify upper and lower limits for detecting doubtful monthly mean or anomaly values in the gridded ship data set. These limits may vary depending on calendar month and area. Any value which exceeds the upper limit, or is less than the lower limit, is considered a priori doubtful and will be tested for compatibility with monthly mean or anomaly values in adjacent or neighboring 2° lat \times 2° long boxes. These upper or lower limits determine the number of values which will be tested in the detection procedure for a given gridded data set. Thus, if we want to test nearly all the data values for compatibility, we just have to specify a very low upper limit and/or a very high lower limit in the algorithm. Such a choice means that the algorithm will use more computer time since a lot of data values will be tested; but in any case, a data value will be declared an outlier only on the basis of the probabilities of rare events as outlined in section 2 (see below).
 - (2) For any date, doubtful monthly mean or anomaly values identified in step 1 are arranged from the most outlying to the most inlying compared to the bulk of the data. For this purpose, absolute values of residuals of these doubtful values from the overall mean of the observed data for this date are sorted in descending order, and the doubtful values are ranked accordingly.
 - (3) These doubtful values are then considered consecutively, from the most outlying to the most inlying, and a sample is constructed from adjacent or neighbouring 2° lat \times 2° long area values for any of these possible outliers. The number of 2° lat \times 2° long boxes in the vicinity of each doubtful value which are scanned, in order to construct a sample, may be chosen by the user before running the procedure. It should be noted that the number of items in this sample may vary depending on the date and the area. However, the significance level α of the test will be the same for any suspected raw monthly mean value, as we will see below.
 - (4) At this stage, several different possibilities exist:
 - (a) First, we need to consider the case when it is not possible to pick up a sample to test the doubtful value because none of the surrounding boxes contain data. Frequently, this means that the doubtful raw monthly mean is calculated from very few ship reports. In such a case, the user may decide, before running the procedure, to flag or reject all these unrepresentative values.
 - (b) Second, suppose that there is only one doubtful value in the collected sample, the one we want to test. If this value is at the upper end of the sample, we use T as a test criterion; if it is at the lower end of the sample, the statistic T' is considered instead. In both cases, the doubtful value is declared an outlier if the statistic exceeds the critical value λ_α corresponding to a specified significance level α . In this case, the suspected value is rejected or flagged (a user choice) and the next most outlying doubtful value is processed.
 - (c) Finally, imagine that there is more than one doubtful value in the constructed sample of n items, according to the upper and lower limits specified in step 1. Let K be the number of such doubtful values and x be the

suspected value that we are currently processing. In order to take into account the possibility that the sample contains more than one outlier, a consecutive procedure is applied. One naive approach is to use repeated applications of the single outlier statistical test T^* described above, deleting the 'outlier' detected at each step and applying the test again to the reduced sample until an insignificant result is obtained or the suspected value x is tested for compatibility with the remaining observations. However, this 'forward selection' approach may be quite misleading in practice (Hawkins, 1980). The problem is the so-called masking effect discussed in the preceding section, namely that the presence of two or more outliers may produce an insignificant result in the initial single outlier test. In view of this defect, the following variant is recommended: remove the K most extreme values of the sample (absolute values of residuals from the sample mean of the successively reduced sample are used to rank the observations). If the current doubtful value x is not thrown away in this process, declare x as 'clean' and process the next most outlying doubtful value for the current date. Otherwise, apply the following 'backward selection' algorithm: starting with the $n-K$ 'clean' observations, test the most inlying of the K extreme values for compatibility with the clean observations by the statistic T^* at a nominal significance level α . If it is compatible, then include it with the clean observations and repeat the procedure with the next most outlying suspected value in the sample until the current doubtful value x is processed and declared as compatible. This sequence of tests is immediately stopped when an observation is rejected by the statistical test T^* since all the subsequent outlying raw monthly means, including x , are then incompatible with the clean observations. In this case, the 2° lat \times 2° long area mean value corresponding to x is rejected or flagged, and the next doubtful value for the current date is processed. Note, however, that the other rejected values in the sample are not set to missing at this stage.

It is fair to say that, while the backward consecutive algorithm described in 4(c) is immune to masking (providing that the actual number of outliers in the sample does not exceed the number of suspected values K in the test procedure), it provides important distributional difficulties associated with finding suitable fractiles λ_α if we require (as we do) an actual significance level α for each of the successive null hypotheses which are tested in the backward selection algorithm. A comprehensive discussion of this problem is given by Hawkins (1980), and we omit the details owing to the lack of space and the difficulty of the problem. Suffice to say here, that it is necessary to resort to simulation if we require exact fractiles, but that there is little error by approximating these fractiles, as outlined in the preceding section, excepted for small n , say $n < 15$. The latter solution was adopted in this study. Consequently, the sequence of tests used in the backward consecutive algorithm described in 4(c) may have actual significance levels in excess of 25 per cent of the specified nominal significance level α according to Hawkins (1980). We will try to correct this deficiency in a future version of our outlier detection procedure by carrying out the required simulations.

4. **EXAMPLE** The outlier detection algorithm has been applied to several ship data sets and various examples were presented during the workshop. In particular, an experiment was undertaken on a pre-COADS marine product with known systematic errors, in order to show the benefit of this type of procedure in the context of marine climatology.

An extensive description of the ship data set used in this experiment may be found in Terray (1994). Briefly, SST data are presented as raw monthly means in 2° lat \times 2° long boxes in a domain extending from 30° to 100° E longitude and from 30° S to 30° N latitude. The period of analysis extends from 1900 to 1986. Figure 1 documents the irregular space-time sampling associated with this gridded ship data set.

Many well-known deficiencies were observed in this data set before and around the Second World War (Terray, 1994). In addition, a suspicious warming

trend is apparent on the SST time series during 1954-1976, and it was anticipated that this trend may be linked to important changes in the origin of the 'source-decks' merged into this marine product, or to the presence of a large amount of erroneous ship reports that were not rejected during basic quality control of the ship reports. Suspect raw monthly means were mainly confined along the shipping routes from Madagascar to Sumatra and from Sumatra to the Northern Arabian Sea for the 1968-1974 period.

In view of this, the outlier detection algorithm of the preceding section has been applied to this SST gridded ship data set in a two-step procedure:

- First, the algorithm was applied to all the raw monthly SST fields with 15°C as a lower limit and 35°C as an upper limit to identify doubtful 2° lat × 2° long monthly means which must be tested by the algorithm. A nominal significance level of 0.05 was chosen for all the tests. This first step was only intended as a check on 'evident' outliers far away from the bulk of the data. In this first step, 481 raw monthly values were tested for all the monthly fields of 1900-1986 and, among them, 361 were identified as outliers by the statistical tests (this number includes isolated monthly mean values) and rejected.
- The second step is designed to remove outliers with respect to anomaly fields. For this purpose, the raw monthly mean SST fields were expressed as monthly anomaly fields by using a monthly climatology obtained from a weighted EOF analysis on COADS SST data (Terry, 1998). The outlier detection algorithm was applied to these anomaly fields with -3°C as a lower limit and 3°C as an upper limit. Again, a nominal significance level of 0.05 was chosen for all the tests. In this second step, 10 917 anomaly values were tested; among them, 2 826 were identified as outliers and the corresponding raw monthly mean values were rejected.

The 15°C-35°C limit in the first step and -3°C-3°C limit in the second step were chosen as a compromise between a good use of computer resources and the quality of the final product. Lower upper limits or higher lower limits in both steps of the algorithm give roughly the same final results, but are more expensive with respect to CPU time.

On average, five to ten grid squares have been removed for each date from the beginning of this century until the Second World War. During the 1940-1968 period, the number of outliers is quite low, being less than five for each date on average. Finally, recent decades have witnessed a substantial increase in outlier losses. The number of outlier rejections may be as high as one hundred during 1968-1974 and outliers are still very common after this period. It is interesting to note that, except for the 1968-1974 period, outlier losses are very similar to the trimming losses observed for computing COADS trimmed monthly summaries for

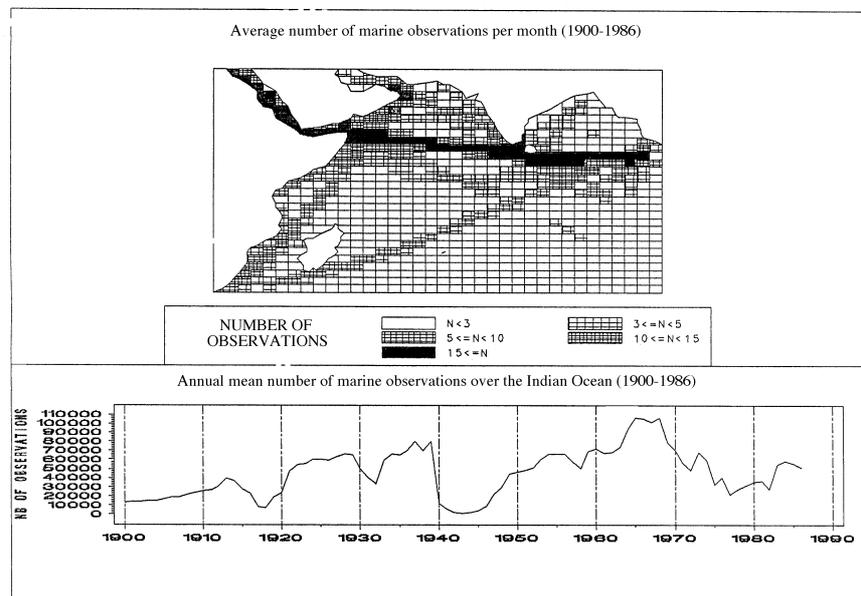


Figure 1—Space-time sampling of the ship reports associated with the gridded data set.

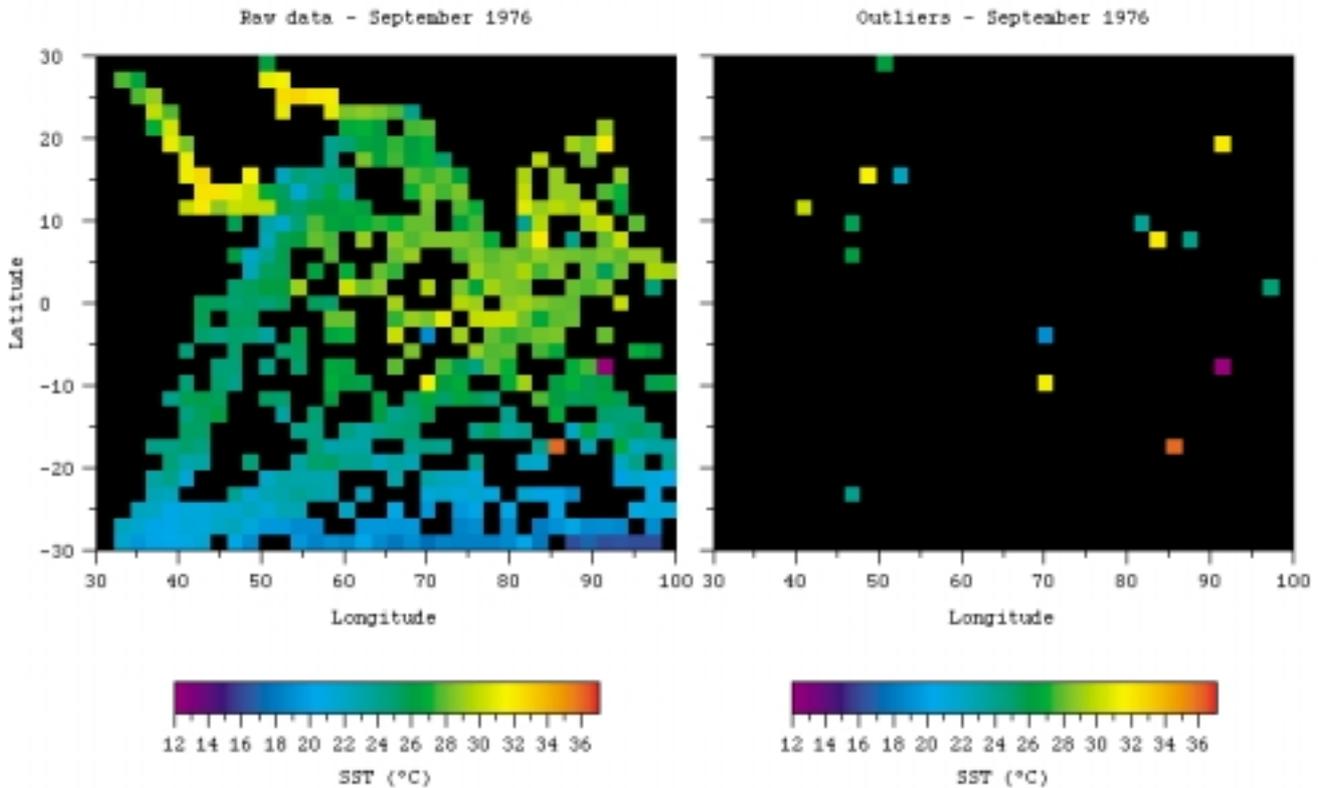


Figure 2—Results of the two-pass outlier detection algorithm for September 1976.

the global ocean (Wolter, 1997). Figure 2 presents the results of the two-pass outlier detection algorithm for September 1976 and may be used to obtain some understanding of the grid squares which were rejected by the outlier detection process.

To investigate the impacts of the outlier detection algorithm, the following computations were also undertaken on the SST ship data set both before and after the ‘cleaning’ of the data:

- (i) First, the 1954-1976 interval was used as a reference period for calculating a climatology for each calendar month and each 2° box, provided that data for at least 10 years with more than 5 observations per month were available in the period. The monthly means for each *i* grid point and *j* month were computed as a weighted average:

$$\bar{X}_{ij} = \left(\sum_{k=1954}^{1976} W_{ijk} X_{ijk} \right) / \left(\sum_{k=1954}^{1976} W_{ijk} \right)$$

where $W_{ijk} = 1 - \exp(-N_{ijk}/5)$

Here X_{ijk} is the value computed for the *i*th box, *j*th month and *k*th year. N_{ijk} is the number of ship observations used in computing X_{ijk} . W_{ijk} is in the neighbourhood of 1 if $N_{ijk} > 10$ and near 0.5 if N_{ijk} equals 5.

- (ii) After this first step, time monthly anomaly series for each 2° box during the 1900-1986 period were computed by simply subtracting this climatology from each value, provided that neither the datum nor the climatology was missing. These anomalies were then subsequently spatially averaged over the whole Indian Ocean with the same weighting scheme (e.g., W_{ijk}) as used in the computation of the climatology.

The two SST anomaly time series computed, respectively, before and after the ‘cleaning’ of the data, were then subjected to the X11 monthly additive scheme (Terray, 1994), a powerful technique for describing a time series, to assess their consistency. In the X11 procedure, the analysed X_t monthly time series is decomposed into three terms:

$$X_t = T_t + A_t + I_t$$

The T_t term is used to quantify the trend and low-frequency variations in the time series. The A_t term describes the annual cycle and I_t can be used to assess the level of noise in the data, though this term can also contain some signal in a climatological sense. All the terms are estimated with specific moving averages of various lengths.

Figures 3 and 4 give the results of the analysis for the SST time series computed before and after outliers were rejected, respectively. The monthly number of observations is also plotted on the bottom of each figure as an aid for interpreting the results and detecting accurately any change in the composition of the 'source-decks' contributing to the time series. While the two series and their associated X11 components are similar in many aspects, an important discrepancy may be noted during 1968-1974: the unlikely warm anomalies observed in the data before running the outlier detection procedure (Figure 3) are considerably reduced on the time series computed after outliers were rejected (Figure 4). As a consequence, the trend components of the two series are different during 1968-1974. This difference is consistent with the hypothesis of the artificial nature of the warming trend observed during 1968-1974 over the Indian Ocean. Finally, it may be noted that the 'clean' series is less noisy, as demonstrated by the irregular components.

Figure 3—SST monthly anomalies relative to 1954-1976 for the whole Indian Ocean before outlier detection. The series has been broken down into annual, trend and irregular components by the X11 procedure. The monthly number of ship reports used to construct the series is given at the bottom of the figure.

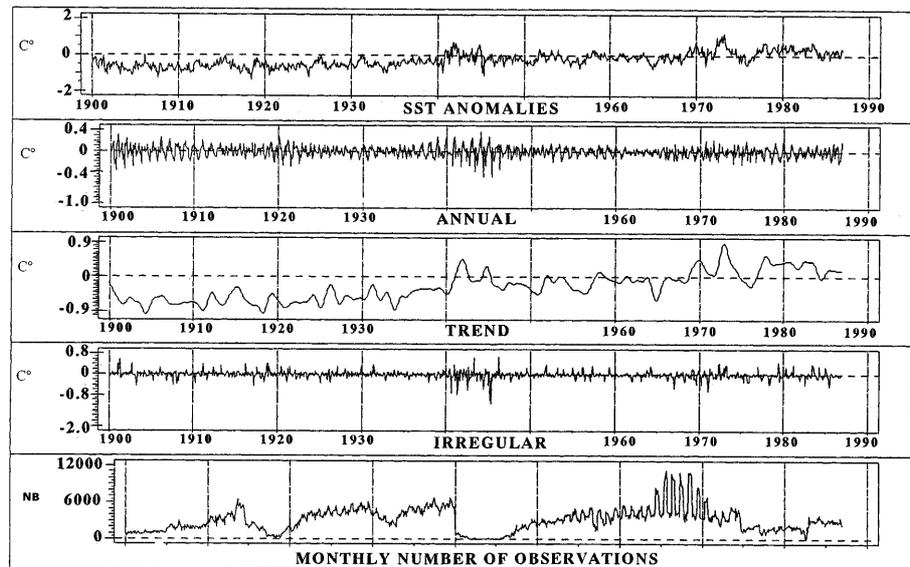
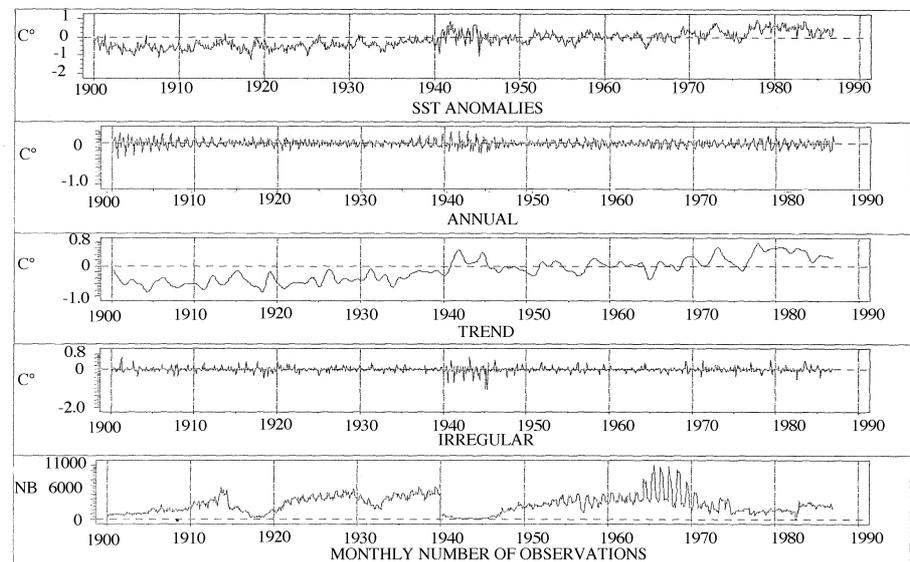


Figure 4—SST monthly anomalies relative to 1954-1976 for the whole Indian Ocean after outlier detection. The series has been broken down into annual, trend and irregular components by the X11 procedure. The monthly number of ship reports used to construct the series is given at the bottom of the figure.



5.
CONCLUSIONS

A recurring problem in the creation and maintenance of large gridded ship data sets is the accuracy of the information entering these products. The fact that large volumes of data are involved suggests that, as far as possible, the reliability of such data sets should be assessed through a computerized screening procedure. For this purpose, a new method for detecting outliers in gridded ship data sets has been proposed. It is our hope that this approach will aid climate scientists in determining which, if any, of the raw monthly area values included in a particular ship data set may be outliers.

Once potential outliers have been identified, it is suggested that these values may be flagged or, more drastically, rejected. In any case, the impact of these doubtful values in a particular data analysis may be easily assessed by comparing the results obtained before and after these potential outliers are rejected. In this way, it may be possible to obtain more reliable results in marine climatology.

The proposed approach may also be considered as a valuable alternative to trimming procedures which are applied to ship reports before computing monthly mean summaries for 2° lat \times 2° long boxes in order to reduce erroneous data losses.

REFERENCES

- Barnett, V. and T. Lewis, 1978: *Outliers in Statistical Data*. John Wiley & Sons, Inc., New York.
- Doornbos, R., 1966: *Slippage Tests*. 1st ed. Mathematical Centre Tracts, No. 15, Mathematisch Centrum, Amsterdam.
- Feller, W., 1968: *An Introduction to Probability Theory and its Applications*. vol. 1, 3rd ed., John Wiley & Sons, Inc., New York.
- Grubbs, F.E., 1950 : Sample criteria for testing outlying observations. *Ann. Math. Statist.*, 21, 27-58.
- Grubbs, F.E. and G. Beck, 1972: Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics*, 14, 847-854.
- Hawkins, D.M., 1980: *Identification of Outliers*. Chapman and Hall, London.
- Terray, P., 1994: An evaluation of climatological data in the Indian Ocean area. *J. Meteor. Soc. Japan*. 72, 359-386.
- Terray, P., 1999: Detecting climatic signals from ship's datasets. *Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata* (15-17 September 1997, Toledo, Spain). H.F. Diaz and S.D. Woodruff Eds., WMO/TD-No. 957.
- Wolter, K., 1997: Trimming problems and remedies in COADS. *J. Climate*, 10, 1980-1997.
- Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A Comprehensive Ocean-Atmosphere Data Set. *Bull. Amer. Meteor. Soc.*, 68, 1239-1250.

A METHODOLOGY FOR INTEGRATING WAVE DATA FROM DIFFERENT SOURCES PERMITTING A MULTISCALE DESCRIPTION OF WAVE CLIMATE VARIABILITY

G.A. Athanassoulis, Ch.N. Stefanakos¹; S.F. Barstow²

1. INTRODUCTION

Long-term wave data (e.g. spectral parameters) present themselves as time series of data, exhibiting random variability, serial correlation, seasonal periodicity and, possibly, a long-term climatic trend. The first three features are obvious in any long-term time series of wave data; see, for example, Figure 1, where a nine-year long time series of three-hourly sampled significant wave height H_s is shown. The long-term trend is a disputable character that may be seen after a careful statistical analysis of the annual mean values of H_s (see, for example, Athanassoulis and Stefanakos, 1995; WASA Group, 1998; Carter, 1999 and references cited there). In any case, the availability of multi-year long time series of data is a prerequisite for investigating the presence of a multi-year trend. In the present work, we shall disregard this question, since the data we have at our disposal are not long enough to resolve this feature.

To obtain a time series like the one shown in Figure 1 (nine years in situ measurements) is a time-consuming and expensive procedure. Nevertheless, this kind of data is necessary for a number of important applications such as, for example,

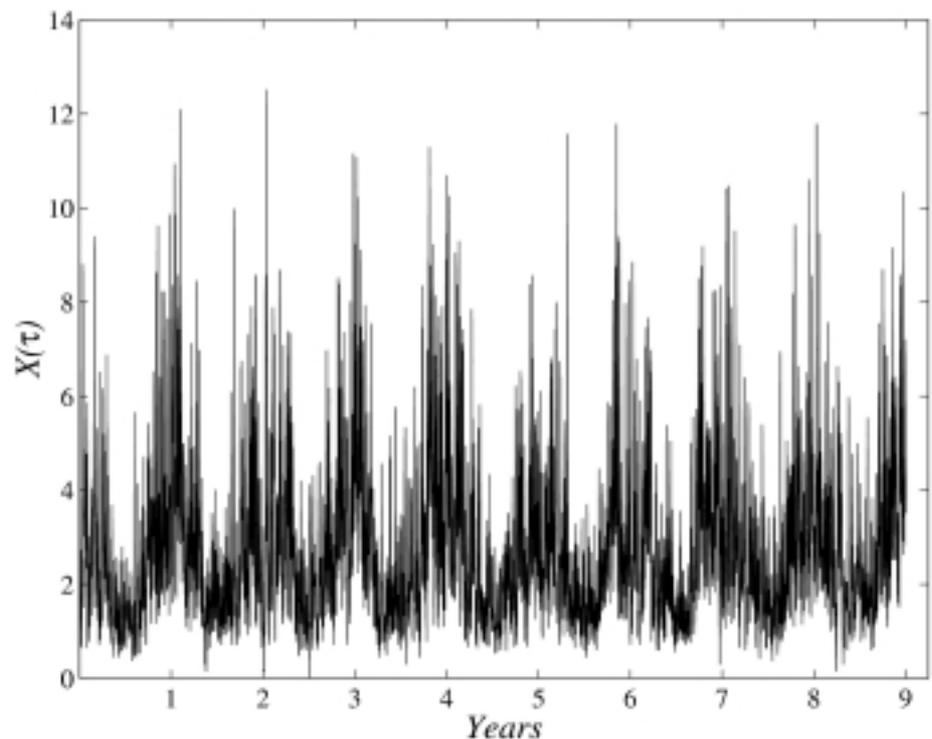


Figure 1 — Nine-year (1980–1989) long time series of significant wave height H_s from Haltenbanken, Norway, (65.08°N, 7.57°E) [buoy measurements].

1 National Technical University of Athens, Dept. of Naval Architecture and Marine Engineering, P.O. Box 64033, 15710, Zografos, Athens, Greece; e-mail: mathan@central.ntua.gr; chstef@central.ntua.gr

2 OCEANOR, Oceanographic Company of Norway, Pir-Senteret, N-7462, Trondheim, Norway; e-mail: sbarstow@oceanor.no

coastal morphodynamics (sediment transport and beach erosion) and direct numerical simulation of nonlinear long-term responses of offshore structures. Satellite altimeter data, that are being continuously collected, cannot substitute for the time series data in this kind of application, since the satellite footprint is moving across the ocean. Thus, the question arises of whether it is possible to estimate the different features of long-term time series of H_s , combining data from different sources, e.g. satellite altimeter data in conjunction with a restricted amount of buoy measurements.

The above question can be treated by exploiting an appropriate modelling of the time series (Athanasoulis and Stefanakos, 1995), which makes it possible to distinguish the different time scales (features) and, eventually, to associate the random variability and the correlation structure (hourly scale) with the buoy measurements, and the seasonal periodicity with the satellite measurements. In this way, the first two features can be estimated by means of a restricted amount of buoy data (say, one year), after a deseasonalization of these data by means of monthly mean values and monthly standard deviations obtained from several years of satellite altimeter measurements.

By further exploiting the time series modelling, it is possible to derive a many-year long time series by simulation, which combines all the basic statistical structure of the wave data. The whole methodology can be considered as an efficient way of blending (integrating) already available satellite data with a short (thus affordable and feasible) period of in situ measurements, to obtain an artefact of a long-term measured time series.

The structure of this paper is as follows. In section 2, the underlying stochastic modelling is presented and reformulated in accordance with the needs of the present work. In section 3, various statistics are introduced, defined by means of a time series of in situ measurements, which will be found to be comparable with appropriate statistics based on satellite measurements. The structure of satellite measurements that can be associated with a given site is discussed in section 4, where appropriate statistics are also defined. Systematic comparisons of the various statistics based on the two data sources, revealing which ones are interchangeable, are presented in section 5. After this assessment, the whole methodology for integration of satellite and in situ measurements leading to the construction of a simulated long-term time series is recapitulated in section 6. A general discussion and some conclusions concerning the extent of applicability and the necessary precautions in using the present approach are presented in section 7.

2. MODELLING AND ANALYSIS OF α -HOURLY LONG-TERM TIME SERIES

¹ We use the terminology ' α -hourly' (time series) in order to denote any time series of measurements with time step $\Delta\tau = \alpha$ hours. Usually, spectra or spectral parameters are recorded (or calculated) every 1, 3, 6 or 12 hours, thus $\alpha = 1, 3, 6, 12$. However, any value $1 \leq \alpha \leq 12$ is possible.

² Let it be noted that the whole methodology presented herein can be equally well applied to the case where a climatic trend $\bar{X}_{tr}(\tau)$ is present, if the data (from the same or other sources) allow us to identify such a trend.

Let us denote by $X(\tau_i)$, $i = 1, 2, \dots, I$, the α -hourly¹ many-year long time series of significant wave height $H_s(\tau)$ or an appropriate transform thereof. Usually, the shifted logarithms of $H_s(\tau)$ are considered, i.e. $X(\tau) = \log[H_s(\tau) + c]$, where c is a small positive constant between 0.2 m and 1 m. The constant $X(\tau)$ is introduced to avoid zeros and minimize the skewness of the probability distribution of $X(\tau)$. The log-transformed data are often approximately Gaussian, which greatly facilitates the analysis and the simulation procedure. According to the modelling introduced by Athanasoulis and Stefanakos (1995, 1998), such a time series $X(\tau)$ allows the following decomposition (see also Stefanakos, 1999):

$$X(\tau) = \bar{X}_{tr}(\tau) + \mu(\tau) + \sigma(\tau) W(\tau) \quad (1)$$

where $\bar{X}_{tr}(\tau)$ is any possible long-term (climatic) trend, $\mu(\tau)$ and $\sigma(\tau)$ are deterministic periodic functions with a period of one year, and $W(\tau)$ is a zero-mean, stationary, stochastic process. The functions $\mu(\tau)$ and $\sigma(\tau)$ are called seasonal mean value and seasonal standard deviation, respectively. In the sequel, we shall consider that $\bar{X}_{tr}(\tau) = \bar{X} = const$ and this constant will be incorporated into $\mu(\tau)$ ².

Thus, in the present work, decomposition (1) will be rewritten as:

$$X(\tau) = \mu(\tau) + \sigma(\tau) W(\tau) \quad (2)$$

The principal aim of the present work is to examine if, and how, it is possible:

- (i) To obtain reasonable estimates of $\mu(\tau)$ and $\sigma(\tau)$ by means of satellite data and, if the answer to this question is positive;
- (ii) To obtain the statistical characteristics of the residual process $W(\tau)$ by exploiting the satellite-based estimates of $\mu(\tau)$ and $\sigma(\tau)$, and a short-period (say, one year) of in situ measurements.

The methodology will be checked a posteriori by studying the stationarity of $W(\tau)$ and making a comparison with the corresponding results of direct α -hourly time-series analysis, in cases for which long-term in situ measurements are available.

The time series $X(\tau)$ is usually reindexed, in order to properly treat variability at different time scales, by using the double Buys-Ballot index (j, τ_k) , where j is the year index and τ_k ranges within the annual time (Athanasoulis and Stefanakos, 1995). In the present work, a triple index of similar philosophy is introduced, denoted by (j, m, τ_k) . The first component j is again the year index. The second component m is a month index, ranging through the set of integers $\{1, 2, \dots, M = 12\}$. The third component τ_k represents the monthly time, the index k ranging through the set of integers $\{1, 2, \dots, K_m\}$, where K_m is the number of α -hourly observations within the m -th month. Clearly, the meaning of the symbol τ_k in the triple index (j, m, τ_k) used herewith is different from the meaning of the same symbol in the double index (j, τ_k) used in previous studies (e.g. Athanasoulis and Stefanakos, 1995; 1998).

According to the new, three-index notation, the time series $X(\tau)$ is reindexed as follows:

$$\{X(j, m, \tau_k), \quad j = 1, 2, \dots, J, \quad m = 1, 2, \dots, M, \quad k = 1, 2, \dots, K_m\} \quad (3)$$

Note that there is a one-to-one correspondence between single index i and the triple index (j, m, τ_k) . Indeed, if the triple index (j, m, τ_k) is given, then the corresponding single index i is obtained by means of the relation:

$$i = (j-1)K + \sum_{q=1}^{m-1} K_q + k \quad (4)$$

where $K = \sum_{q=1}^M K_q$ is the total number of observations within a year. Conversely, if the single index i is given, then the corresponding triple index (j, m, τ_k) is calculated as follows:

$$\bullet \quad j = \left[\frac{i-1}{K} \right] + 1 \quad (5a)$$

• m is the unique integer for which:

$$\sum_{q=1}^{m-1} K_q + 1 \leq i - (j-1)K \leq \sum_{q=1}^m K_q \quad (5b)$$

and

$$\bullet \quad k = i - (j-1)K - \sum_{q=1}^{m-1} K_q \quad (5c)$$

The three indices j, m, τ_k represent three different timescales, making it possible to explicitly define statistics with respect to each one of them, separately. In the following sections, use will be made of the subscripts 1, 2, 3 to denote various statistics (mean value and standard deviation) with respect to the corresponding (first, second and third) index. To clarify the structure of this notation, we present a number of examples, some of which will also be used in the sequel:

$$M_1(m, \tau_k) = \frac{1}{J} \sum_{j=1}^J X(j, m, \tau_k) \quad (6a)$$

$$S_1(m, \tau_k) = \sqrt{\frac{1}{J} \sum_{j=1}^J [X(j, m, \tau_k) - M_1(m, \tau_k)]^2} \quad (6b)$$

$$M_3(j, m) = \frac{1}{K_m} \sum_{k=1}^{K_m} X(j, m, \tau_k) \quad (6c)$$

$$S_3(j, m) = \sqrt{\frac{1}{K_m} \sum_{k=1}^{K_m} [X(j, m, \tau_k) - M_3(j, m)]^2} \quad (6d)$$

We also define two-index statistics, which are obtained by successively taking mean values with respect to two indices. For example:

$$M_{13}(m) = \frac{1}{K_m} \sum_{k=1}^{K_m} \frac{1}{J} \sum_{j=1}^J X(j, m, \tau_k) = M_{31}(m) \quad (6e)$$

3. STATISTICS OF TIME SERIES OF MONTHLY MEAN VALUES

It is a straightforward matter to define the time series of monthly mean values (MMV) of $X(\tau_j)$. In fact, Equation (6c) defines this time series by averaging α -hourly observations over each month. In Figure 2, the MMV time series, obtained from the α -hourly time series shown in Figure 1, is presented (continuous line). By averaging $M_3(j, m)$ over all the examined years, we obtain the overall MMV (per month):

$$\tilde{M}_3(m) = \frac{1}{J} \sum_{j=1}^J M_3(j, m) = M_{31}(m) = M_{13}(m) \quad (7a)$$

The time series of monthly standard deviations (MSD) of $X(\tau_j)$ is defined by means of the Equation (6d). See also Figure 2 (dashed line). Averaging $S_3(j, m)$ over all the examined years, we obtain the overall MSD (per month):

$$\tilde{S}_3(m) = \frac{1}{J} \sum_{j=1}^J S_3(j, m) \quad (7b)$$

It should be noted that $\tilde{S}_3(m)$ is not the standard deviation of the time series $M_3(j, m)$. The selection of $\tilde{S}_3(m)$ as the representative quantity for the variability of

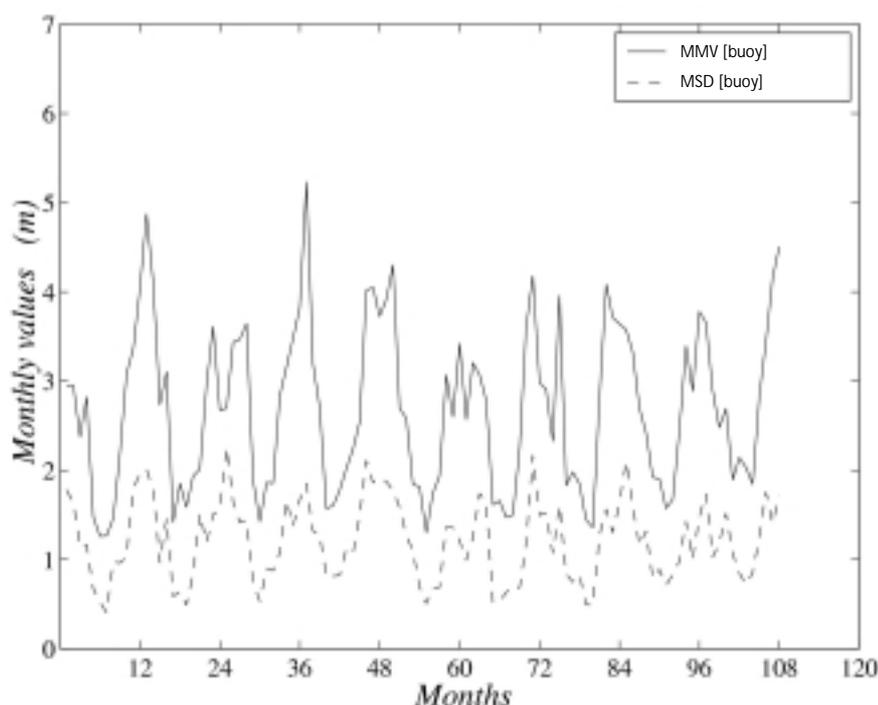


Figure 2—Time series of monthly mean values (MMV) and monthly standard deviations (MSD) from the Haltenbanken buoy measurements.

MMV $M_3(j, m)$ about the overall MMV $\tilde{M}_3(m)$ has been dictated by the data analysis. Indeed, after extensive numerical experimentation, it was found that it is exactly this quantity, i.e. $\tilde{S}_3(m)$, that can be related with (estimated by) an appropriately defined quantity obtained from satellite altimeter measurements.

4. STATISTICS OF SATELLITE MEASUREMENTS

Let us now turn our attention to satellite altimeter measurements of H_5 , obtained along specific (satellite dependent) ground tracks. Clearly, successive satellite observations are not referred to the same point in the sea. Thus, satellite wave data do not have the structure of a time series. If, however, we assume that the wave field is spatially homogeneous for an area S_A , surrounding a specific site of interest A^3 , then we can associate to this site all satellite observations within the area (Tournadre and Ezraty, 1990). This set of observations (population) can be given the structure of a three-index data set:

$$\{X^{sat}(j, m, \chi_\ell), \quad j = 1, 2, \dots, J, \quad m = 1, 2, \dots, M, \quad \ell = 1, 2, \dots, L_m\} \quad (8)$$

where again j is the year index, m is the month index, and X_ℓ is just a monthly counter, i.e. an index counting the number of observations within the area S_A , during the month m of the year j . Clearly, for given values of j and m , the individual values $X(j, m, \tau_k)$, $k=1, 2, \dots, K_m$, and $X^{sat}(j, m, X_\ell)$, $\ell=1, 2, \dots, L_m$ are not directly comparable.

³ The extent and shape of the area S_A are satellite dependent (sufficient data), and site dependent (local meteorological conditions).

Despite the structural differences between the data sets $X(j, m, \tau_k)$ and $X^{sat}(j, m, X_\ell)$, it can be expected that appropriate statistics of $X(j, m, \tau_k)$ can be approximated by analogous statistics of $X^{sat}(j, m, X_\ell)$, provided that the sea area S_A has been chosen appropriately. This expectation is based on the following assumptions concerning the time-space field of significant wave height $H_5(\tau, \vec{r})^4$:

- (i) $H_5(\tau, \vec{r} = \text{const})$ is (approximately) stationary within each month⁵;
- (ii) $H_5(\tau = \text{const}, \vec{r})$ is (approximately) homogeneous within the area S_A ; and
- (iii) a dispersion relation holds for the wave field $H_5(\tau, \vec{r})$ (Tournadre, 1993, Section 5.3).

Some results concerning the correspondence of temporal and spatial scales of $H_5(\tau, \vec{r})$ have been presented by Monaldo (1988, 1990), Tournadre (1993), and Krogstad and Barstow (1999).

⁴ Observations $X(j, m, \tau_k)$ and $X^{sat}(j, m, X_\ell)$ are considered as two different samples from the field $H_5(\tau, \vec{r})$.

The triple-index notation greatly facilitates the definition of various statistics on $X^{sat}(j, m, X_\ell)$, and the comparison with analogous statistics on $X(j, m, \tau_k)$. We present below some definitions of MMV and MSD related with $X^{sat}(j, m, X_\ell)$:

Definitions (9a,b) and (10a,b) correspond to (6c,d) and (7a,b), respectively.

⁵ Of course, in finer scales, short-duration energetic events (e.g. frontal passages) may occur that do not comply with the stationarity assumption. These events, which should be modelled using different (finer scale) stochastic processes, will not be considered herewith.

$$M_3^{sat}(j, m) = \frac{1}{L_m} \sum_{\ell=1}^{L_m} X^{sat}(j, m, \chi_\ell) \quad (9a)$$

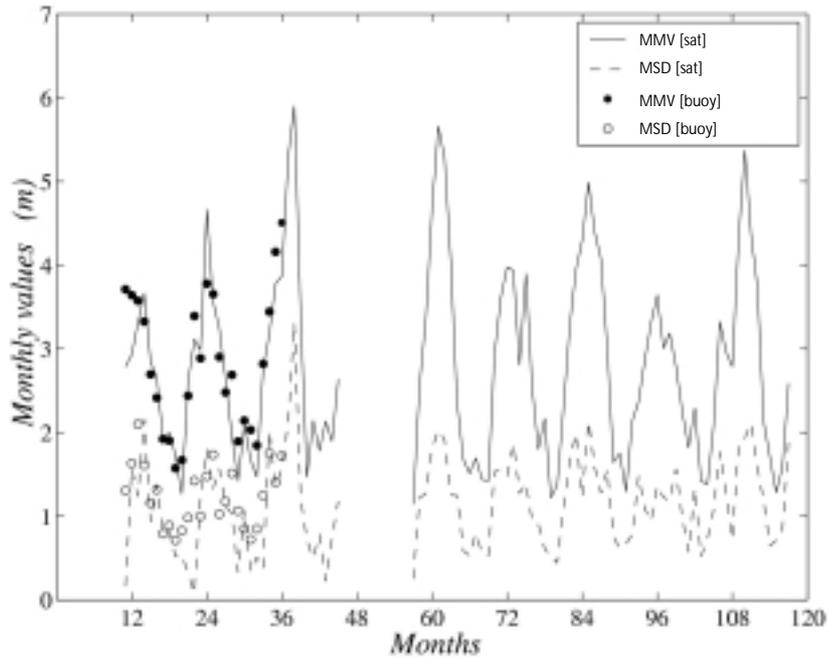
$$S_3^{sat}(j, m) = \sqrt{\frac{1}{L_m} \sum_{\ell=1}^{L_m} [X^{sat}(j, m, \chi_\ell) - M_3^{sat}(j, m)]^2} \quad (9b)$$

$$\tilde{M}_3^{sat}(m) = \frac{1}{J} \sum_{j=1}^J M_3^{sat}(j, m) \quad (10a)$$

$$\tilde{S}_3^{sat}(m) = \frac{1}{J} \sum_{j=1}^J S_3^{sat}(j, m) \quad (10b)$$

Clearly $M_3^{sat}(j, m)$ and $S_3^{sat}(j, m)$ are monthly time series generated by spatial averaging over the area S_A . In Figure 3, these time series, calculated from Geosat (1986-1989) and Topex (1992-1997) altimeter data, for an area near the Haltenbanken site, are shown. For the first two years, buoy measurements are also available, and they are depicted in the same figure by circles. As can be seen, the agreement between satellite monthly values and buoy monthly values is satisfactory. See also Figure 4, where results of the same analysis are shown for another site (NOAA 41001, 34.68°N, 72.64°W) in the western part of the North Atlantic Ocean. Note that this buoy is one of several buoys used in deriving the calibration

Figure 3—Time series of monthly mean values (continuous line) and monthly standard deviations (dashed line) from satellite measurements near Haltenbanken buoy (circles).



procedure applied to satellite altimeter data sets (see, for example, Krogstad and Barstow, 1999). Note also that the period of measurements for this NOAA buoy is 16 years (1982-1997), completely covering the period of the satellite measurements, which is the same as in the previous case. It seems, thus, reasonable to consider $M_3^{\text{sat}}(j, m)$ and $S_3^{\text{sat}}(j, m)$ as substitutes for $M_3(j, m)$ and $S_3(j, m)$. However, for the needs of our study, only the weaker assumption that the monthly time series $M_3^{\text{sat}}(j, m)$ and $S_3^{\text{sat}}(j, m)$ are statistically equivalent to the monthly time series $M_3(j, m)$ and $S_3(j, m)$, respectively, is necessary. On the basis of the above discussion, we can expect that $\tilde{M}_3^{\text{sat}}(m) \approx \tilde{M}_3(m)$ and $\tilde{S}_3^{\text{sat}}(m) \approx \tilde{S}_3(m)$.

5. COMPARISON OF MONTHLY STATISTICS FROM BUOY AND SATELLITE DATA

In this section, the statistical equivalence of MMV $\tilde{M}_3^{\text{sat}}(m)$ and $\tilde{M}_3(m)$, and MSD $\tilde{S}_3^{\text{sat}}(m)$ and $\tilde{S}_3(m)$ is established using the two aforementioned data sets (Haltenbanken and NOAA 41001).

In Figure 5, the monthly mean values from Haltenbanken are shown and compared. The nine values $\{M_3(j, m), j=1,2,\dots,9\}$ for each month, obtained from buoy data, are depicted as filled circles, while the eight values $\{M_3^{\text{sat}}(m), j=1,2,\dots,8\}$ for each month, obtained from satellite data, are depicted as open-faced stars.

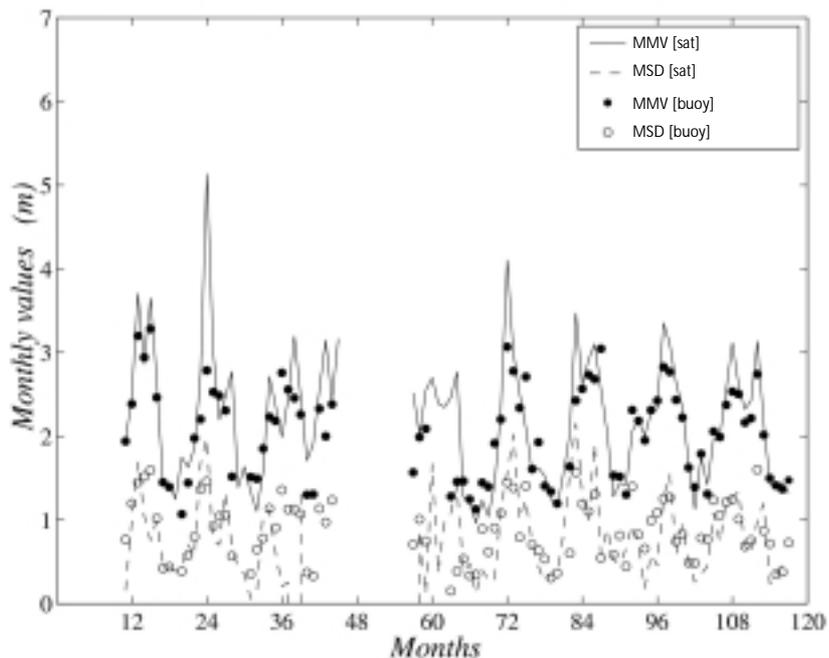


Figure 4—Time series of monthly mean values (continuous line) and monthly standard deviations (dashed line) from satellite measurements near NOAA 41001 buoy (circles).

These two sets of values (per month) are not comparable one-to-one since the periods of satellite and buoy measurements only partially overlap. However, based on the assumption of statistical periodicity of the wave climate, we expect that their mean values $\tilde{M}_3(m)$ and $\tilde{M}_3^{\text{sat}}(m)$ will be approximately the same, being two estimators of the same quantity. These two quantities are shown in Figure 5 by a solid and a dashed line, respectively. The agreement between $\tilde{M}_3^{\text{sat}}(m)$ and $\tilde{M}_3^{\text{sat}}(m)$ is, in general, impressively good, except for January and February. This discrepancy may (at least partly) be explained as follows:

After the end of the buoy measurement campaign in 1988, the winter wave climate in the Haltenbanken area, particularly during the months January to March, and for several years, had been surprisingly severe in comparison with the statistics of 1980-88 (buoy measurements period). This phenomenon has been confirmed by Barstow and Krogstad (1993) and other subsequent studies, using satellite and additional proprietary buoy data from the same area. This fact may be an indication of a long-term trend or periodicity in the wave climate in this area, but the amount of data is not yet enough to enable a statistical justification of such behaviour.

Similar comparisons are presented in Figure 6 for the monthly standard deviations from Haltenbanken. The conclusions are, in general, the same, though the agreement between $\tilde{S}_3(m)$ and $\tilde{S}_3^{\text{sat}}(m)$ is even better.

In Figure 7, the quantities $\tilde{M}_3(m)$ and $\tilde{S}_3(m)$, as calculated from three different data sources, namely, buoy data, satellite data and model (WAM) data⁶, are shown. Again, the overall agreement is very good, with maximum discrepancy (less than 15 per cent) in the winter months.

A similar analysis has also been performed for the site of the NOAA buoy 41001. Figure 8 shows the overall monthly mean values $\tilde{M}_3(m)$ and $\tilde{M}_3^{\text{sat}}(m)$ (continuous line), and the overall monthly standard deviations $\tilde{S}_3(m)$ and $\tilde{S}_3^{\text{sat}}(m)$ (dashed line). The agreement between buoy and satellite results is even better in this case.

6 The model data used is a six-year long, six-hourly sampled time series of H_s (July '92-June '98) for the grid point (64.5°N, 7.5°E), obtained from the WAM model operating at ECMWF, Reading, UK.

6. A METHODOLOGY FOR INTEGRATION OF SATELLITE AND BUOY MEASUREMENTS

Having established that $\tilde{M}_3^{\text{sat}}(m)$ and $\tilde{S}_3^{\text{sat}}(m)$ are reasonable estimates of $\tilde{M}_3(m)$ and $\tilde{S}_3(m)$, respectively, we are in a position to obtain the estimates $\mu^{\text{sat}}(\tau)$ and $\sigma^{\text{sat}}(\tau)$ as smoothed periodic extensions of the discrete estimates $\tilde{M}_3^{\text{sat}}(m)$ and $\tilde{S}_3^{\text{sat}}(m)$. In this way, we can obtain the residual time series:

$$W^{\text{sat}}(\tau) = \frac{X(\tau) - \mu^{\text{sat}}(\tau)}{\sigma^{\text{sat}}(\tau)} \quad (11)$$

and compare it with the corresponding $W(\tau)$ obtained by using the seasonal patterns $\mu(\tau)$ and $\sigma(\tau)$ estimated from the buoy data.

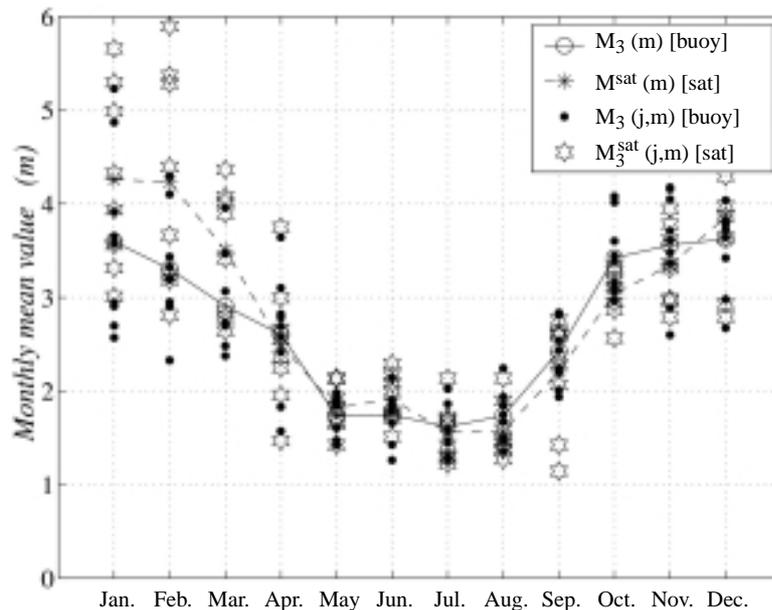


Figure 5—Comparison of monthly standard deviations of H_s from buoy (continuous line) and satellite (dashed line) measurements for the Haltenbanken site.

Figure 6—Comparison of monthly standard deviations of H_s from buoy (continuous line) and satellite (dashed line) measurements for the Haltenbanken site.

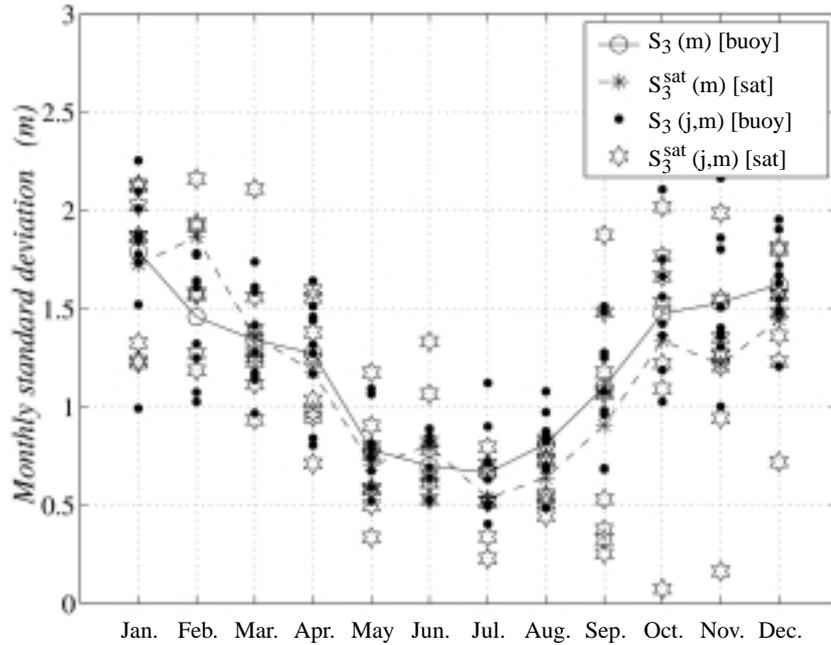
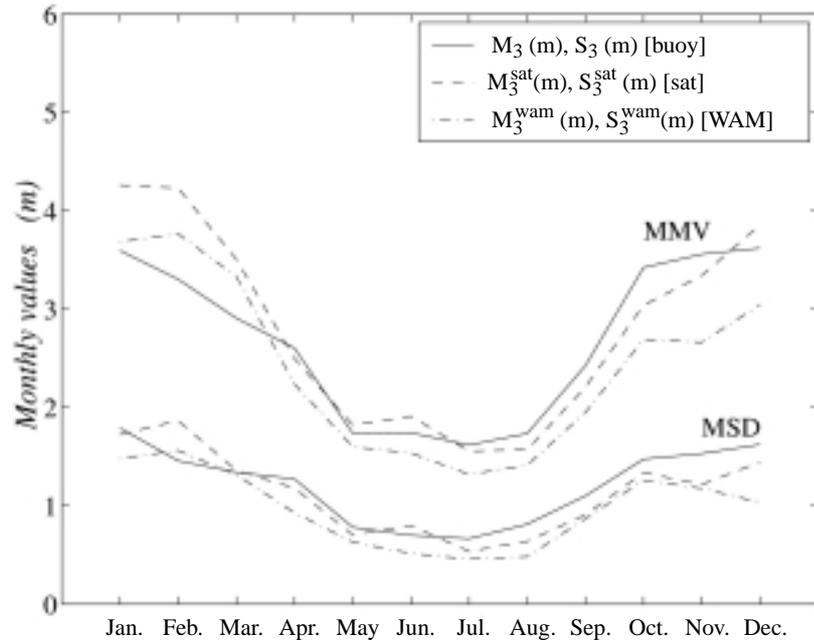


Figure 7—Comparison of monthly mean values and monthly standard deviations of H_s from buoy (continuous line) and satellite (dashed line) and hindcast (dashed-point line) data for Haltenbanken.



In order that the examined stochastic process is nearly Gaussian, we will work in this section with the time series $X(\tau) = \log [H_s(\tau) + c]$ with $c = 1m$. Clearly, the whole analysis procedure presented above using the original time series $H_s(\tau)$ must be repeated, using the log-transformed time series $X(\tau)$, in order to estimate the residual series $W^{sat}(\tau)$ and $W(\tau)$ of the transformed data.

Once the residual time series $W^{sat}(\tau)$ and $W(\tau)$ have been obtained, we can check their equivalence by comparing, for example, their spectral densities. Let us denote by $S_w^{sat}(f)$ the spectral density of $W^{sat}(\tau)$, and by $S_w(f)$ the corresponding one for $W(\tau)$. Figure 9 shows the two spectral densities $S_w^{sat}(f)$ and $S_w(f)$, calculated using various yearly segments of $W^{sat}(\tau)$ and $W(\tau)$, respectively. It can be seen that, even in the case of a one-yearly segment, there is good agreement between the various versions of the spectral density. This finding leads to the conclusion that the one-year buoy measurements, deseasonalized by means of the satellite seasonal patterns, describe well the state-to-state correlation structure of the α -hourly time series.

Figure 8—Comparison of monthly mean values and monthly standard deviations of H_s from buoy (continuous line) and satellite (dashed line) data for the NOAA buoy 41001 site.

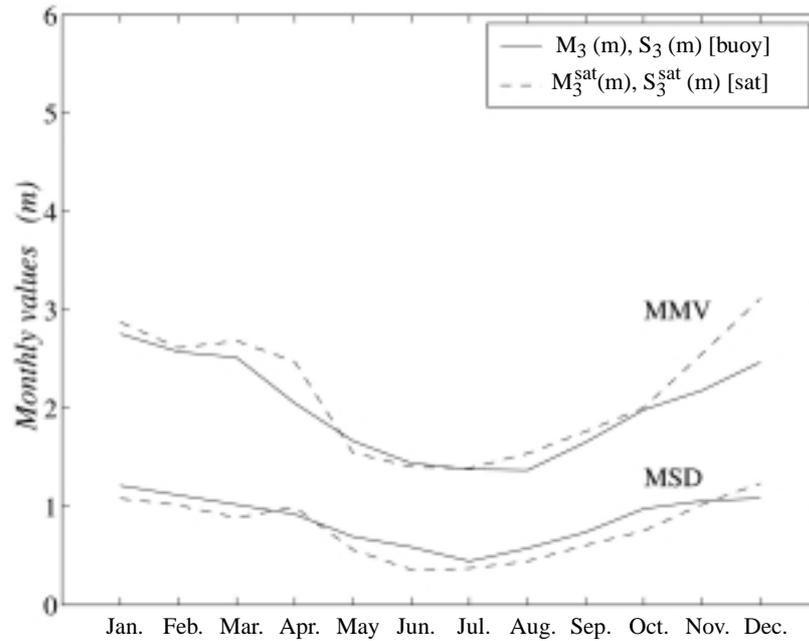
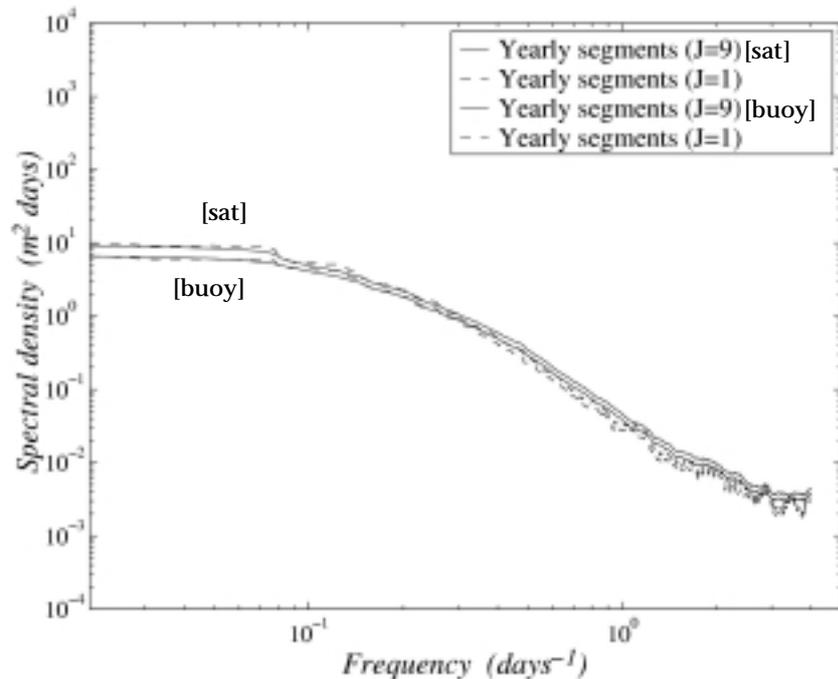


Figure 9—Spectral density functions $S_w^{\text{sat}}(f)$ and $S_w(f)$ calculated by analysing nine-year (continuous line) and one-year (dashed line) measurements. Site: Haltenbanken.

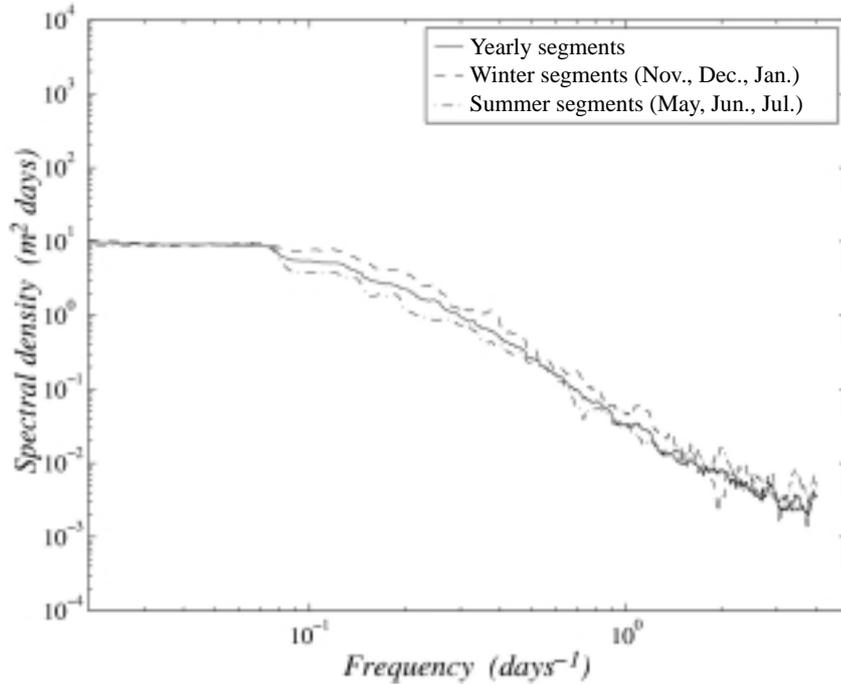


Next, the stationarity of $W^{\text{sat}}(\tau)$ is examined. In Figure 10, the spectral density $S_w^{\text{sat}}(f)$ is shown, estimated using various seasonal segments (yearly, winter and summer). The three versions exhibit a remarkable seasonal independence, confirming that the residual part $W^{\text{sat}}(\tau)$ can be modelled as a stationary stochastic process.

Furthermore, $W^{\text{sat}}(\tau)$ is given the structure of an autoregressive moving average model of order (2,2) (ARMA (2,2) model), the coefficients of which are estimated by least-square fitting to the raw spectral density (see, for example, Priestley, 1981 or Spanos, 1983). It is interesting to note that the ARMA coefficients, estimated either by means of a one-year segment, or by means of all nine-yearly segments, are very similar.

The ARMA modelling can be further exploited for simulation purposes. By generating zero-mean uncorrelated normal variates and using the estimated ARMA coefficients, a family of realizations of the same stationary stochastic

Figure 10—Spectral density function $S_w^{\text{sat}}(f)$ calculated by analysing yearly (continuous line) winter (dashed line) and summer (dashed-point line) segments. Site: Haltenbanken.



process can be produced. Then, using Equation (2), we obtain realizations of the initial time series $X(\tau)$.

In Figures 11 and 12, $S_w^{\text{sat}}(f)$ of the initial $W^{\text{sat}}(\tau)$ is compared with $S_w^{\text{sat}}(f)$ of the new $W^{\text{sat}}(\tau)$, which has been produced by simulation. Their agreement is found to be very good.

7. DISCUSSION AND CONCLUSIONS

In connection with various engineering applications, the long-term time series of significant wave height H_s (as well as various other wave and environmental parameters) can be considered as random series spanning at least three well-separated time scales: the sea-state duration (of the order of some hours), the yearly period (resolving the mean seasonal pattern), and a multi-year long time scale associated with possible long-term climatic trends.

Systematic long-term measurements covering all these scales are very impractical (time-consuming and expensive) and thus very rare. In fact, the evolution patterns of an environmental quantity in such different time scales may

Figure 11—Comparison of the initial residual time series $W(\tau)$ and the corresponding simulated series. Site: Haltenbanken.

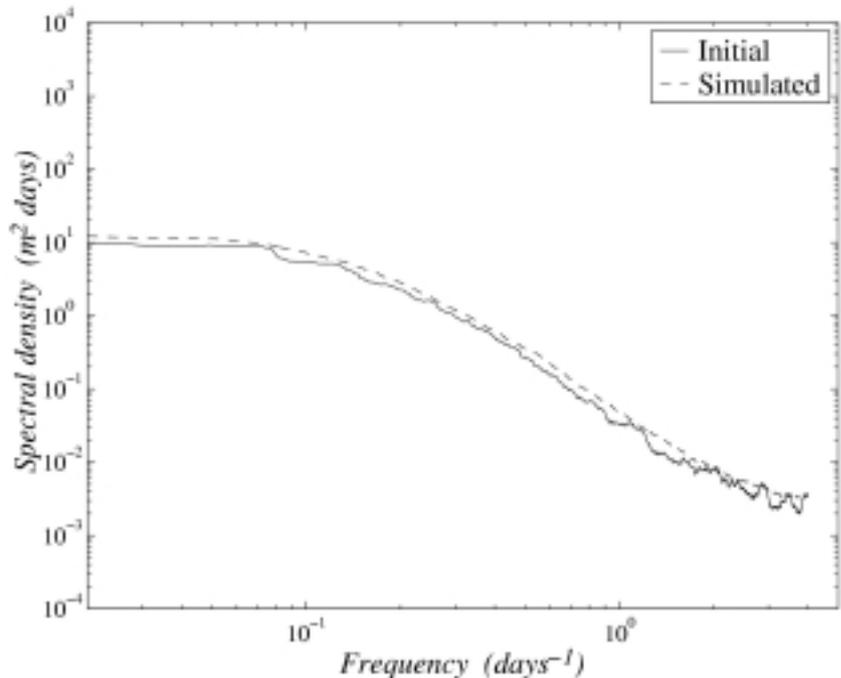
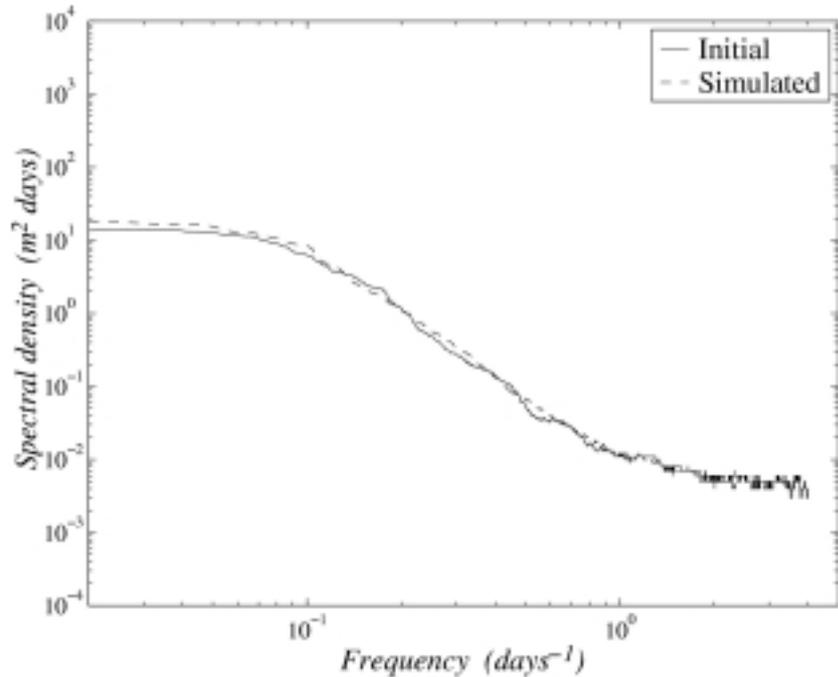


Figure 12—Comparison of the initial residual time series $W(\tau)$ and the corresponding simulated series. Site: NOAA buoy 41001.



be studied by different scientific or engineering disciplines, using entirely different types of devices and models.

Different kinds of measurements, from different sources, each one resolving a different scale, are available nowadays. In the present work we have established that a restricted period of buoy measurements can resolve the state-to-state correlation structure (a continuum of scales associated with various weather patterns), while an appropriate spatial averaging of satellite altimeter measurements can describe the mean seasonal pattern and the seasonal variability. The key point is that we can use the seasonal pattern, as obtained from appropriately defined satellite monthly values, in order to deseasonalize the buoy measurements. Then, by analysing the deseasonalized buoy measurements (which can be assumed to be a stationary stochastic process; see, for example, Athanassoulis and Stefanakos, 1995), we can estimate the correlation structure associated with the state-to-state scale by calculating the corresponding autocorrelation (or spectral density) function.

If we have at our disposal data (from any source) or even indications about a possible long-term trend, we can extend the scope of this methodology by using the decomposition (1). In this case, both the satellite and the buoy measurements first will be detrended by subtracting $\bar{X}_T(\tau)$, and then the detrended time series will be treated as previously.

In applying the models (1) or (2) (having determined their parameters using data from different sources), we should bear in mind that, in principle, they contain exactly those characters that have been resolved in the stage of the analysis procedure. For example, intermediate scale phenomena (e.g. energetic frontal passages) not complying with the constitutive assumptions of our model are not included therein. There are, however, various benefits in using a carefully estimated model like (1) or (2), instead of a unique measured sample. For example, the model is free from gaps (missing values), it enables the performance of sensitivity studies either by obtaining a population of realizations (by using various independent identically distributed (iid) samples of the generating random sequence) or by varying the parameters of the model, and, also, it gives us the ability to treat more complex problems by combining the present model with other ones.

Among various possible generalizations of the models (1) and (2) and their applications, the following two seem to be the most interesting ones. First, the generalization towards multivariate data, e.g. $\vec{X}(\tau) = (H_s(\tau), T_m(\tau))$ or $\vec{X}(\tau) = (H_s(\tau), U_{wind}(\tau))$, where T_m is the mean wave period and U_{wind} is the wind speed. The possibility of such an extension is also related to the quality and accuracy of the

satellite measurements for the additional parameters. Second, the generalization towards the inclusion of other phenomena evolving in different scales. For example, finer scale phenomena may be modelled by pulse-like processes (see, for example, Lopatoukhin *et al.*, 2000, Lopatoukhin *et al.*, 2001), while longer-scale phenomena might be included by introducing additional (longer) periods in the cyclostationary model (1).

ACKNOWLEDGEMENTS

The authors wish to thank the two anonymous referees for their constructive comments and suggestions.

This work has been partially supported by EC funds under contract MAS3-CT97-0109 (project EUROWAVES). The partners of the EUROWAVES project are: the Oceanographic Company of Norway (OCEANOR, coordinator), the National Technical University of Athens, Greece (NTUA) and the Instituto Studio Dinamica Grandi Masse, Italy (ISDGM).

The second author has also been supported by a postdoctoral fellowship from the Greek Foundation of State Scholarships.

REFERENCES

- Athanassoulis, G.A. and Ch.N. Stefanakos, 1995: A non-stationary stochastic model for long-term time series of significant wave height. *J. Geophys. Res.*, 100(C8), 16149 -16162.
- Athanassoulis, G.A. and Ch.N. Stefanakos, 1998: Missing-value completion of nonstationary time series of wave data. 17th International Conference on Offshore Mechanics and Arctic Engineering (OMAE) (Lisbon, Portugal).
- Barstow, S.F. and H.E. Krogstad, 1993: Analysis of extreme waves and recent storms in the Norwegian Sea. Climatic Trends and Future Offshore Design and Operation Criteria Workshop (Reykjavik, Iceland).
- Carter, D.J.T., 1999: Variability and trends in the wave climate of the North Atlantic: A review. 9th International Offshore and Polar Engineering Conference (ISOPE) (Brest, France).
- Krogstad, H.E. and S.F. Barstow, 1999: Satellite wave measurements for coastal engineering applications. *Coastal Engineering*, 37, 283-307.
- Lopatoukhin, L.J., I.V. Lavrenov, V.A. Rozhkov, A.V. Boukhanovsky, A.B. Degtyarev, 2000: Wind and wave climate of oil fields in the Barents, Pechora, and Kara Seas. 6th International Conference on Ships and Marine Structures in Cold Regions (ICETECH) (St. Petersburg, Russia, 475-482) (in Russian, English abstract).
- Lopatoukhin, L.J., V.A. Rozhkov, V.E. Ryabinin, V. Swail, A.V. Boukhanovsky, A.B. Degtyarev, 2001: *Estimation of Extreme Wind Wave Heights*. WMO/TD-No. 1041.
- Monaldo, F., 1988: Expected differences between buoy and radar altimeter estimates of wind speed and significant wave height and their implications on buoy-altimeter comparisons. *J. Geophys. Res.*, 93(C3), 2285-2302.
- Monaldo, F., 1990: Corrected spectra of wind speed and significant wave height. *J. Geophys. Res.*, 95(C3), 3399-3402.
- Priestley, M.B., 1981: *Spectral Analysis and Time Series*. Academic Press, 890 pp.
- Spanos, P-T.D., 1983: ARMA algorithms for ocean wave modeling. Trans. of the ASME, *Journal of Energy Resources Technology*, 105, 300-309.
- Stefanakos, Ch.N., 1999: Nonstationary stochastic modelling of time series with applications to environmental data. Ph.D. Thesis, National Technical University of Athens, Greece, 200 pp. (Supervisor: G.A. Athanassoulis).
- Tournadre, J., 1993: Time and space scales of significant wave heights. *J. Geophys. Res.*, 98(C3), 4727-4738.
- Tournadre, J. and R. Ezraty, 1990: Local climatology of wind and sea state by means of satellite radar altimeter measurements. *J. Geophys. Res.*, 95(C10), 18255-18268.
- The WASA Group, 1998: Changing waves and storms in the Northeastern Atlantic. *Bulletin of the American Meteorological Society*, 79(5), 741-760.

REDUCED SPACE APPROACH TO THE OPTIMAL ANALYSIS OF HISTORICAL MARINE OBSERVATIONS: ACCOMPLISHMENTS, DIFFICULTIES, AND PROSPECTS

A. Kaplan, M.A. Cane and Y. Kushnir, Lamont-Doherty Earth Observatory of
Columbia University Palisades, New York, USA

ABSTRACT Observed historical climate fields are characterized by comparatively precise data and good coverage in the last few decades, and by poor observational coverage prior to then. The technique of the reduced space optimal analysis of such fields (i.e. estimating them in projections onto a low-dimensional space spanned by the leading patterns of the signal variability) is presented in the context of more traditional approaches to data analysis. Advantages of the method are illustrated on examples of reconstructions of near-global monthly fields of sea surface temperature and sea level pressure from the 1850s to the present, along with verified error bars. The limitations of the technique as regards quality and robustness of estimating a priori parameters, representation of long-term and small-scale types of variability, assumption of stationarity of means and covariances, and incompleteness of coverage are discussed, and possible ways to overcome these problems are suggested.

1. INTRODUCTION Less than two centuries of observational records which have made their way from the hand-written ship logs into the modern data banks constitute the main source of our knowledge of the variability associated with the ocean-atmosphere interaction. For use in climate research, the ship measurements are customarily being compiled into monthly binned averages on regular longitude-latitude grids with quality control and other statistics (e.g. Comprehensive Ocean-Atmosphere Data Set (COADS) - Woodruff *et al.*, 1987; Global Ocean Surface Temperature Atlas (GOSTA) - Bottomley *et al.*, 1990). The resulting products still reflect the historical variations in the intensity of marine traffic, being incomplete at present, quite 'gappy' before the 1950s, and extremely sparse for the most of the 19th century. Satellite observations can complete the modern part of this record (almost two decades for sea surface temperature (SST), and much less for other climate variables), but cannot provide a record lengthy enough for the studies of decadal and longer time scale climate variability. As a result, in climate studies, one faces the necessity of using very incomplete data fields which are affected by observational and sampling errors. In contrast, the two main approaches to modern climate research, namely statistical techniques (like principal components analysis, singular vector decomposition, singular spectrum analysis, etc.) and model experiments (use of observed fields for boundary conditions), both expect gapless and error-free input data. Because of this, a great deal of attention has been paid in the last few decades to various methods of data analysis, and to those which are supposed to interpolate gaps and suppress data error.

A majority of existing approaches to interpolating historical data are drawn from the idea of minimization of least squares. It is well-known (by Gauss-Markov theorem; e.g. Mardia *et al.*, 1979; Rao, 1973) that the systematic use of this method makes it possible to produce an optimal estimate (an unbiased one with the smallest error among all linear estimates). However, this involves a few assumptions, including the knowledge of error covariances. In the absence of

this knowledge, some additional assumptions are usually made. Statistical techniques such as kriging (e.g. Cressie, 1991) or successive corrections (Daley, 1993) normally assume a ‘localized’ covariance structure and produce useful results if the gap size does not exceed the data decorrelation scale (e.g. Da Silva *et al.*, 1994; Levitus and Boyer, 1994).

A seemingly different approach to historical data analysis (often also called data reconstruction, to emphasize the scarcity of the input data), which is based on the use of empirical orthogonal functions (EOFs), has become quite popular in recent years (Shriver and O’Brien, 1995; Smith *et al.*, 1996; Rayner *et al.*, 1996; Mann *et al.*, 1998). In fact, this technique can be derived as a straightforward application of a classic least squares estimate with a special EOF-based reduced rank approximation of a signal (or model error) covariance matrix. In this venue, Kaplan *et al.* (1997) formulated reduced space analogues of the traditional technique of optimal analysis (optimal interpolation, Kalman filter, optimal smoother). The application of this technique to the historical data sets of SST and marine sea level pressure (SLP) resulted in near-global monthly analyses of these variables going back to more than 140 years, accompanied by the error bars (Kaplan *et al.* (1998, 2000)) which are publicly available. The assumptions underlying the method, namely the stationarity of the mean field and covariance of the signal, have been recently criticized (Hurrell and Trenberth, 1999). Additionally, the current settings of the analysis result in globally incomplete fields of comparatively sparse resolution ($4^\circ \times 5^\circ$ grid size) which limits considerably the utility of such analyses in climate model experiments. In section 2, we bring the reduced space optimal analysis into the context of more traditional objective data analyses and summarize its advantages and existing applications. Section 3 discusses the current difficulties in applications of the method and suggests ways of resolving them. Section 4 concludes the paper by emphasizing the prospects of the method and directions for further applications.

2. ACCOMPLISHMENTS

2.1 THEORETICAL BACKGROUND

The generic problem of the optimal analysis of time-evolving fields T_n (n is the time index) requires reconciliation of information coming from two sources: an imperfect model of time transitions A_n and incomplete and erratic observations T_n^o connected to the estimated field via a linear (or linearized) operator H_n same as Figure 1. Note that error of linearization (or interpolation) of the operator H_n is included in the effective observational error e^{obs} .

This problem is central for two areas of climate research which traditionally are considered separately: assimilation of data into numerical models and objective analyses (reconstructions) of data sets of historical observations. In fact, the main difference between these two types of problems is the relative amount of information brought by the model versus observations: it is high in the former problem and low in the latter. If model and observational errors in the equations

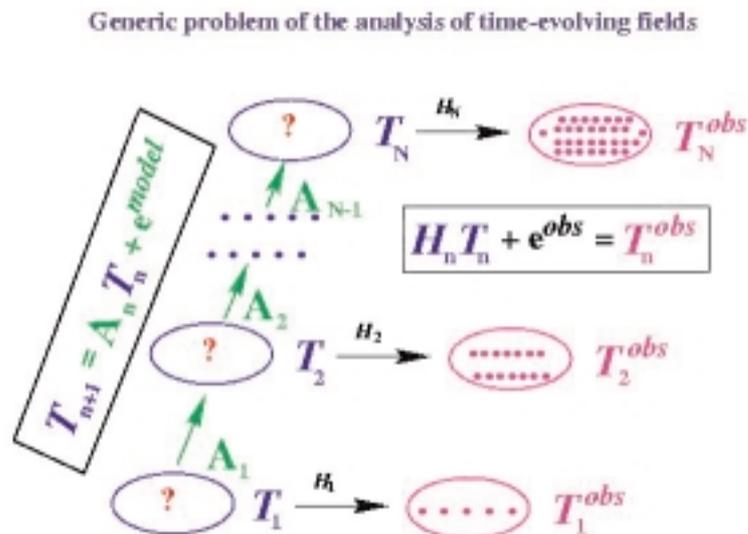


Figure 1—Generic scheme of the informational content for the analysis of time-evolving fields.

shown in Figure 1 are white in time, uncorrelated with each other and the estimated fields T_n , have zero mean and known spatial covariances Q_n and R_n respectively, we have a classic Gauss-Markov estimation problem for T_n , whose solution can be found as a minimizer of the quadratic cost function:

$$S[T_1, T_2, \dots, T_N] = \sum_{n=1}^N (H_n T_n - T_n^o)^T R_n^{-1} (H_n T_n - T_n^o) + \sum_{n=1}^{N-1} (T_{n+1} - A_n T_n)^T Q_n^{-1} (T_{n+1} - A_n T_n) \quad (1)$$

(for a detailed explanation of notation, terminology and basic facts of optimal estimation, readers are referred to Kaplan *et al.*, 1997). According to the Gauss-Markov theorem (e.g. Mardia *et al.*, 1979; Rao, 1973), this solution has minimum error variance among all linear estimates of T , and it is usually referred to as the ‘optimal’ solution. In fact, if additional assumptions on the Gaussian distribution of errors or the signal are made, the same solution receives an interpretation as the maximum likelihood estimate, or becomes the best solution among all, even nonlinear, estimates for a wide class of optimality criteria. Because the solution minimizes a quadratic cost function, it is often referred to as a ‘least-squares solution’.

There are well-known algorithms to find a minimizer of (1) in its complete form (fixed-interval optimal smoother (OS)), or somewhat truncated forms (fixed-lag optimal smoother, Kalman filter (KF)), or its simplification for a single-time estimation (optimal interpolation (OI)). They are supposed to give optimal solutions if assumptions on errors are satisfied, including the requirement that the covariance matrices of errors (Q and R) are known. However, in actual applications to the problems of climate research, the realistic dimensions of data are usually large enough to warrant two outcomes:

- (1) error covariance matrices are not known in all their details since there are not enough data to resolve them completely, so some crude parameterizations are used instead;
- (2) if no simplifications are carried out, optimal data analysis procedures are very expensive (OI), extremely expensive (KF), or prohibitively expensive (fixed-interval OS).

Both these difficulties, however, can be dealt with at once if certain features of optimal solutions of realistic climate fields are taken into account.

Consider as an example a standard OI problem whose solution is a minimizer T of the cost function:

$$S[T] = (HT - T^o)^T R^{-1} (HT - T^o) + (T - T^b)^T C^{-1} (T - T^b) \quad (2)$$

Here T^o is a (column-) vector of observations, T^b is a first guess (background) solution, H is a transfer matrix from a complete field to the set of observed points, R and C are covariances of observational and first guess errors, respectively. The two terms of the cost function S ‘punish’ the solution T for deviation from observations and from the background values.

The solution to this OI problem is:

$$T = P(H^T R^{-1} T^o + C^{-1} T^b)$$

where

$$P = (H^T R^{-1} H + C^{-1})^{-1}$$

is estimated covariance of its error.

Let us subtract the first guess solution from the estimated field, so that the new T is $T - T^b$ and new T^o is $T^o - HT^b$. If the first guess solution is a climatological field, then we have redefined the signal to be a field of anomalies. After such a change in definitions, the first guess solution equals zero, so that the first guess error equals the entire value of the signal T , and the matrix C becomes the covariance of the signal $\langle TT^T \rangle$. It can be expanded into its canonical representation:

$$C = E \Lambda E^T \quad (3)$$

E being a matrix of eigenvectors (EOFs if C is effectively a sample covariance estimate), and Λ is a diagonal matrix of eigenvalues. We can use eigenvector patterns to rotate an estimated field:

$$T = E\alpha \tag{4}$$

so that $\alpha = E^T T$ becomes a new unknown: a vector of projections of a target field on eigenvectors.

For simplicity, let us consider the case of a completely observed system ($H=I$, I being an identity matrix) with white uniform error ($R=rI$). The OI solution for such a system has a closed form for each component of α :

$$\alpha_i = \frac{\lambda_i}{\lambda_i + r} \alpha_i^o \tag{5}$$

($i=1\dots N$ is an index of components, eigenvalues and eigenvectors, $\alpha^o = E^T T^o$ is a vector of projections of the observed field T^o on eigenvectors). We assume that eigenvalues are arranged in descending order. The usual case then is that $\lambda_N \ll r \ll \lambda_1$. This means that $\alpha_1 \approx \alpha_1^o$ and $\alpha_N \approx 0$.

In other words, the standard least squares procedure of OI in its search for the optimal solution will damp the observed values of all eigenvector amplitudes whose energy in the signal does not dominate over the observational error. Eigenvector modes which are expected to have energy much below the level of observational error will not be represented in the OI solution. In the case of global SST anomaly fields, a realistic observational error level of 0.5°C (see Kaplan *et al.*, 1998 for details on the data set and its error model) results in the reduction by the factor of 2 or more of the variance in the modes beyond top 100 (Figure 2).

Consequently, computing the OI solution in all its details (projection to all EOFs) is superfluous: equally good results can be achieved by computing only projections on some set of leading eigenvectors. It should be noted that for many physical variables, the most energetic modes are those of the largest spatial scale. Details of the solution on small scales (projection to high number eigenvectors) is controlled by the fine details of the covariance matrix C which usually cannot be reliably estimated from the data. Large scale patterns of C (leading eigenvectors), however, can be estimated in a more reliable way. Approximation of C in (3) by only a few leading terms (truncation) results in infinite coefficients in the second term of the cost function (2) which totally disallow projection of the solution on truncated modes (in terms of the solution (5), if λ_i is assumed to be zero, then α_i will also be zero). The same result, of course, can be achieved by truncating the eigenvector representation of the solution (4) to begin with. We call such a

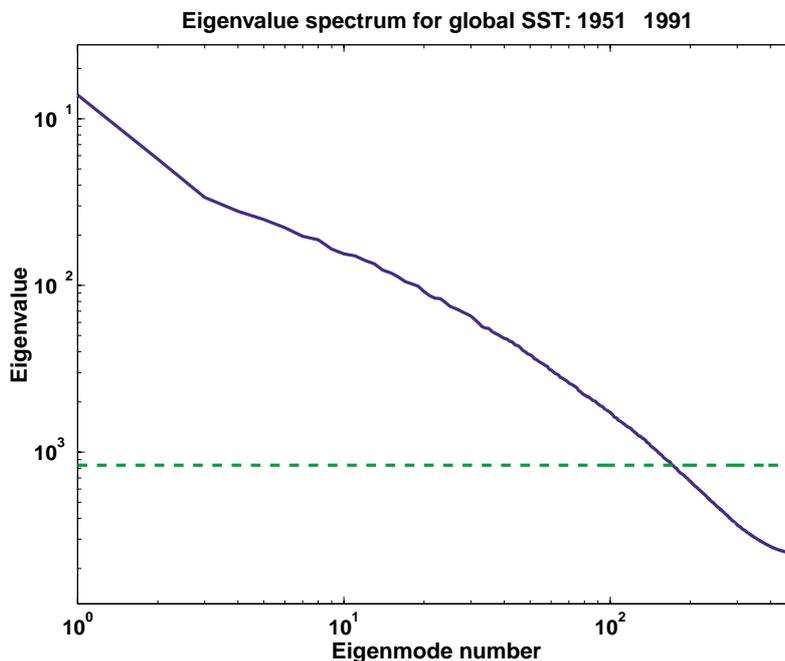


Figure 2—Normalized eigenvalue spectrum of global SST fields. Dashed line corresponds to the magnitude of the characteristic observational error in the historical data (standard deviation of 0.5°C): variance in the eigenmodes with the eigenvalues below this level are reduced by the factor of 2 or more in the least-squares optimal estimates.

truncation a reduced space representation of the solution; inserting the truncated form of (4) into cost functions followed by their minimization with regards to the low-dimensional vector allowed the development of the reduced space analogues of the OI, KF, and OS algorithms (Cane *et al.*, 1996; Kaplan *et al.*, 1997). If certain assumptions are held, these solutions are in fact projections of the ‘complete’ full-grid optimal solutions onto the low-dimensional reduced space.

Figure 3 emphasizes the contrast between the reduced space solutions and those obtained via the more traditional kriging (or successive corrections) approach. Both types of solutions are based on the least-squares, the difference being in the approximation used for the baseline error covariance. The reduced space approach uses the most effective type of low-rank covariance approximation: via its leading eigenvectors in its canonical expansion (Golub and Van Loan, 1996). For most climatic fields, this approximation will retain the part of the covariance with the longest spatial (and often temporal) scales, i.e. it corresponds to that part of the signal which we usually presume to be ‘climatic’. The residual of this representation will have predominantly short decorrelation scales and in fact will not be an effective representation of the true climatic covariance in any matrix norm. Yet it is being used in the standard applications of kriging and successive correction techniques for the sole reason that such ‘localized’ covariance structures are easy to model statistically.

While the reduced space solutions are formally suboptimal among full grid solutions, they are optimal among all reduced space solutions, being also far cheaper and much easier to feed by a priori error covariance information. For the settings which allow direct comparison, the solutions in the reduced space prove to be not inferior to the actually existing full grid solutions (Cane *et al.*, 1996). The reason for that is the poor representation (or inadequate parameterization) of small scales in full grid error covariance estimates. As a result, the full grid data analysis of small scales sometimes does more harm than good. Moreover, the analysis for those scales represents the major computational expense of the entire procedure. Hence, the savings of reduced space analysis occur at the scales which are not really constrained by the data. Estimation on such scales is often meaningless, but the traditional schemes cannot selectively cut off computation there. The tunable nature of the dimension of a reduced space makes it possible to put into the solution all scales down to the smallest resolved by available data, and the choice of leading EOFs for a basis that guarantees to some extent the minimal dimension of the analysis space.

When the covariance of a climatic variable is dominated by a few large-scale modes, the generic objective analysis with correctly estimated covariance structures will predominantly reconstruct the patterns manifested in the large-scale climate dynamics. This is true for both full-grid and reduced space analyses, the latter being particularly effective in such settings. When this is not the case, the results of covariance estimation and of the full-grid objective analysis applied to the sparsely observed data are likely to be less robust and more error-prone, with space reduction not being effective either.

APPROXIMATING COVARIANCE

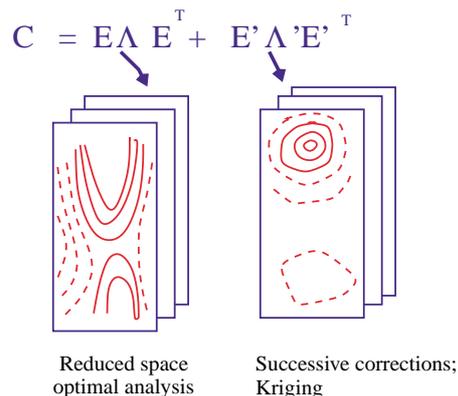


Figure 3—Separation of covariance into large- and small-scale portions in various optimal analysis techniques.

2.2 APPLICATIONS

We applied the reduced space OS to produce the near-global analysis of $5^\circ \times 5^\circ$ SST monthly anomaly grids for the 1856-1991 period (Kaplan *et al.*, 1998). For a model of time transitions we used an empirically fitted first order autoregressive model which was assumed to be diagonal in the reduced space coordinates. The observational data used in this work are known as the MOHSST5 compilation of ship observations produced by the UK Met Office (Bottomley *et al.*, 1990), Parker *et al.*, 1994). Covariance of the SST field was derived from the 1951-1991 period, then its leading 80 EOFs were used for the optimal estimation in the entire time range from 1856 to 1991.

Extensive tests proved the analysis to be robust and self-consistent. As Figure 4 illustrates, even under the sparse spatial coverage of December 1877 (known to be a strong warm ENSO event), the analysis produces a believable structure for a very strong El Niño known to have occurred that year (panels (a), and (b)). We verified the credibility of that reconstruction by taking data for December 1986 (a well sampled month, panels (c) and (d)), sampled them per the 1877 sampling pattern, and corrupted them by noise (to reflect the increase in the error at each grid box due to less frequent sampling). The OS analysis produced the 1986 El Niño pattern with only slightly weaker amplitude than that obtained with

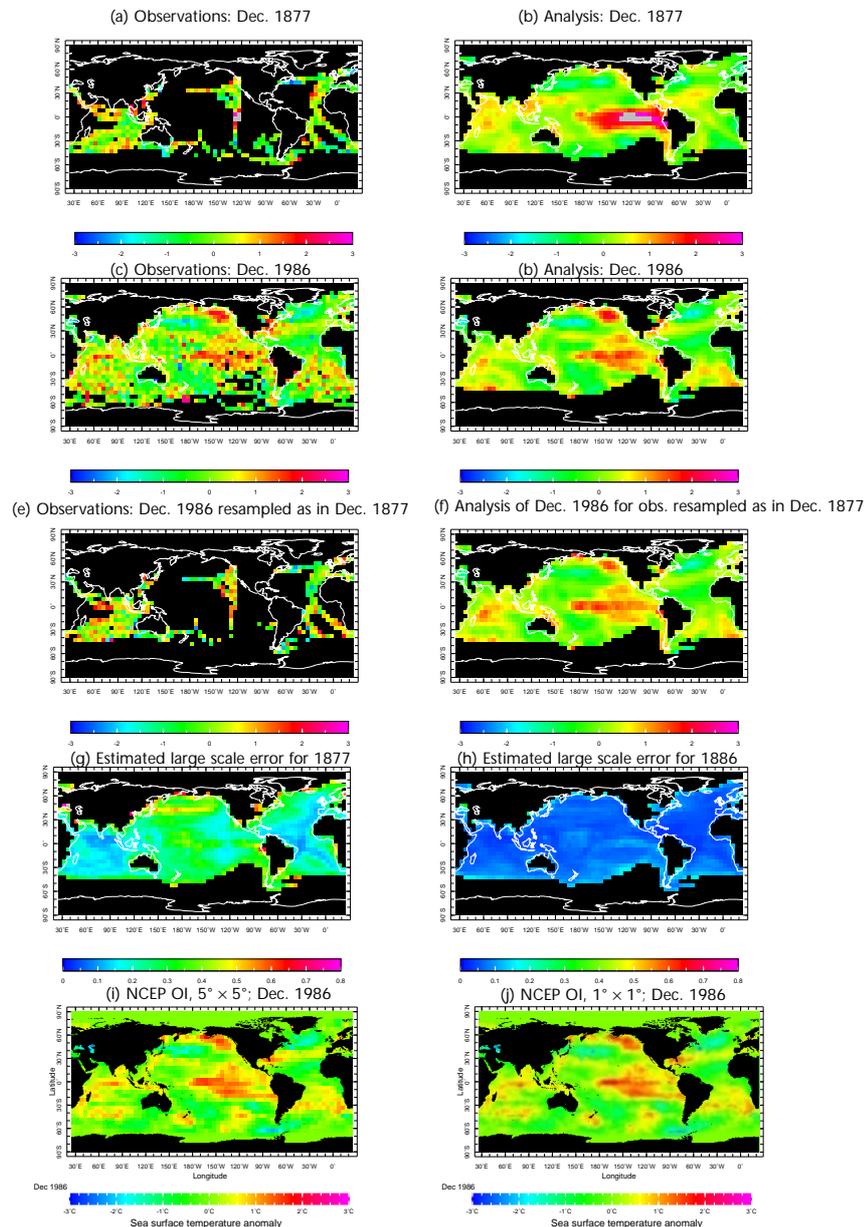


Figure 4—Available SST observations and their reduced space OS analysis for December 1877 (panels (a) and (b)) with verification through the experiment with 1986 data: simulated OS analysis for December 1986 using the data distribution of 1877 (panels (e) and (f)) versus the standard OS analysis for December 1986 with all available data (panels (c) and (d)). Also shown are large-scale errors in the two reconstructions (panels (e) and (f)) and the NCEP OI December 1986 field presented in (i) $5^\circ \times 5^\circ$ and (j) $1^\circ \times 1^\circ$ resolution. Units are $^\circ\text{C}$.

the full data (panels (e) and (f)). As expected, the magnitude of the large-scale estimated error is much larger for the reconstruction from the December 1986 reduced quality simulation, than for the reconstruction from the complete data (panels (g) and (h)). Further tests show that our reconstructions are very similar to the Reynolds and Smith (1994) NCEP OI estimates of December 1986 SST anomaly (the NCEP OI combines in situ and AVHRR satellite data), though the latter is richer in small-scale details, particularly when presented in its full $1^\circ \times 1^\circ$ resolution (panels (i) and (j)).

To test the analysis for a period not used in estimating the covariance structures, we carried out additional experiments as follows: the Reynolds and Smith (1994) NCEP OI SST anomaly fields for 1992-1996 were chosen as the ‘true’ solution. These ‘true’ data were resampled and corrupted by noise according to the data availability and our estimates of observational error for the 1916-1920 period (Figure 5). The average rms error for available observations is 0.74°C , and there are many locations where the SST is not observed at all (panel (a)). The analysis of the simulated data differs from the NCEP OI fields by 0.48°C on average (panel (b)). However, the major part of this difference is in the error of truncation: the variance of NCEP OI fields which cannot be represented by the 80 EOFs used in our reconstruction (cf. Figure 6f from Kaplan *et al.*, 1998). Projecting the NCEP OI fields on the linear subspace defined by the 80 EOFs from our analysis provides the ‘reduced space version’ of the true SST field (and incidentally allows for a statistically homogeneous extension of reduced space historical analyses by higher quality modern period data sets: extension of our OS by the reduced space projection of the NCEP OI is now publicly accessible, see Acknowledgments). Our simulated analysis differs on average by 0.31°C from this reduced space version of truth (panel (c)), which is in good agreement with the average theoretical error estimate, 0.28°C (panel (d)). The years 1992-1996 are outside the period used in constructing the covariance estimate and are marked by strikingly different behaviour. Thus, these experiments demonstrate that even with limited data, the reduced space OS is able to reconstruct the global SST in a period when the covariance structure is somewhat different from the one used by the analysis procedure.

We also applied the reduced space OI analysis to the SLP data of Comprehensive Ocean-Atmosphere Data Set (COADS, Release 1 extended by standard Release 1a; Woodruff *et al.*, 1987, 1993) to produce $4^\circ \times 4^\circ$ fields of SLP monthly anomaly for the 1854-1992 period (Kaplan *et al.*, 2000). Note that both our SST and SLP analyses utilize only ship observations presented in the form of monthly ‘superobservations’ (Smith *et al.*, 1996) - mean values for $5^\circ \times 5^\circ$

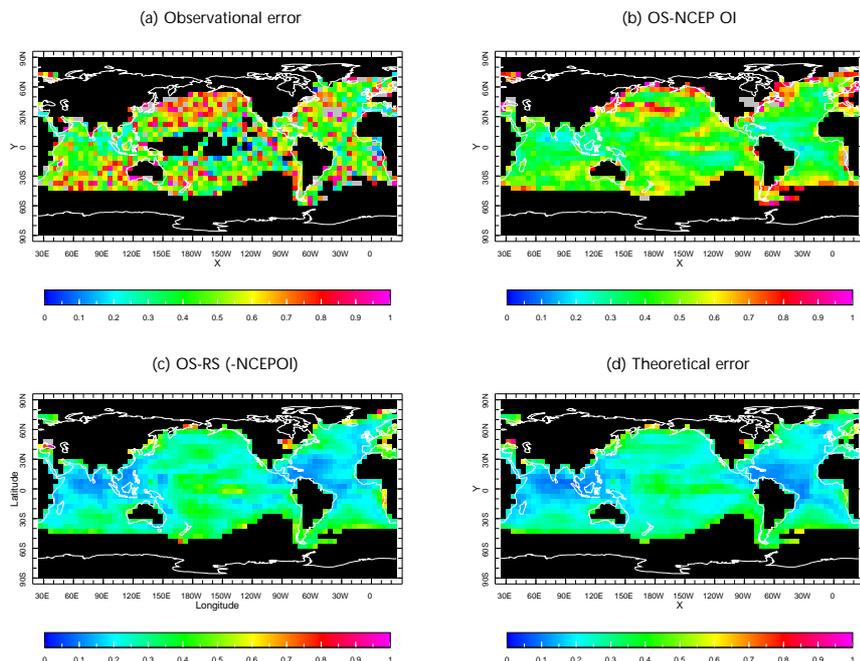


Figure 5—Statistics of the experiment with the NCEP OI data for the 1992-1996 period resampled according to the 1916-1920 observational coverage. See text for explanation.

(MOHSST5) or $2^{\circ} \times 2^{\circ}$ (COADS) bins. The UK Met Office applies so-called ‘winsorization’ (Bottomley *et al.*, 1990) to the content of their bins which makes the bin average more similar to a median. The COADS maintains a variety of statistical characteristics of the bin contents in its ‘monthly summaries’: in addition to the mean, it provides a number of observations, their standard deviation, median, sextiles, etc. Pre-war SST data of the UK Met Office has Folland and Parker (1995) ‘bucket corrections’ applied to it.

Figure 6 shows the monthly values of the analysed NINO3 (mean SST for the eastern equatorial Pacific 5°S - 5°N , 150° - 90°W), a familiar El Niño - Southern Oscillation index, with 3σ error bars supplied by the analysis. Obviously, the analysis eliminates a great deal of noise present in direct NINO3 estimates from the observed data, and agrees well with the Quinn (1992) list of El Niño events which is based on a variety of land-based, historical factors known to be associated with El Niño. The summary comparison of annual mean NINO3 with Quinn’s data is shown in Figure 7. The relation is strong but not perfect: six El Niño events, rated as ‘moderate’ or weaker by Quinn have in fact negative (as large as -1°C for 1874) annual NINO3 from our analysis. The latest of them happened in 1943, others occurred in the 19th century. However, the analysis of the Southern Oscillation (SO) and associated coastal phenomena for the period 1926-1986 by Deser and Wallace (1987) suggests that the coastal SST index might show a stronger connection to Quinn’s index of El Niño events. For this purpose we created a coastal SST index by averaging the results of the OS analysis over the NE triangular half of the $[15^{\circ}\text{S}$ - 0°N , 90° - $75^{\circ}\text{W}]$ square (the diagonal included). Indeed, the comparison presented in Figure 7 qualitatively supports this suggestion: only 2 of the events (1871 and 1907) have a negative coastal SST value. The

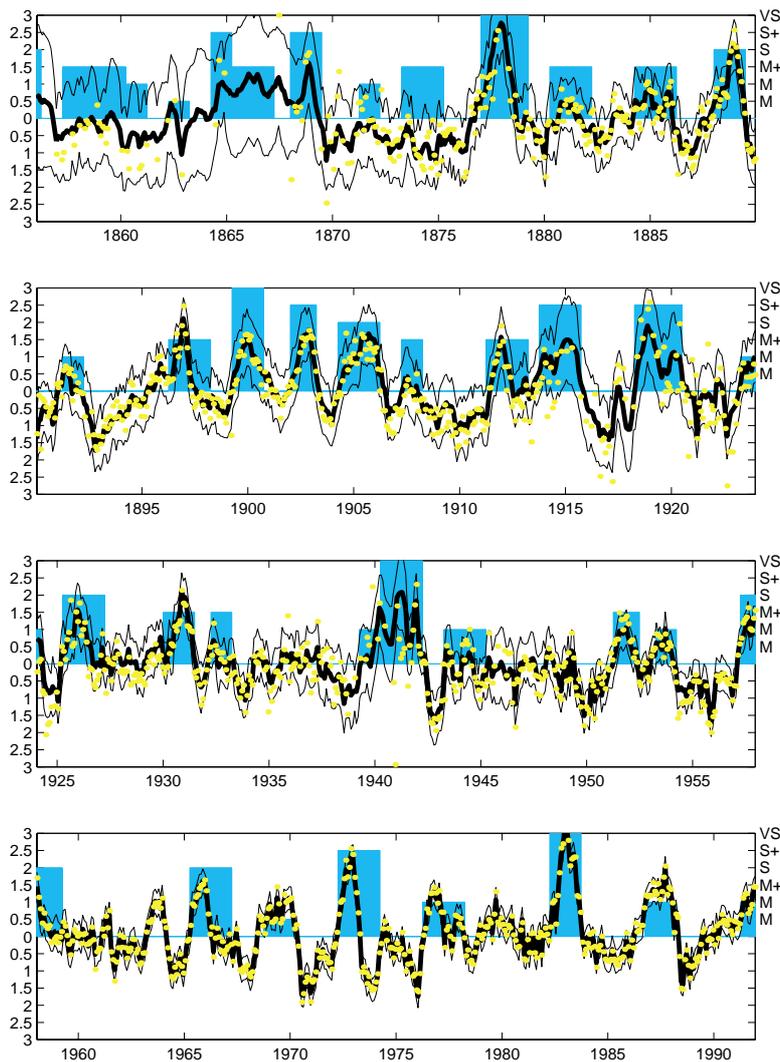
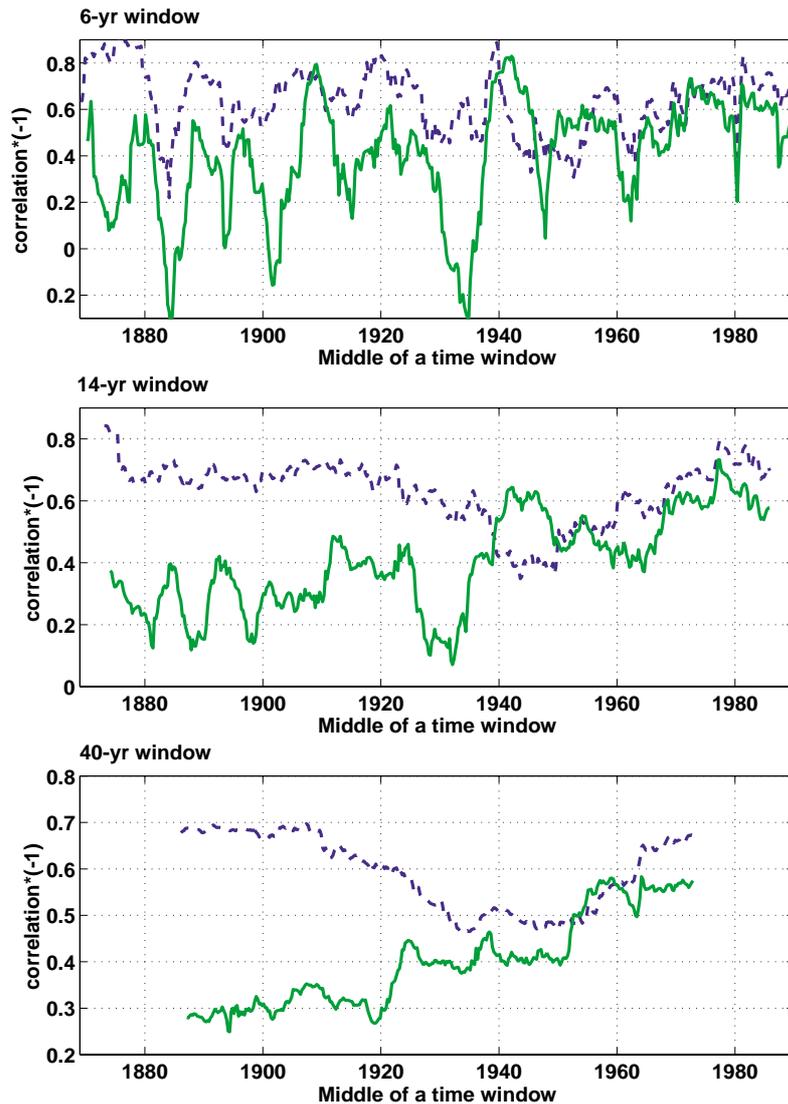


Figure 6—The reduced space optimal smoother reconstruction of the NINO3 index based on ship observations of SST (thick solid line) (Kaplan *et al.*, 1998). Also shown are 3 error bars on the analysis values (thin lines), the straight estimates of NINO3 from raw data (dots), and the ENSO event ratings of Quinn *et al.*, 1992 (histogram bars). Histogram bars are scaled to the Quinn *et al.* (1992) ratings: M-, M (moderate), M+, S (strong), S+, and VS (very strong).

Figure 8—Correlation coefficient (multiplied by -1) between Darwin and Tahiti land station records (solid lines) and their marine-based proxies (dashed lines) shown as a function of time for time windows of 6, 14, and 40 years.



computed in different width time windows and presented as functions of time. The correlation coefficients are close for the modern period, but the land station values are lower during earlier periods. We suggest that the correlation between the land-based data weakens for the early part of the record owing to degraded data quality. Factors like instrument defects and replacements, changes in observational times and location can create systematic problems in early fragments of station records; some of these problems for Darwin and Tahiti records are documented, and corrections are customarily applied (Ropelewski and Jones, 1987; Allan *et al.*, 1991). It is most likely, however, that there are uncorrected biases still left in these records, particularly in the one for Tahiti (Kaplan *et al.*, 2000).

On the other hand, the sparser and more erratic marine data force the analysis to reproduce less smaller scale (and thus more error-prone) phenomena, and to leave mostly the large-scale SO-associated pressure changes in the reconstruction. That strengthens the correlation between the analysis proxies for Darwin and Tahiti which are located near antinodes of the SO. Note that this correlation increase occurs as the response of the analysis procedure to a systematic decrease in the quality of marine data, despite the underlying assumption of constant covariance for an estimated field.

In fact, correlation between Darwin and Tahiti SLP records has traditionally been interpreted (Trenberth, 1984) as an indicator of the signal-to-noise ratio when these station records are used as the indices of the SO (in this case the 'signal' is the SO, everything else is the 'noise'). Note that most of the weakening episodes in six-year window correlations exhibited by the land stations in the

early part of the record are mimicked by the marine proxy correlations. Those episodes are most likely the realistic changes in the strength of SO relative to the background atmospheric noise. Those which are present only in the land records might be either spurious or missed in the marine records because of the sparsity of COADS coverage, at those particular times. The level of certainty of the latter possibility may change significantly when the SLP from the ‘Dutch’ deck, a major COADS component prior to the Second World War, is included in the monthly summaries in further COADS releases (Woodruff *et al.*, 1998). Even at the present level of coverage, the indices based on ship observations may provide a cleaner indication of the large-scale phenomena than the local land-based records.

3. DIFFICULTIES AND WAYS TO RESOLVE THEM

3.1 SPECIFICATION OF OBSERVATIONAL UNCERTAINTY

The advantages of the reduced space optimal analysis do not come for free: they are based on our knowledge of a priori estimates, namely covariances of observational error R and of the first two statistical moments of the solution: its mean field T_m and its covariance C . All these necessary values can be computed only approximately from the observations.

In computing R (which allows the analysis to distinguish between poor and high quality superobservations), we use intrabox variability and a number of observations for the superobservational bins. When we analyse the UK Met Office SST data, we have to estimate their intrabox standard deviations from the COADS monthly summaries because the UK Met Office does not maintain any intrabox statistics but winsorized means in its official data format. Our estimates of observational error are far from perfect. Figure 9 shows the map of our estimated single ship observational error (values used in the analysis by Kaplan *et al.*, 1998). These errors are standard deviations of individual measurements taken during one month within a given $5^\circ \times 5^\circ$ box. Such deviations account for both instrumental and sampling error (for a single measurement the latter is equal to the natural variability of SST in the given space-time box). Note that these deviations from mean values are computed for monthly bins, so they do not reflect month-to-month or longer climate variability. These values can be easily computed for larger bins, if mean and standard deviation statistics are available for their parts (Kaplan *et al.*, 2000, p. 2989).

Comparison of Figure 9 with the map of random error estimates by Kent *et al.*, 1999 (their Figure 3d) brings uneven conclusions. The latter map does not include any kind of sampling error. This explains the much larger values of Figure 9 in the regions of Gulf Stream and Kuroshio Current. However, outside these areas, the map of Figure 9 should also give larger values. This does not seem to be the case everywhere: insufficient density of observations does not allow for an adequate sampling of the SST natural variability in many areas of the world ocean. For the SLP, the contribution of sampling variability into our estimates of a single ship error (not shown) is so large, that our COADS-based estimates (used by

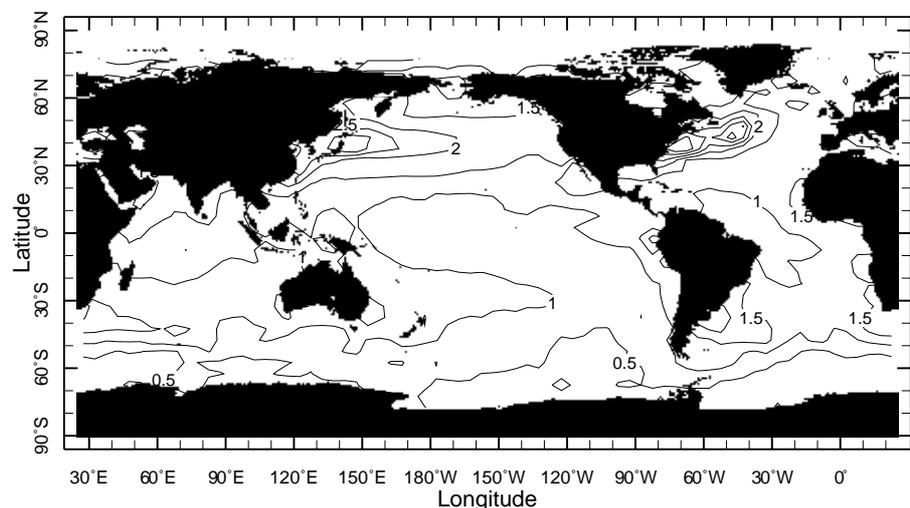


Figure 9—Intrabox SST variability ($^\circ\text{C}$) inside $5^\circ \times 5^\circ$ bins, estimated from COADS, and used as a single ship error by Kaplan *et al.* (1998).

Kaplan *et al.*, 2000) exceed those of Kent *et al.*, 1999 (their Figure 3b) by the factor of 3 in the mid-latitudes and marginally in the tropics.

Clearly, a lot more work should be carried out in this direction until really reliable observational error estimates enter gridded analyses of climate variables. An important step in this direction would be to bring to the attention of all data centres the necessity to include the statistics of intrabox distributions in their standard data formats, rather than just providing box mean values. This seems to be particularly crucial in the planned blending project of the COADS and UKMO data banks (Woodruff *et al.*, 1998). The comparison of ship-based estimates, like that of Figure 9, with those obtained from satellite data suggests that the ship-based estimates are affected by the sampling error even for the periods of the best coverage. Hence, the satellite data must be used to supplement the ship-based estimates of the small-scale variability.

3.2
CHARACTERIZATION OF THE
SIGNAL

The problems with the reliable estimation of T_m and C are even more fundamental. Ideally, these statistical characteristics of the signal are supposed to be applicable to the entire period of the analysis. In fact, poor data quality and sparse coverage in the early part of the record forces us to use only the modern data period for the derivation of T_m and C . In the applications described above, we used climatological means for the 1951-1980 period and estimated the covariance for the period from 1950 to the beginning of the 1990s. An analysis is then made using these values for as far back as the middle of the 19th century.

Problems with the mean

It was observed by Hurrell and Trenberth (1999) that a linear trend for the 20th century computed from our SST analysis shows somewhat less warming than other estimates. They suggested that this is due to the ‘stationarity’ assumption: the hypothesis that the modern-period mean and covariance are applicable for the entire record. If, in fact, the long-term variability of SST (e.g. trend) resulted in a much different mean SST state for the first half of the century, and the pattern of this change is not well-represented by the modern-period covariance, the analysis might underestimate this change.

At present, we are addressing this issue through the analysis of data residuals, the difference between the observed data and our analysis. These residuals presumably consist of two major components: observational and sampling error and part of long-term variability unresolved by the analysis. Because of the very different characteristics of these components, it should be relatively easy to isolate the latter. Prospective methods of isolation include the application of the reduced space OI and OS technique to the residuals and covariance reestimation (Kaplan *et al.*, 1997, 2000) and polynomial spline smoothing of the residuals (Wahba, 1990). Note that bestfitting a straight line or other slowly changing functions of time to the residuals can be brought into the prospect of optimal estimation and provide error bars for the trend estimates, because all other variability in the residuals is expected to be temporally uncorrelated errors. The same approach does not work for fitting slowly changing functions of time to the actual temperature changes, as the latter contains a complete spectrum of temporally correlated variability, from secular to intermonthly. If those are not removed, one should not assume the ‘whiteness’ (mutual statistical independence) of errors, for such an assumption will result in unrealistically low theoretical estimates for the uncer-

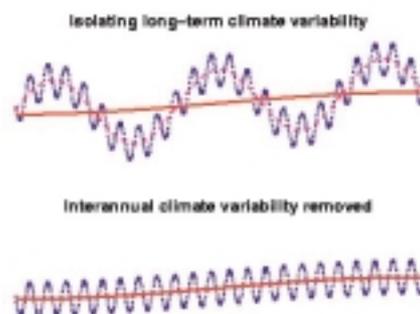
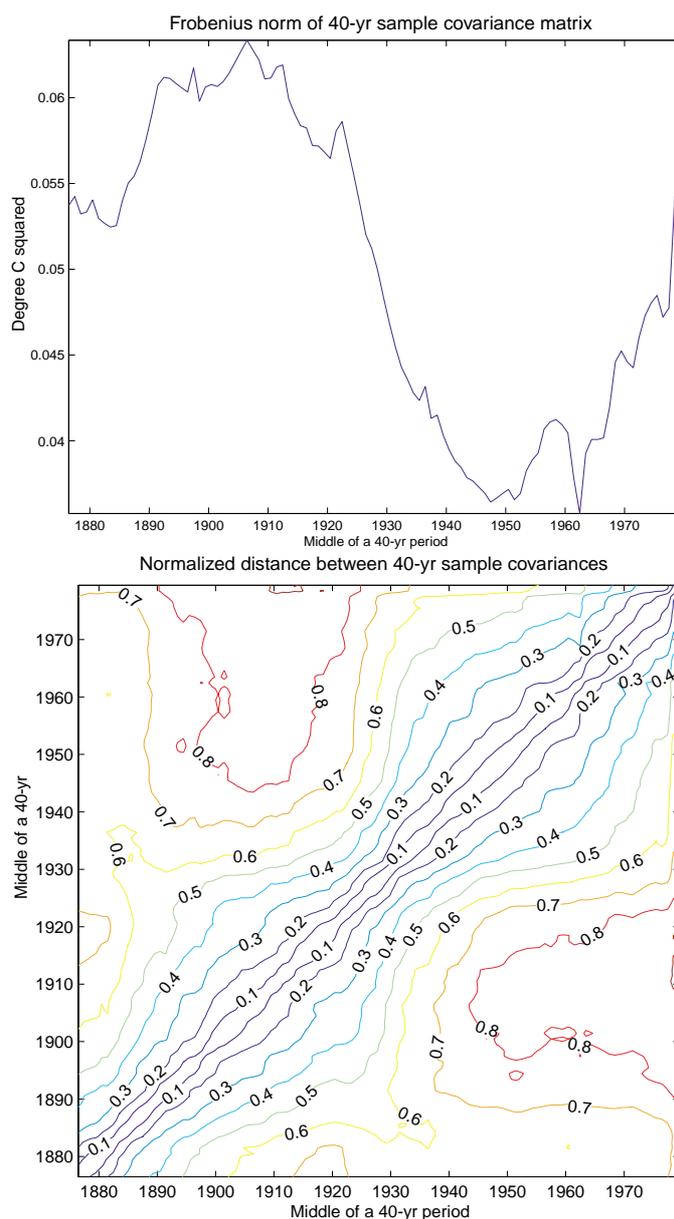


Figure 10—Removal of interannual climate variability (dashed line) from the observed data leaves the mixture of long-term variability (solid line) and error (dots).

Figure 11—(top) The Frobenius norm of SST covariance estimated in 40-year time windows, (bottom) normalized distance between sample covariance matrices estimated for different 40-year windows.



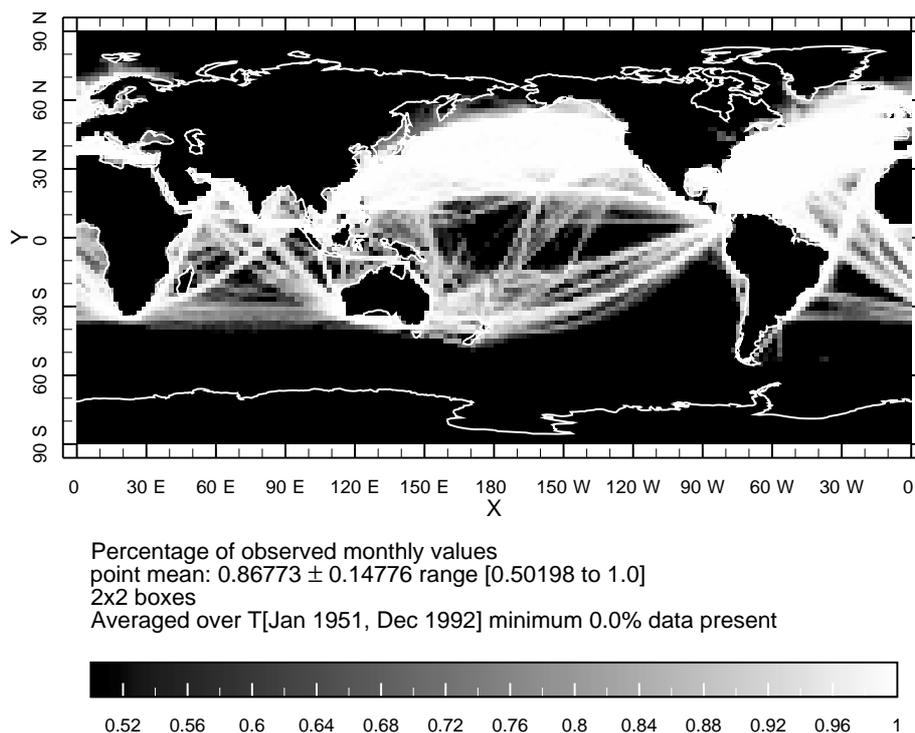
tainty of the fit. Figure 10 presents a drawing emphasizing the advantage of removing the interannual climate variability from the observed data prior to estimating long-term changes.

Our preliminary analysis indeed shows some long-term variability left in the residuals. Once we are done with its complete estimation, we will be able to estimate the total long-term variability in the SST record and measure its uncertainty. Then we will either separate it from the raw observations before applying the analysis procedure, or make sure that it is properly represented in the covariance structure.

Problems with the covariance: stationarity

The assumption of covariance stationarity and the possibility of its negative consequences comes up quite often in discussions, but has not, to the best of our knowledge, been systematically researched. As a first attempt at this, we compared covariance matrices estimated in different 40-year time windows for our SST analysis. Since the current SST analysis was performed under a conservative assumption of stationary covariance, this comparison probably underestimates the actual covariance variability. As a measure of distance between two covariance matrices, we use the Frobenius norm: the square root of the sum of squares of all elements in the matrix difference (Golub and Van Loan, 1996). Even the norm of covariance matrix itself seems to change dramatically over

Figure 12—Percentage of months with observed SLP in COADS 22° monthly summaries for the 1951-1992 period.



time, with the minimum in 1930-1970, and the maximum in 1890-1920 (Figure 11, top). The diagram of normalized distances (norm of a difference divided by the norm of the covariance matrix computed for the recent 40 years) suggests that every period in the last 1.5 centuries was in some sense unique: the farther from each other the middles of sample periods get, the larger the normalized distance between matrices (Figure 11, bottom). During some periods (like the one centered on 1930), this change happens very fast; in others (1910 and 1950) it occurs more slowly.

The successful validation of many aspects of our SST analysis so far shows that the exhibited instability of the covariance matrix does not render the analysis completely wrong or useless: the inherent robustness of the least squares estimates can absorb some level of inadequacy of a priori estimates. Moreover, all the different covariance matrices compared in Figure 11 were produced by the analysis of Kaplan *et al.* (1998) under the assumption that the covariance of the SST field is constant and equal to the sample covariance of 1951-1991. It seems reasonable, however, to involve data from all time periods in the computation of the covariance and to either use the estimate which would be applicable to the entire analysis period, or to account for slow changes in time of the covariance structure in our analysis methodology.

Problems with the covariance: resolution and coverage

The significant volatility of the covariance structure discourages the use of only the modern period of particularly good (helped by satellite coverage) data for covariance estimation. If we are determined to estimate the large-scale covariance structures from a period of no shorter than a few decades, this imposes certain restrictions on the spatial resolution with which covariance can be estimated. Before analysing COADS SLP data we tried to estimate covariance for $2^\circ \times 2^\circ$ spatial bins, and found that the analysis domain had large holes (shown in black in Figure 12) in the tropical Pacific. It took averaging to a $4^\circ \times 4^\circ$ grid to 'close' these holes. As a result, the analysis domain we obtain has quite a coarse resolution and still is globally incomplete. This severely limits the usage of such analyses in the climate model studies. It seems important to be able to generalize the technique of the reduced space optimal estimation to the stage at which it can produce high resolution and globally complete analyses.

In fact, the reduced space reconstruction technique can be empowered by the multivariate approach. The principal modification of the reduced space

optimal analysis that can produce high resolution globally-complete fields is to separate an estimated field into a few terms which correspond to different scales of resolution (and thus variability). Different terms can be observed through different sources. For example, most of the ocean $5^\circ \times 5^\circ$ resolution term is well observed by ships during last 50 years, and $1^\circ \times 1^\circ$ covariability within $5^\circ \times 5^\circ$ boxes, plus all variability in the Southern Ocean can be estimated from the NCEP OI (Reynolds and Smith, 1994) for the last 15 years, etc. The set of all terms can be subjected to multivariate EOF analysis, each piece being a separate variable in this analysis. These multivariate EOFs are then used for the reconstruction of all pieces together, and thus for the entire high resolution globally-complete field. This approach has a certain ‘modular’ nature because it makes it possible to push further in both directions: very large scale variability can be estimated for very long periods from the paleodata, extending the analysis to very long periods, and certain areas of high gradients and/or good observational networks can be ‘refined’ by adding special high resolution ‘patches’.

Problems with the covariance:
representing small scales

A seemingly fruitful direction for producing high resolution objective analyses is to literally combine analyses represented by the left-hand and right-hand parts of Figure 3. Note that the exact solution for the full grid OI can be separated into two parts:

$$\begin{aligned} T &= (H^T R^{-1} H + C^{-1})^{-1} H^T R^{-1} T^o = CH^T (R + HCH^T)^{-1} T^o = \\ &= E\Lambda E^T H^T (HE\Lambda E^T H^T + HE'\Lambda'E'^T H^T + R)^{-1} T^o \\ &+ E'\Lambda'E'^T H^T (HE\Lambda E^T H^T + HE'\Lambda'E'^T H^T + R)^{-1} T^o = \\ &= E\alpha + C'H^T (HE\Lambda E^T H^T + HC'H^T + R)^{-1} T^o = E\alpha + \Delta T \end{aligned}$$

The first term $E\alpha$ here is our standard reduced space OI solution. The second part, $C'H^T (HC'H^T + R)^{-1} \Delta T^o$, represents a correction to it towards the complete (exact) solution. This correction is defined by the covariance piece C' and contributes predominantly to the small-scale variability. It is easy to check that ΔT is a formal OI solution to the estimation problem:

$$H\Delta T = \Delta T^o + \tilde{\epsilon}^o, \langle \Delta T \Delta T^T \rangle = C', \langle \tilde{\epsilon}^o \tilde{\epsilon}^{oT} \rangle = R + HE\Lambda E^T H^T$$

where $\Delta T^o = T^o - HE\alpha$ is an observational residual to the reduced space OI solution. We do not expect to be able to estimate C' from the data without any special assumptions. However, this part of covariance can be modelled statistically under certain assumptions of spatial stationarity, e.g. as a function of spatial lag, in the style of the traditional kriging or successive correction approach. Thus, these traditional techniques can be successfully used for complementing the reduced space solution with small-scale corrections.

Problems with the covariance:
consistent estimation and
uncertainty

When an a priori estimate of the signal covariance is correct, the statistics of the solution should be consistent with it, i.e. certain balance equations should be satisfied. If this is found not to be the case, a priori values can be reestimated to satisfy the balance, and then the analysis solution can be recalculated. These steps can be repeated iteratively until the solution satisfies the balance. However, the use of different balance formulations might result in somewhat different solutions.

Kaplan *et al.* (1997) introduced the balance in the form of the system of equations:

$$\begin{aligned} A_p &\stackrel{\text{def}}{=} \langle \alpha^p \alpha^p T \rangle = \Lambda + P^p \\ A_{OI} &\stackrel{\text{def}}{=} \langle \alpha^{OI} \alpha^{OI T} \rangle = \Lambda (\Lambda + P^p)^{-1} \Lambda \end{aligned}$$

which ties together covariances of the projection and reduced space OI solutions (α^p and α^{OI} respectively), error covariance for the projection solution P^p , and the reduced space representation of the covariance. The projection solution consists of the best fit coefficients of EOF patterns to the observed data. P^p is the

theoretical covariance of the error in these coefficients. Originally they used the one-parametric heuristic formula for ‘redistributing’ the spectrum of Λ . This seemed to give satisfactory results for SST analyses, but failed when applied to the SLP analysis by Kaplan *et al.* (2000). Because of that, the latter work reduced the system to a single nonlinear matrix equation for Λ :

$$A_p = \Lambda A_{OI}^{-1} \Lambda$$

and presented an exact solution to it. The results of the analysis satisfied the balance after the first iteration.

An alternative way to state the analysis balance can be based on the expectation maximization (EM) procedure (Schneider 2000 and references therein). In the reduced space version, and taking into account the observational error, the EM balance for the OI solution can be written as:

$$\bar{\alpha} = \langle \alpha^{OI} \rangle, \Lambda = \langle (\alpha^{OI} - \bar{\alpha})(\alpha^{OI} - \bar{\alpha})^T \rangle + P^{OI}$$

Our initial trials of this procedure for the SST analysis have shown convergence after approximately 10 iterations.

It should be noted that because of their reduced space nature, the procedures described above cannot bring the estimates of the leading EOFs outside the initially defined reduced space. However, if the small-scale correction is added after every iteration, and the full-grid covariance is reestimated, that might result in substantially better estimates of the signal covariance and perhaps overcome the limitation of ‘gappy’ and erratic data from which it is derived.

On the other hand, however complicated the technique we use, the covariance is always estimated with some uncertainty. The explicit modelling of this uncertainty, transferring it into the uncertainty of the analysed fields, perhaps in the Bayesian framework, is an important task for the future.

4. CONCLUSIONS AND PROSPECTS

We have shown that the reduced space optimal estimation is a computationally effective restructuring of the process of obtaining the full-grid optimal solution, and that it delivered verifiable analyses of climatic fields in both systematic applications to date (for SST and SLP).

The problems of the method are the same as those of any objective analysis technique: difficulty in deriving reliable a priori estimates from the sparse and erratic data. These problems might be solved, in part, if new significant volumes of data for the early periods become available (Woodruff *et al.*, 1999). It is very important that all data centres involved provide extensive statistics of intrabin distributions (as opposed to providing means only), for example, the current COADS model of monthly summaries. The use of satellite data is another prospective way of improving a priori estimates of in situ error statistics.

Land station data is another powerful information resource that can be combined in the analyses with marine observations to the advantage of the product (cf. recent SLP analysis of the UK Met Office by Basnett and Parker, 1997).

Further improvement of the analysis technique should include the systematic a priori estimation of mean, covariance, or long-term variability and changing covariance structure from the entire period of available data. Separation of the estimated fields into large- and small-scale varying components allows for the generalization of the technique which can produce high resolution globally-complete products.

The technique of reduced space optimal estimation should be more systematically applied to all climate variables for which historical (COADS) data sets are available, e.g. meridional and zonal winds, marine air temperature, humidity, or (non-COADS) precipitation, sea ice concentration, and possibly sea surface height. It also opens interesting prospects for historical analyses of ocean-atmosphere fluxes with the possible modification of applying the analysis to the system of a few physical variables (e.g. surface wind components and SLP) and using a linearized physical model (e.g. geostrophic or frictional balance) as an additional analysis constraint.

ACKNOWLEDGEMENTS

This work was supported by NOAA grant UCSIO-10775411D/NA47GPO-188, NOAA/NASA Enhanced Data Set Project grant NA06GP0567, and NSF/NOAA Earth System History grant NA86GP0437. AK is thankful to the NOAA OGP for their CLIMAR99 travel grant (administered through the UCAR) that served as encouragement and facilitated this work. Discussions with the CLIMAR99 participants Scott Woodruff, Dick Reynolds, David Parker, Liz Kent and Henry Diaz are gratefully acknowledged. The hospitality of the NCAR Geostatistics Project during AK's July 2000 visit (supported by the NSF grant DMS-9815344), their workshop on "Statistics for Large Data Sets" and stimulating discussions with Doug Nychka, Dave Higdon, Chris Jones and Chris Wikle helped to bring this work into its final shape. Discussions with Misha Chechelnsky and his technical help with Figure 1 were invaluable. Remarks made by both anonymous reviewers helped to improve the clarity of the manuscript. Benno Blumenthal's Ingrid software is responsible for all ocean domain plots in this work, and his Data Library system uses Senya Basin's CUF format for providing easy public access to the analysed data presented in this work at:

http://ingrid.lidgo.columbia.edu/SOURCES/.KAPLAN/.RSA_MOHSST5.html for the SST analysis,

http://ingrid.lidgo.columbia.edu/SOURCES/.KAPLAN/.RSA_COADS_SLP1.html for the SLP analysis, and

<http://ingrid.ldeo.columbia.edu/SOURCES/.KAPLAN/.EXTENDED/> for the OS SST analysis extended monthly to the present (this is achieved by concatenating the NCEP OI projections onto the analysis' reduced space). This work is a Lamont-Doherty Earth Observatory contribution number 6149.

REFERENCES

- Allan, R.J., N. Nicholls, P.D. Jones and I.J. Butterworth, 1991: A further extension of the Tahiti-Darwin SOI, early SOI results and Darwin pressure. *J. Climate* 4, 743-749.
- Basnett, T.A. and D.E. Parker, 1997: Development of the global mean sea level pressure data set GMSLP2. Hadley Centre of the UK Met Office. *Clim. Res. Tech. Note* 79. Unpublished document available from the Hadley Centre for Climate Prediction and Research, Meteorological Office, London Road, Bracknell, RS12 2SY, U.K.
- Bottomley, M., C.K. Folland, J. Hsiung, R.E. Newell, D.E. Parker, 1990: *Global Ocean Surface Temperature Atlas*. HMSO, London.
- Cane, M.A., A. Kaplan, R.N. Miller, B. Tang, E.C. Hackert and A.J. Busalacchi, 1996: Mapping tropical Pacific sea level: data assimilation via a reduced state space Kalman filter. *J. Geophys. Res.*, 101, 22599-22617.
- Cressie, N.A.C., 1991: *Statistics for Spatial Data*, Wiley-Interscience.
- Deser, C. and J.M. Wallace, 1987: El Niño events and their relation to Southern Oscillation: 1925-1986, *J. Geoph. Res.*, 92, 14,189-14,196.
- Daley, R., 1993: *Atmospheric Data Analysis*, Cambridge University Press, 457 pp.
- Da Silva, A.M., C.C. Young and S. Levitus, 1994: *Atlas of Surface Marine Data*. Vol. 1, Algorithms and Procedures, NOAA, 83 pp [Available from U.S. Department of Commerce, National Oceanographic Data Center, User Services Branch, NOAA/NESDIS E/OC21, Washington, DC 20233].
- Folland, C.K. and D.E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data, *Q. J. R. Meteorol. Soc.*, 121, 319-367.
- Golub, G.H. and C.F. Van Loan, 1996: *Matrix Computations*. Third edition, The John Hopkins University Press, Baltimore. 694 pp.
- Hurrell, J.W., and K.E. Trenberth, 1999: Global sea surface temperature analyses: multiple problems and their implications for climate analysis, modeling, and reanalysis. *Bull. Amer. Meteor. Soc.*, 80, 2661-2678.
- Kaplan, A., Y. Kushnir, M.A. Cane and M.B. Blumenthal, 1997: Reduced space optimal analysis for historical datasets: 136 years of Atlantic sea surface temperatures. *J. Geoph. Res.*, 102, 27,835-27,860.
- Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856-1991, *J. Geoph. Res.*, 103, 18,567-18,589.

- Kaplan, A., Y. Kushnir and M.A. Cane, 2000: Reduced space optimal interpolation of historical marine sea level pressure: 1854-1992. *J. Climate*, 13, 2987-3002.
- Kent, E.C., P. Challenor and P. Taylor, 1999: A statistical determination of the random observational errors present in voluntary observing ships meteorological reports. *Journal of Atmospheric and Oceanic Technology*, 16, 905-914.
- Konnen, G.P., P.D. Jones, M.H. Kaltofen and R.J. Allan, 1998: Pre-1866 extensions of the Southern Oscillation Index using early Indonesian and Tahitian meteorological readings. *J. Climate*, 11, 2325-2339.
- Levitus, S. and T.P. Boyer, 1994: *World Ocean Atlas 1994*. Vol. 4, Temperature. NOAA Atlas NESDIS 4, U.S. Department of Commerce, Washington, DC, 117 pp. [Available from U.S. Department of Commerce, National Oceanographic Data Center, User Services Branch, NOAA/NESDIS E/OC21, Washington, DC 20233].
- Mann, M.E., R.S. Bradley, M.K. Hughes, 1998: Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392, 779-787.
- Mardia, K.V., J.T. Kent and J.M. Bibby, 1979: *Multivariate Analysis*, Academic, San Diego, Calif., 521 pp.
- Parker, D.E., P.D. Jones, C.K. Folland and A. Bevan, 1994: Interdecadal changes of surface temperature since the late nineteenth century. *J. Geophys. Res.*, 99, 14,373-14,399.
- Parker, D.E., C.K. Folland and M. Jackson, 1995: Marine surface temperature: Observed variations and data requirements, *Clim. Change*, 31, 559-600.
- Quinn, W.H., 1992: A study of Southern Oscillation-related climatic activity for A.D. 622-1900 incorporating Nile River flood data, in *El Niño Historical and Paleoclimatic Aspects of the Southern Oscillation*, edited by H.F. Diaz and V. Markgraf, pp. 119-149, Cambridge Univ. Press, New York.
- Rao, C.R., 1973: *Linear Statistical Inference and its Applications*, John Wiley, New York, 625 pp.
- Rayner, N.A., E.B. Horton, D.E. Parker, C.K. Folland and R.B. Hackett, 1996: Version 2.2 of the global sea-ice and sea surface temperature data set, 1903-1994. *Clim. Res. Tech. Note* 74. Unpublished document available from the Hadley Centre for Climate Prediction and Research, Met Office, London Road, Bracknell, RS12 2SY, U.K.
- Reynolds, R.W. and T.M. Smith, 1994: Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, 7, 929-948.
- Ropelewski, C.F. and P.D. Jones, 1987: An extension of the Tahiti-Darwin Southern Oscillation Index. *Mon. Weather Rev.*, 115, 2161-2165.
- Schneider, T, 2001: Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate*, in press.
- Shriver, J.F. and J.J. O'Brien, 1995: Low-frequency variability of the equatorial Pacific ocean using a new pseudostress dataset: 1930-1989. *J. Climate*, 8, 2762-2786.
- Smith, T.M., R.W. Reynolds, R.E. Livezey and D.C. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, 9, 1403-1420.
- Trenberth, K. E., 1984: Signal versus noise in the Southern Oscillation. *Mon. Wea. Rev.*, 112, 326-332.
- Wahba, G., 1990: Spline models for observational data. CBMS-NSF, *Regional Conference Series in Applied Mathematics*, Vol. 59, Society of Industrial and Applied Mathematics, 169 pp.
- Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. *Bull. Amer. Meteor. Soc.*, 68, 521-527.
- Woodruff, S.D., S.J. Lubker, K. Wolter, S.J. Worley and J.D. Elms, 1993: Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a : 1980-92. *Earth System Monitor*, 4, No. 1, 1-8.
- Woodruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, 23, 517-526.

ANALYSIS OF WAVE CLIMATE TRENDS AND VARIABILITY

Val R. Swail¹; Andrew T. Cox and Vincent J. Cardone²

1. INTRODUCTION

This paper describes the analysis of wave climate trends and variability from two long-term (40 years) wave hindcasts recently carried out by Environment Canada and Oceanweather. In this study, the NCEP/NCAR Reanalysis (NRA) surface (10 m) wind fields at six-hourly intervals were used to drive a global spectral ocean wave model for the 1958-1997 period. The detailed North Atlantic hindcast was based on kinematically reanalysed NRA wind fields, as described by Swail and Cox (1999). These enhanced wind fields were demonstrated to be a significant improvement over the NRA winds. A description of the evaluation of both hindcasts against in situ and satellite data is given in this publication by Cox *et al.* (2003).

The issue of ocean wave variability and trend has been investigated in recent years by many researchers, using different data sets. Investigations using instrumental measurements in the North Atlantic were carried out by Carter and Draper (1988), and Bacon and Carter (1991, 1993). Bouws *et al.* (1996) studied operational wave analyses for ship routing prepared by the Koninklijk Nederlands Meteorologisch Instituut (KNMI). Gulev and Hasse (1999) used visual wave observations from voluntary observing ships. In recent years, several wave hindcast studies have been undertaken, including the Kushnir *et al.* (1997) ten-year hindcast of the North Atlantic, the Sterl *et al.* (1998) 15-year global hindcast based on the European Centre for Medium-range Weather Forecasts Reanalysis (ERA15), and the European Union Waves and Storms in the Atlantic (WASA) project 40-year hindcast of the Northeast Atlantic ocean (WASA, 1998). In general, all of these works showed an increase in significant wave height in the North Atlantic over the different periods, although details of the patterns and the magnitudes of the changes varied somewhat.

One disturbing property of earlier hindcast studies, and of real time NWP operations, is that changes over time in data sources, improvements in data analysis techniques and evolution and upgrades in numerical models have tended to impart a temporal or 'creeping' inhomogeneity into the real-time products of such centres. When the wind fields produced by these centres are used to drive a wave model, these creeping inhomogeneities are translated into the wave climate simulations. Therefore, output data quality varies over time and subtle changes in climate may be masked. By using the NRA wind fields (which were derived from one version of the NCEP model for the entire 40 years) as a base, much of the inhomogeneity should be removed. However, it also must be noted that the assimilation input changed with time, and this could still be a source of inhomogeneity. White (2000) noted that many trends in the NRA were correlated with the change in the number of observations. The largest impact was found in the southern hemisphere, and the North Atlantic is probably less influenced by changing data coverage. There are specific inhomogeneities in the North Atlantic as well, such as the termination of the Ocean Weather Ship programme in the early 1970s, that may have an inverse effect.

Fifteen statistics were computed for both the resultant wave heights and input wind fields on monthly, seasonal and annual time scales; trend and variability analysis was carried out for each grid point in both hindcasts. In addition,

1 Environment Canada, 4905 Dufferin Street, Downsview, Ontario, Canada M3H 5T4;
e-mail: Val.Swail@ec.gc.ca

2 Oceanweather, Inc. 5 River Road, Cos Cob, CT 06807, USA

a spatial analysis was carried out for the northern hemisphere oceans relating the wave climate to the surface pressure patterns. The results of the two wave hindcasts were compared with each other, and with homogeneous point time series of waves to investigate potential biases in the trend analyses. The paper is organized as follows. Section 2 briefly describes the global and North Atlantic hindcasts. Section 3 provides a description of the wind and wave climate derived from each hindcast. Section 4 gives an assessment of the potential biases in the hindcast and in situ data. Section 5 contains our conclusions.

2. WAVE HINDCASTS

The global wave hindcast (GROW – Global Reanalysis of Waves) was carried out using Oceanweather's ODGP2 1-G fully discrete spectral wave model with a grid resolution of 2.5° longitude by 1.25° latitude. Wind fields are derived directly from the NCEP Reanalysis surface 10 m winds, updated at six-hourly intervals, and the model time step is three hours. The only modification to the wind fields in the global model was to convert them to effective neutral stability using the NRA 2 m temperature and sea surface temperature fields. Details of the selection of these particular NRA wind fields and their validation are given in Cardone *et al.* (2003; this publication). In the global model, ice fields were specified on a monthly basis, using long-term monthly historical ice concentration data. Details of the global hindcast methodology are given by Cox and Swail (2001).

The North Atlantic wave hindcast (AES40) was carried out using the ODGP 3-G wave model, with a grid resolution of 0.625° latitude by 1.25° longitude. The ice edge was based on the actual monthly ice concentration. The NRA wind fields were reanalysed and enhanced with the aid of analyst-interactive techniques, during which in situ data were correctly re-assimilated, wind fields in extratropical storms were intensified as necessary, and tropical cyclone boundary layer winds were included. Swail and Cox (1999) describe the generation of these wind fields in detail, and show the significant improvement in the reanalysed wave fields, particularly in the specification of storm peaks.

The wave height fields produced in this hindcast showed excellent agreement with in situ wave measurements and satellite wave estimates. Cox and Swail (2001) show detailed comparisons with the global hindcast, while Cox *et al.* (2003; this publication) show overall comparisons with both the global and detailed North Atlantic hindcasts.

3. CLIMATE ASSESSMENT

Fifteen statistics were computed for both the resultant wave heights and input wind fields on monthly, seasonal and annual time scales; trend and variability analysis was carried out for every grid point in each hindcast. Among the statistics computed were mean, standard deviation, skew, kurtosis, 50th, 90th, 95th and 99th percentiles, and exceedance above selected thresholds.

Figure 1 shows the mean annual wind speed and wave height distribution for the period 1958 to 1997 for the global (GROW) and North Atlantic (AES40) hindcasts. The maxima in the high-latitude areas in both hemispheres and along the prevailing storm tracks are very evident in these charts. It is interesting to note that wind speeds over land are far less than those over the oceans. As found by Sterl *et al.* (1998), the waves in the North Atlantic are higher than those in the North Pacific. There are no wave hindcast data poleward of 70° in either hemisphere in GROW, or north of 76°N in AES40.

Figure 2 shows the geographical distribution of the annual 99th percentile wind speed and wave height for each hindcast for 1958-1997. The patterns in both hindcasts are very similar to those of the means, although the areas of highest wind speed and wave height are even more accentuated. The areas of strongest winds in the GROW hindcast for the 99th percentile are between Iceland and Canada, while in the mean charts the Southern Ocean showed higher values. The 99th percentile GROW wind and wave charts do not reflect areas where episodic high winds and waves might be expected due to tropical storms, such as the south-eastern US coast and the Gulf of Mexico, the South China Sea, north Australia or the Indian Ocean. This is certainly due to the inability of the NRA to adequately resolve these relatively small atmospheric features.

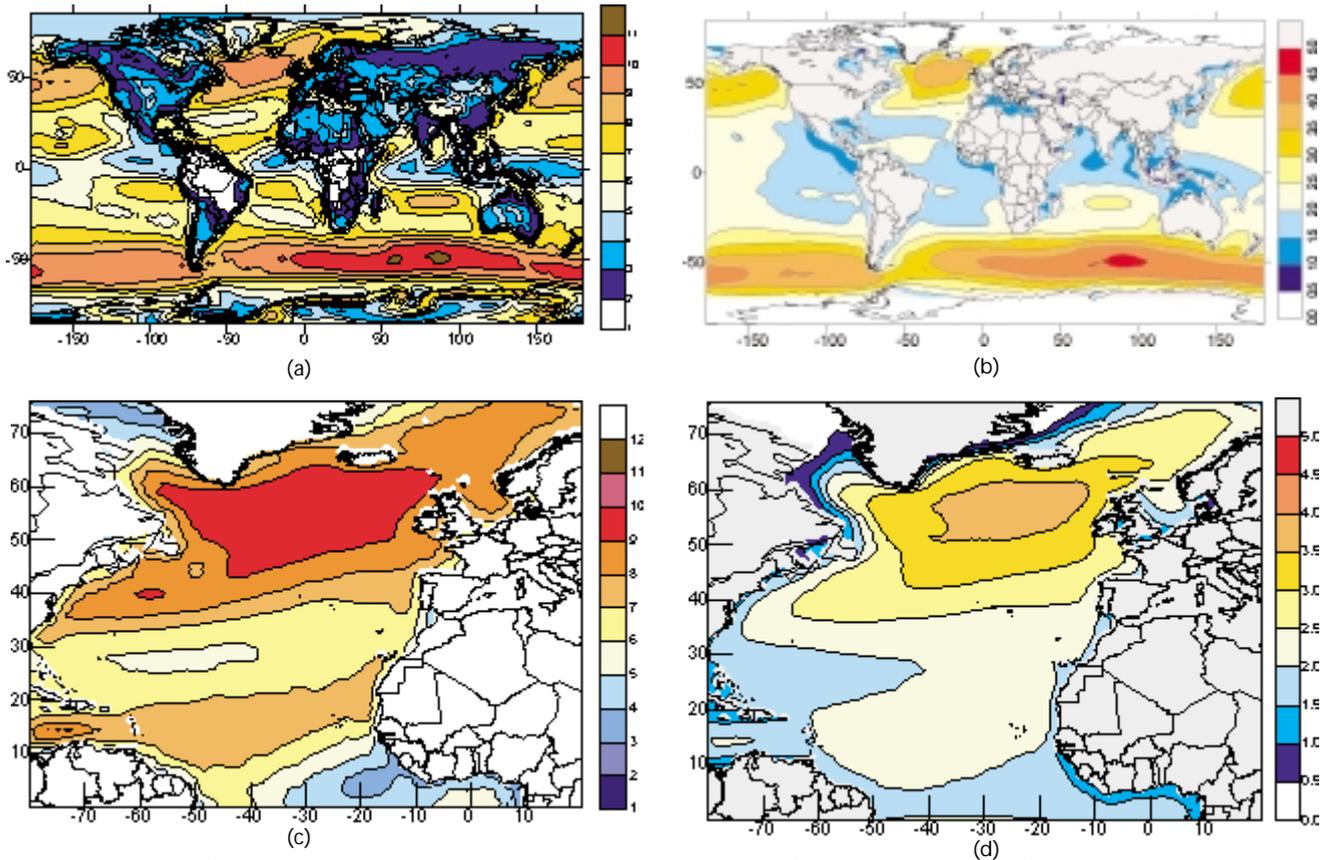


Figure 1 — Annual mean wind speed (m/s) 1958–1997 for GROW (a) and AES40 (c); and significant wave height (m) for GROW (b) and AES40 (d).

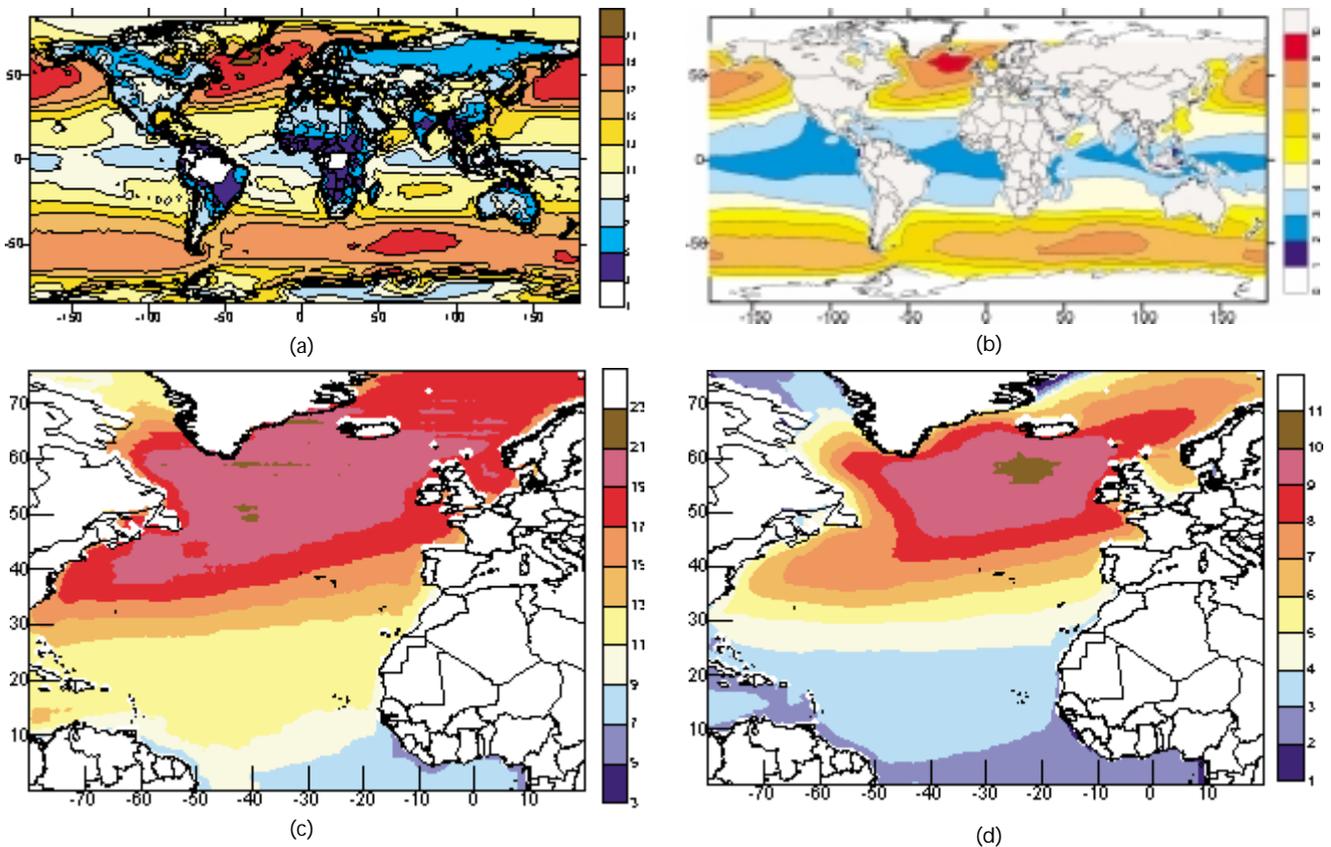
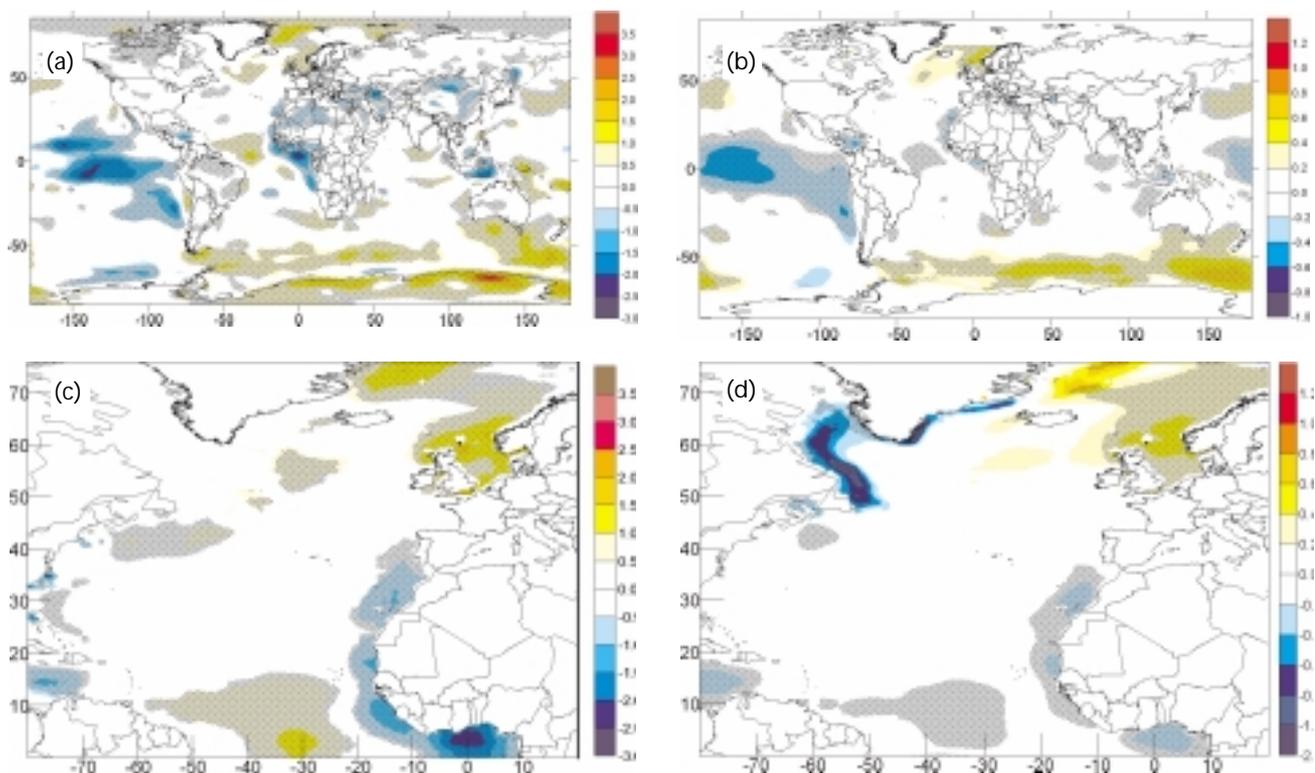


Figure 2 — 99th percentile wind speed (m/s) 1958–1997 for GROW (a) and AES40 (c); and significant wave height (m) for GROW (b) and AES40 (d).

A series of statistical analyses of the wind and wave trends was carried out for both hindcasts at each point on their respective grids. Trends were computed as simple linear trends over the 40 years of each hindcast using least squares fitting techniques. Figures 3 and 4 show the trends in the mean and 99th percentile wind speed and wave heights for the two hindcasts; trends are expressed as the inferred change over the 40-year period 1958-1997 based on the slope of the linear trend line. The areas where the null hypothesis (i.e. that the time series in question is random) is rejected at the 99 per cent level are also shown; the effect of series autocorrelation was also taken into account in determining the rejection levels. Increasing trends are most noticeable in the north-east Atlantic Ocean, across the northern edge of the North Pacific Ocean, and along the margins of Antarctica. The Antarctic trends are considered to be rather unreliable due to the data scarcity in the Southern Ocean as a whole, and documented problems in the NRA with the southern hemisphere, particularly south of 50°S. Negative trends in wave height are found mostly in equatorial regions, particularly in the Pacific Ocean, and also in the Labrador Sea. Particularly noticeable is the bi-polar nature of the trends in the North Atlantic, with strong increases in the north-east, and strong decreases in the south central North Atlantic. This pattern follows the dominant mode of the North Atlantic Oscillation. The spatial patterns of the trend in the mean and extreme (99th percentile) significant wave height are very similar. However, the magnitudes of the trends are much greater for the extreme wave heights than for the mean conditions, with large areas of increases in wave height of more than 1 m.

Wang and Swail (2001, 2002) describe in detail the results of spatial statistical analysis performed on both hindcasts; only a brief summary is included here. These studies used the Mann-Kendall test for trend against randomness at each grid point, accounting for autocorrelation; the time series were 'pre-whitened' and the autocorrelation and regressions coefficients were computed using an iterative scheme. Redundancy analysis techniques (described by Wang *et al.*, 1999) were used to carry out detailed seasonal spatial statistical analyses for both the global and North Atlantic hindcasts. Like canonical correlation analysis (CCA), redundancy analysis is a technique used to associate patterns of variation in a predictor field with patterns of the predictand field through a regression model. It differs from CCA in that it seeks to find pairs of predictor and predictand patterns that maximize the associated predictand variance, rather than the correlation only. In the North Atlantic hindcast, significant increases in

Figure 3 — Inferred change over the 1958–1997 period with 99 per cent statistical significance in annual mean wind speed (m/s) for GROW (a) and AES40 (c); and significant wave height (m) for GROW (b) and AES40 (d).



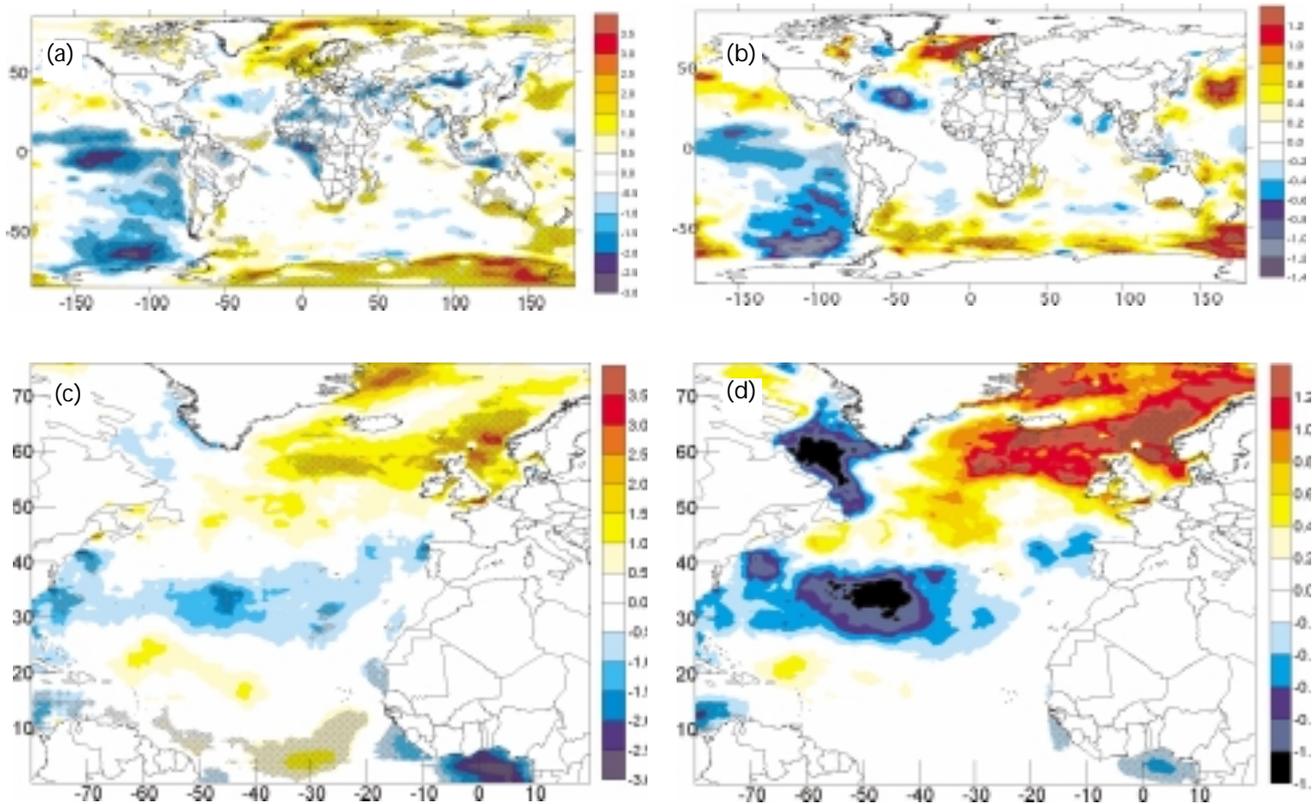


Figure 4 — Inferred change over the period 1958–1997 with 99 per cent statistical significance in annual 99th wind speed (m/s) for GROW (a) and AES40 (c); and significant wave height (m) for GROW (b) and AES40 (d).

the north-east Atlantic in the 90th percentile wave heights were matched by significant decreases in the subtropical North Atlantic, for the winter (JFM) season. The rates of increase/decrease are generally greater than those found in the global wave hindcast. Linear trends detected for the 99th percentiles are generally less significant than those for the 90th percentiles. The correlation between sea level pressure (SLP) and the 90th percentile wave height (H90) is significant at the 99th confidence level. Both time series possess a significant increasing trend at the 95 per cent confidence level, indicating that the Icelandic low has deepened during recent decades while the Azores high intensified, and, consequently, significant wave height (SWH) extremes have increased in the north-east NA, accompanied by decreases of SWH extremes in the subtropical NA. Both SLP and H90 are highly significantly correlated with the NAO index. Similar results were also found for winter (JFM) 99th percentile wave heights. In the global hindcast, changes in North Pacific winter (JFM) SWH are found to be significant at the 90 per cent confidence level; increases in SWH in the central North Pacific are found to be associated with a deepened and eastward extended Aleutian low. For both the North Atlantic and North Pacific, no significant trends of seasonal SWH extremes are found for the last century, though significant changes do exist in the last four decades; multi-decadal fluctuations are quite noticeable, especially in the North Pacific.

4. ASSESSMENT OF HOMOGENEITY

While the NRA used the same numerical prediction scheme for the 40-year period, thus removing the bias associated with ever-changing operational models, there still remain probable biases due to increased observational densities, and, particularly for ocean areas, an increase in shipboard anemometer heights coupled with an increased fraction of measured versus estimated winds. These are often referred to as ‘creeping inhomogeneities’, and are potentially serious constraints to any attempt to derive long-term trends. Therefore, we would like to verify the trend analyses derived from the two hindcasts against some long time histories of homogeneous measured data at selected points. Unfortunately, there exist very few locations in the global ocean where such data are available.

One location for which we do have reasonably homogeneous wind measurements over the 40-year period is at Sable Island, just off the east coast of Canada. We are also able to analyse the surface atmospheric pressure record from

Sable Island, along with records from two other sites in Nova Scotia (Halifax, Sydney) to compute pressure triangle wind records. As shown by Schmidt and von Storch (1993), the pressure triangle winds are most likely the least biased wind estimator available, since inhomogeneities in pressure records are much less than for most atmospheric variables.

Table 1 and Figure 5 show the trends for the Sable Island area from the hindcasts, Sable Island and the pressure triangle. In both the Sable Island measurements and the triangle winds, the trends in the percentiles are decreasing; the magnitude of the decreasing trend is comparable in both analyses, with the triangle wind trend being slightly more negative. The hindcast wind speed trends show a near-zero, but very slightly positive, trend. This most likely indicates an inhomogeneity introduced into the NRA winds. This could be a result of increased data densities in later years. For the NRA hindcast, it could also be a result of assimilating ship wind observations at an anemometer height of 10 m, when in fact the heights have increased from about 20 m at the beginning of the period to more than 30 m by the end of the period, with many observations coming from anemometers at heights exceeding 45 m. Coupled with an increase in the percentage of measured winds from ships, this could induce an artificial positive trend in the winds (and waves). In the 1990s, an increasing volume of moored buoy data would have been included in the NRA winds. These winds are taken at 5-m height, but are also assimilated at 10 m into the model. This would have the effect of reducing the wind speed trends, and thereby reducing, but not eliminating, the

Table 1—Summary of trends (per cent change/year) in winds and waves near Sable Island (1958-1997).

Per cent ILE	NRA wind	AES40 wind	SABLE IS wind	TRIANGLE wind	SHIP wind	NRA wave	AES40 wave
99	0.01	0.07	-0.12	-0.19	0.31	-0.01	0.15
90	0.03	0.09	-0.11	-0.14	0.13	-0.02	0.01
50	0.05	0.10	-0.24	-0.20	0.05	0.13	0.19

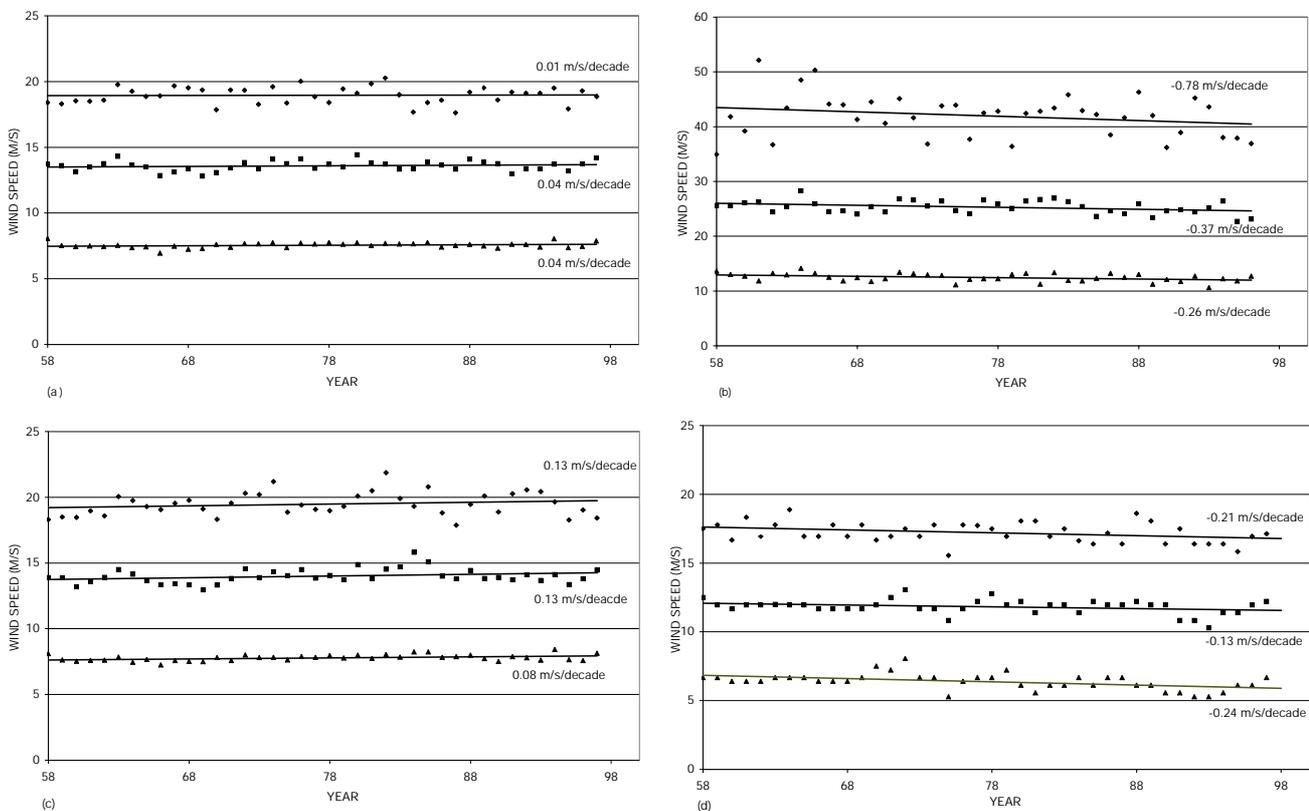


Figure 5 — Climate trends in 50th, 90th and 99th percentile wind speeds near Sable Island for (a) GROW, (b) pressure triangle, (c) AES40, (d) Sable Island measurements.

positive bias in areas near the buoys, i.e. we would expect the trends to be more positive if the buoy winds were assimilated at the correct heights. This is, in fact, what we see from the AES40 hindcast where both the ship winds and the buoy winds are assimilated at their actual anemometer heights. The AES40 trends in both winds and waves are consistently higher in this region, dominated by buoy observations in the 1990s, than the NRA trends. The AES40 trends should be a truer indication of the more intangible creeping inhomogeneities in the reanalysis process, such as increased data density, since the other sources, such as changing anemometer heights, have been mostly removed. Table 1 also shows the trends from Comprehensive Ocean-Atmosphere Data Set (COADS) ship observations. These wind speeds have been corrected where possible following the approach of Cardone *et al.* (1990). However, there remains a strong positive trend in wind speeds, particularly at the higher percentiles. Based on the Sable Island and triangle winds, this trend is most likely spurious, indicating that even these methods are unable to remove all of the artificial trend introduced by changing observational procedures on ships.

A second area for which 'ground truth' information is available for trends is off the Norwegian coast. WASA (1998) computed winds from two pressure triangles: (1) T-B-M (Thorshavn-Bergen-Mike (OWS)); and (2) T-A-B (Thorshavn-Aberdeen-Bergen). Table 2 and Figure 6 show the comparative results of the hindcasts and the WASA triangles. In this area both trends are positive, the hindcast winds being slightly more positive than the triangles. This indicates that the hindcast trends are reasonable, but probably slightly too high, or a good upper bound on real trends. Trends from adjusted ships in these areas similarly show increases in wind speed which are too strong, especially in the higher percentiles.

We have compared trends from ship wind observations in other areas with the hindcasts. In addition to the Sable Island box, and a box selected near the Hibernia oil field on the Grand Banks (47N, 47W), we have arbitrarily selected a mid-Atlantic 2° box (49N, 35W) and a box near the Bay of Biscay (45N, 9W). Table 3 shows that, except for the mid-Atlantic box, the adjusted ship trends show much larger increases than the hindcasts. It is also evident that the AES40 wind and wave trends are less than the NRA trends in the eastern Atlantic, while near Sable Island, where buoys dominate the later years, the AES40 trends exceed those from the NRA. Trends from the Labrador Sea, away from the influence of the buoys, show the same pattern as the eastern Atlantic. An anomaly appears for the Grand Banks, just outside the northern edge of the buoy coverage, where the AES40 wind trends exceed the NRA trends, but the AES40 wave trends are less.

Table 3 shows trend results from OWS Papa and OWS Bravo. Unfortunately, the overlapping period between the weathership records and the hindcasts is restricted to 24 years (Papa) and 16 years (Bravo). At Bravo, trends are negative for both the OWS and hindcasts. Consistent with a general artificial upward trend in hindcasts, the weathership trend is more negative (i.e. less positive). The same applies at the ship Papa location, although the OWS Papa trend looks somewhat suspicious, particularly the 99th percentile trend.

5. CONCLUSIONS

In this paper we have described the analysis of wave climate trend and variability from two ocean wave hindcasts: (1) a coarse mesh global hindcast based on wind

Table 2—Trends (per cent change/year) in winds and waves for WASA triangles, nearest hindcast points and 2° latitude-longitude adjusted COADS boxes (1958-1997).

	Per cent ILE	NRA wind	AES40 wind	WASA wind	SHIP wind	NRA wave	AES40 wave
TRIANGLE T-A-B	99	0.22	0.26		0.56	0.30	0.40
	90	0.28	0.28	0.23	0.44	0.40	0.44
	50	0.27	0.33		0.56	0.34	0.44
TRIANGLE T-B-M	99	0.29	0.34		0.73	0.45	0.54
	90	0.25	0.27	0.23	0.42	0.42	0.46
	50	0.22	0.30		-0.17	0.29	0.40

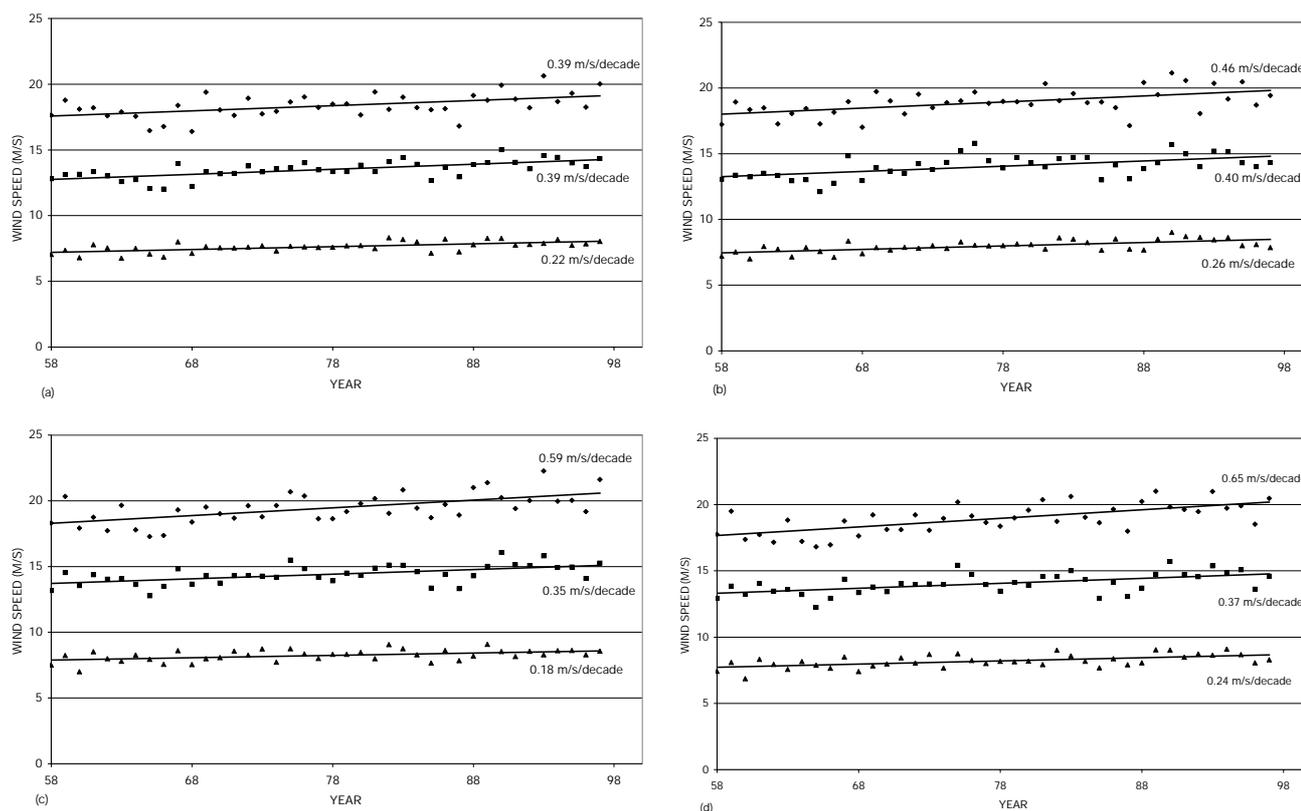


Figure 6 — Climate trends in 50th, 90th and 99th percentile wind speeds at eastern Atlantic pressure triangle locations: (a) triangle T-A-B (GROW); (b) triangle T-A-B (AES40); (c) triangle T-B-M (GROW); (d) triangle T-B-M (AES40).

Table 3—Trends (per cent change/year) in winds and waves at selected locations (1958-1997).

	Per centile	NRA wind	AES40 wind	SHIP wind	NRA wave	AES40 wave
SCOTIAN SHELF	99	0.01	0.07	0.31	-0.01	0.15
	90	0.03	0.09	0.13	-0.02	0.01
	50	0.05	0.10	0.05	0.13	0.20
GRAND BANKS	99	0.10	0.14	0.58	0.17	0.11
	90	0.13	0.18	0.53	0.07	-0.05
	50	0.14	0.17	0.41	0.09	-0.11
BAY OF BISCAY	99	-0.01	-0.03	0.32	0.03	-0.10
	90	-0.01	-0.04	0.16	-0.02	-0.06
	50	-0.02	-0.03	0.25	0.02	-0.02
MID-ATLANTIC	99	0.15	0.01	-0.05	0.17	0.18
	90	0.13	0.11	0.11	0.20	0.16
	50	0.11	0.11	-0.09	0.14	0.08
OWS BRAVO	99	-0.31	-0.64	-0.51	-0.64	-0.01
	90	-0.22	-0.35	-0.20	-0.35	-0.85
	50	-0.30	-0.30	-	-0.80	-0.99
OWS PAPA	99	-0.01	-	-0.94	-0.17	-
	90	-0.02	-	-0.46	0.06	-
	50	0.08	-	-	0.25	-

taken directly from the 40-year NCEP Reanalysis Project, and (2) a fine mesh hindcast of the North Atlantic Ocean based on manual kinematically reanalysed surface wind fields which have been shown to be significantly more representative in storm conditions. Each of these hindcasts has been analysed for trend. Both the global and North Atlantic trend analysis showed statistically significant

areas of both increasing and decreasing winds and waves. The increasing trend in the north-east Atlantic and decreasing trend in the central north Atlantic are particularly well defined and consistent with reported changes in the NAO. Other increasing trends were found in the North Pacific in the global hindcast.

It is essential to verify trends derived from the modelled winds and waves with those computed from long-term homogeneous point time series of measured data. In the absence of such data in the southern hemisphere, we have low confidence in the rather large trends found in parts of the Southern Ocean. In the North Atlantic and North Pacific Oceans, the hindcast winds and waves appear to be affected by a creeping inhomogeneity due to the increased observational density. The global hindcast may be further affected by increasing anemometer heights on board ships, and the increased fraction of measured versus estimated ship winds. Nevertheless, the trends are generally consistent with the analysis of measurements from weather ships, transient ships, and the analysis of pressure triangle-derived geostrophic winds. This implies that the hindcasts may provide a good upper bound to true trends in the wind and wave climate.

REFERENCES

- Bacon, S. and D.J.T. Carter, 1991: Wave climate changes in the North Atlantic and North Sea. *Int. J. Climatol.*, 11, 545-588.
- Bacon, S. and D.J.T. Carter, 1993: A connection between mean wave height and atmospheric pressure gradient in the North Atlantic. *Int. J. Climatol.*, 13, 423-436.
- Bouws, E., D. Jannink and G. Komen, 1996: The increasing wave height in the North Atlantic Ocean. *Bull. Amer. Meteor. Soc.*, 77, 10, 2275-2277.
- Cardone, V.J., J.G. Greenwood and M.A. Cane, 1990: On trends in historical marine wind data. *J. Climate*, 3, 113-127.
- Cardone, V.J., A.T. Cox and V.R. Swail, 2003: Evaluation of NCEP reanalysis surface marine wind fields for ocean wave hindcasts. *Guide to the Applications of Marine Climatology, Part II*. WMO/TD-No. 1081, World Meteorological Organization. Geneva, Switzerland.
- Carter, D.J.T. and L. Draper, 1988: Has the North-East Atlantic become rougher? *Nature*, 332, 494.
- Cox, A.T. and V.R. Swail, 2001: A global wave hindcast over the period 1958-1997: validation and climate assessment. *J. Geophys. Res.*
- Cox, A.T., V.J. Cardone and V.R. Swail, 2003: On the use of in situ and satellite wave measurements for evaluation of wave hindcasts. *Guide to the Applications of Marine Climatology, Part II*. WMO/TD-No. 1081, World Meteorological Organization. Geneva, Switzerland.
- Gulev, S.K. and L. Hasse, 1999: Changes of wind waves in the North Atlantic over the last 30 years. *Int. J. Climatol.*, 19, 1091-1018.
- Kushnir, Y., V.J. Cardone, J.G. Greenwood and M.A. Cane: 1997: The recent increase in the North Atlantic wave height. *J. Climate*, 10, 2107-2113.
- Schmidt, H. and H. von Storch, 1993. German Bight storms analyzed. *Nature*: 365:791.
- Sterl, A., G.J. Komen and P.D. Cotton, 1998: Fifteen years of global wave hindcasts using winds from the European Centre for Medium-range Weather Forecasts reanalysis: validating the reanalyzed winds and assessing the wave climate. *J. Geophys. Res.* 103 (C3), 5477-5492.
- Swail, V.R. and A.T. Cox, 1999: On the use of NCEP/NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. *J. Atmos. Ocean. Technology*, 17, 532-545.
- Wang, X.L. and V.R. Swail, 2001: Changes of extreme wave heights in Northern Hemisphere oceans and related atmospheric circulation regimes. *J. Climate*, 14.
- Wang, X.L. and V.R. Swail, 2002: Trends of Atlantic wave extremes as simulated in a 40-year wave hindcast using kinematically reanalyzed wind fields. Submitted to *J. Climate*.
- Wang, X.L., F.W. Zwiers and H. von Storch, 1999: Using redundancy analysis to improve dynamic seasonal mean 500 hPa geopotential forecasts. *J. Climate*.

- WASA Group, 1998: Changing waves and storms in the northeast Atlantic? *Bull. AMS*, 79(5):741-760.
- White, G. H., 2000: Long-term trends in the NCEP/NCAR Reanalysis. *Proc. Second WCRP International Conference on Reanalyses*. WCRP-109, WMO/TD-No. 95. World Meteorological Organization. Geneva, Switzerland. 54-57.

SECTION 6

USER REQUIREMENTS FOR CLIMATE INFORMATION

Offshore industry requirements and recent metocean technology developments	229
Specific contributions to the observing system: sea surface temperatures . .	234
Importance of marine data to seasonal forecasting in Australia	242

OFFSHORE INDUSTRY REQUIREMENTS AND RECENT METOCEAN TECHNOLOGY DEVELOPMENTS

C.J. Shaw, Chairman OGP¹ metocean committee & Shell Global Solutions², Rijswijk, NL

1. INTRODUCTION

This paper reviews the offshore industry's requirements for metocean data, and describes some recent applications. It also identifies the technology issues which we face in a fluctuating, unpredictable oil price market, and how the industry is addressing them.

2. METOCEAN REQUIREMENTS IN A WORLD OF FLUCTUATING OIL PRICES

Over the last few years, the offshore oil industry has experienced company mergers, together with considerable cutbacks in both budgets and staff as a result of oil prices, which averaged at around US\$ 13/bbl in 1998. Since then, the spot price of crude oil rose briefly to over US\$ 35 and has since fallen back to the mid-US\$ 20/bbl at the time of reviewing this paper. The future price remains uncertain and unpredictable, and in the context of such fluctuations, oil companies have been cautious in firming up budget and resource plans for new projects. Three years after the reality of US\$ 10/bbl oil, the spectre remains. As a result, there is considerable pressure to control operating costs and many companies still suffer from a loss of experienced staff from their global skill pool following downsizing.

Despite the continuing pressure on costs, optimism is slowly reappearing and staff are being sought to resource projects in exciting deepwater areas, particularly offshore West Africa, in the Gulf of Mexico and in areas such as the Caspian Sea. Fortunately, a good description of the metocean environment is required very early in the lifetime of a project, and a pick-up in the industry is quickly reflected in an increase in demand for metocean information. In two fairly recent cases, we have made use of deep water current data collected in the very early stages of exploration, to aid decision making: in one case, the selection of the rig, and in the other, the use of riser-fairings. In both cases, decisions needed to be made on whether or not to spend many millions of dollars while the costs involved in collecting the appropriate data were several orders of magnitude lower.

The value of metocean data can be increased through multiple applications - often there is more than one engineering application (e.g. engineering design, operations planning, iceberg management, etc.) for one data set. Similarly, there is frequently more than one set of data needed to achieve significant cost benefits. For example, it is necessary to have a combination of metocean data, the appropriate database facilities and the analysis tools, say for deriving response-based design criteria, in order to produce figures that engineers can use.

As well as an understanding of the effects of the environment on the structure, it is important to be aware of, and able to quantify through measurements, any likely impact of the structure (or operation) on the environment itself. A benchmark survey of the environment is needed before a drilling-rig arrives in a new area in order to measure the undisturbed condition. For this, a preliminary knowledge of the metocean conditions will lead to a better designed, more cost-effective and thorough Environmental Impact Assessment (EIA) of the area.

¹ OGP is the abbreviation for the International Association of Oil and Gas Producers, 25-28 Old Burlington St, W1S 3AN, United Kingdom

² Shell Global Solutions is a network of technology companies of the Royal/Dutch Shell Group and provides technical services and consultancy to customers within and outside of Shell

3. **COST SAVINGS** Some examples of the benefits of the expeditious use of metocean data are listed in Table 1. Of course, not every field development will benefit to the same extent, but, nevertheless, these examples show at least the potential for making better use of the information.

4. **METOCEAN TECHNICAL REQUIREMENTS** It is not necessary in every case to collect field data at a new site of interest. The meteorological and oceanographic variables are in many cases spatially homogeneous over quite a large area. It is also possible to extend the range of existing measured data sets into new areas through the use of calibrated numerical wind, wave and current models (normally referred to as hindcast studies). This is a very cost-effective way of increasing the value of existing measurements. Even for sites where, for example, two to three years of measured data sets are available, properly verified hindcast models can provide a means of extending the data sets to, say, a 25-year period, thereby significantly improving an estimate of the extreme values.

BALANCING TECHNICAL, SAFETY AND FINANCIAL NEEDS Despite the optimistic view outlined above, there are many arguments for not investing in a data collection programme; some of the pros and cons are given in Table 2.

<i>Activity</i>	<i>Type of saving</i>
Construction of new offshore structures translates into a 10 per cent	A 5 per cent reduction in design wave height reduction in steel costs. Savings are ~ 10:1.
On existing structures	Reduces the need for repairs or strengthening and increases the opportunity to add topsides weight for satellite developments. Savings vary from about 5:1 to 100:1 in the case that a completely new platform was required.
On new pipelines	Savings can be achieved on: - The design of the pipe and the near shore installation method. - The amount of concrete coating required in the deeper offshore sections as well as on the class of pipelay-barge needed for installation of the pipeline.
Installation of jackets	During installation, real-time directional wave data (including spectra) are used to plan and align the topsides facilities relative to the jacket.
On existing pipelines	Reduces the need for remedial action (e.g. spans).
Selecting Jack-ups, semi-subs or barges	A cheaper option becomes feasible.
In deep water	Correctly assessing the strength of the current and its direction relative to the waves makes it possible to assess the downtime for production facilities. If incorrectly assessed, significant loss of production may result.
On the Continental Shelf Edge, currents play a significantly more important role in defining development concepts	
Jack-up/semi/barge selection	A cheaper option becomes feasible.
All offshore operations	Measured field data are used to update and improve weather forecasts. Many operational decisions are made on the basis of the forecasts.
But	Increases in criteria which cost money in the short term, can save money in the long-run by protecting the investment. Under-design of facilities is not acceptable.

Table 1

Table 2

<i>Some arguments against an investment in collecting metocean data</i>	<i>The main arguments for an investment in collecting metocean data</i>
<p>We need to think several years ahead in order to have a reasonable chance of getting any return on your investment. The need for data may be realized too late to be of value to a project.</p> <p>The project engineer will probably have moved to a new job by the time any savings materialize in the project. There is also a reasonable chance that the project will be cancelled before it is possible to make any use of the data (let alone make any cost-savings).</p> <p>We probably have to spend several hundred thousand dollars to have any chance of making some savings.</p>	<p>If the project goes to maturity, the return on investment is between 10 and 100:1.</p> <p>If we do not have the data, we do not know how to design your facilities to the appropriate functional and safety standards, with the result that we either spend extra money unnecessarily or the design is inadequate.</p>

For obvious safety reasons, the offshore industry will avoid under-designing an offshore installation, hence the price paid for not having the right design data available when it is needed is essentially the cost of over-designing facilities for the future. However, these arguments must be balanced by considering the use of design criteria which are deliberately conservative for today's solution, but which will allow additional facilities to be added to the structure in the future without the need, and associated high costs, for major offshore structural modifications.

To make a proper comparison of cost benefits, we should take into account the discounted savings over the time period before the savings can actually be realized (which may be several years). This means that a strong argument is needed to persuade people to spend money now rather than save the discounted cash at a later date. Appeals to reason, supported by estimates of the costs of collecting the data, the potential savings and the consequences of not having the data available when the engineers need it, are usually the best approach.

SAFETY ISSUES FOR DESIGN CRITERIA

The approach to deriving criteria which might be used for design purposes is to take note of the uncertainty in the criteria which is caused by the lack of data in the early stages of the life cycle of a field development. This is shown schematically in Figure 1. The aim is to ensure that the criteria can be reduced as more data become available; the corollary is to ensure that the criteria are conservative when there is little data on which to base them. With experience, and by comparison with other areas where there are good quality data sets, this can usually be achieved - but it is essential to be aware of the costs associated with using criteria which may be too conservative.

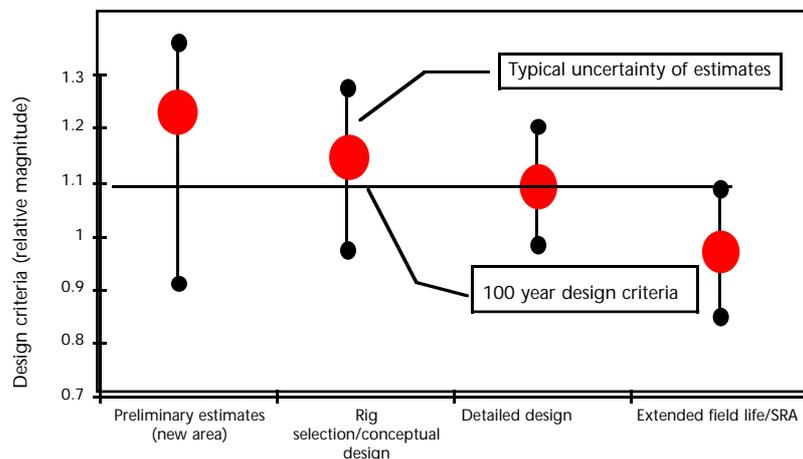


Figure 1—Design criteria according to the length of data sets.

Progress has been made at the front end of the life cycle of a field development; it is now feasible to make use of wave (and to a limited extent wind) data received from remote-sensing satellites which have on-board radar altimeters and scatterometers. The instruments produce their best data sets in remote deep water areas far from the coast and at high latitudes. Unfortunately, there are not yet enough of them to provide coverage that is sufficiently dense to be able to form a satisfactory means of deriving criteria in low latitudes subject to relatively small scale tropical storms. However, it does mean that in areas where they are able to provide good quality data, the conservatism can be confidently removed at a relatively early stage of the field development life cycle.

The uncertainty in the estimate of the 100-year extreme is principally due to the length of the data set. For design purposes (in which the '100-year' return period parameters are required), a hindcast period of 25-30 years is normally considered adequate - beyond this length, the return on the effort invested does not normally warrant the cost involved. However, an increasing number of applications (e.g. for platform reliability studies or minimum deck elevation studies) require estimates of parameters with very low probabilities of occurrence ($\sim 10^{-4}$). In these cases, either the data set needs to be increased substantially, or other more imaginative approaches are needed. However, this does assume that the underlying physical processes are stationary. If it is suspected that climatic cycles are significant and may be influencing the data sets, a much longer data set may be required to identify the period of the cycles and quantify their effect on the extreme values.

ROLE OF REGULATORS

In states where they exist, regulators can play an important role in ensuring that adequate quantities of the right data are collected to ensure that justifiable reliability levels for offshore structures are maintained. If there is a regulation stating that certain data should be collected, this can often ease the process of getting adequate funding at the right time. Most operators are satisfied with a good balance of regulations which enable 'fit-for-purpose' structures to be installed and hence adequate metocean data to be collected.

5. JOINT INDUSTRY PROJECTS

Prior to the recent trend within the industry for company takeovers, significant cost-savings could be made by individual companies if a number of them working in a region or on a particular technology item, decided to work together in joint industry projects (JIPs). In addition to saving costs, the risks (e.g. of losing equipment and data) were also shared. In practice, this meant that additional funds became available to enhance the technical content of the programme and provide resources for improving the safety of the equipment and ensuring good return of the data. Since metocean data is rarely considered to be strategically sensitive or confidential, it has not been difficult to obtain an agreement to join forces in collecting such data. The only real disadvantage is that it takes longer to get the technical scope of work defined and the several contracts agreed and signed.

However, owing to recent company takeovers, the number of companies now available to form JIPs continues to reduce and consequently the cost-saving per participant reduces. For metocean engineers, this has resulted in difficulties in obtaining start-up funds for new JIPs and, inevitably, delays in the projects. Nevertheless, several new regional projects are under consideration and we expect they will be able to start as optimism within the industry continues to improve.

The offshore metocean industry has been successfully instigating joint industry projects for many years, the main vehicle for establishing such collaborative efforts being the International Association of Oil and Gas Producers (OGP). Some examples of recent regional cooperative projects are:

- The Gulf of Mexico Storm Hindcast of Oceanographic Extremes (GUMSHOE).
- The North European Storm Study (NESS/NEXT) - a major hindcast study of north-east European waters of wind, wave and current conditions in storms, as well as over several continuous years.
- The South East Asia Meteorological and Oceanographic Study (SEAMOS) - a wind and wave hindcast study of typhoon and monsoon conditions in the South China Sea.

- The West Africa eXtremes study (WAX) - a wind and wave hindcast study of storms in the South Atlantic which produce primarily swell conditions offshore west Africa. The study has recently been supplemented with the results from a current model study.
- The Caspian Sea Metocean Study (CASMOs) - a proposal to use a hindcast model to derive wind wave and current conditions for the Caspian Sea.

Uncertainty and fluctuations in the oil price may persuade some oil companies to focus on in-house technology developments which have a reasonable chance of delivering results in the short term. There may also be a tendency to farm out non-strategic, longer-term technology into joint industry projects - where this is feasible.

Some of the technology projects in which the industry is involved are:

- SAFETRANS (Definition of criteria for long ocean tows).
- WACSIS (WAVE Crest Sensor Intercomparison Study).
- 'Response'-based design criteria for offshore structures (fixed, floating and pipelines).
- Maximum wind gust speeds in squalls and tropical cyclones.
- 10-4 (developing a better definition of very low probability events).
- ISO (development of international standards for the design of offshore structures (fixed steel, concrete, mobile jack-ups and floating systems)).

There are often synergies between these projects. For example, during work on the ISO code, setting minimum deck elevations was identified as an area of major uncertainty, and it was recognized that resolving the issue required a better understanding of two aspects of the problem. Firstly, we needed to know the distribution of crest heights in storm conditions (and consequently an understanding of which data sets we could consider reliable). Secondly, the new procedure for deck heights requires extrapolation of the distribution of crest heights to very low probabilities (return periods of 10,000 years). Hence, the ISO work led to the WACSIS and 10-4 projects, the combined aims of which are to define a rational procedure, for use anywhere in the world, for defining the minimum air-gaps for fixed structures.

6. DATA MANAGEMENT

6. Enormous quantities of data are produced nowadays (both directly from instruments or from numerical models) and we do not underestimate the task of managing the metocean data once it has been collected. Managing data includes not only storing it, but also being able to access it quickly and efficiently once it has been collected. Effective access to our data is not only a prerequisite of enabling the 'value' to be extracted from the investment, but it is also a necessary form of safeguarding the investment so that further value can be extracted in the future. Although the power of computer hardware is increasing substantially, building and maintaining the appropriate software is still a significant cost item.

7. CONCLUSIONS

7. There is a steady continuation of interest in the collection of metocean data and in continuing development of metocean technology. Certainly the funds presently available are smaller than in the past, and time-scales to get projects off the ground are longer; however, it is still recognized that metocean technology can be a key contributor to developing cost-savings while visibly maintaining safety standards.

SPECIFIC CONTRIBUTIONS TO THE OBSERVING SYSTEM: SEA SURFACE TEMPERATURES

Richard W. Reynolds*, National Climatic Data Center, NESDIS, Camp Springs, Maryland, USA

1. INTRODUCTION

Sea surface temperature (SST) analyses are an important indicator of the coupling between the atmosphere and the ocean and may be the most important field for climate modelling. They are used for climate monitoring, prediction and research, as well as specifying the surface boundary condition for numerical weather prediction, and for other atmospheric simulations using atmospheric general circulation models. The purpose of this paper is to present the current and future status of SST data and SST analyses.

2. SST DATA

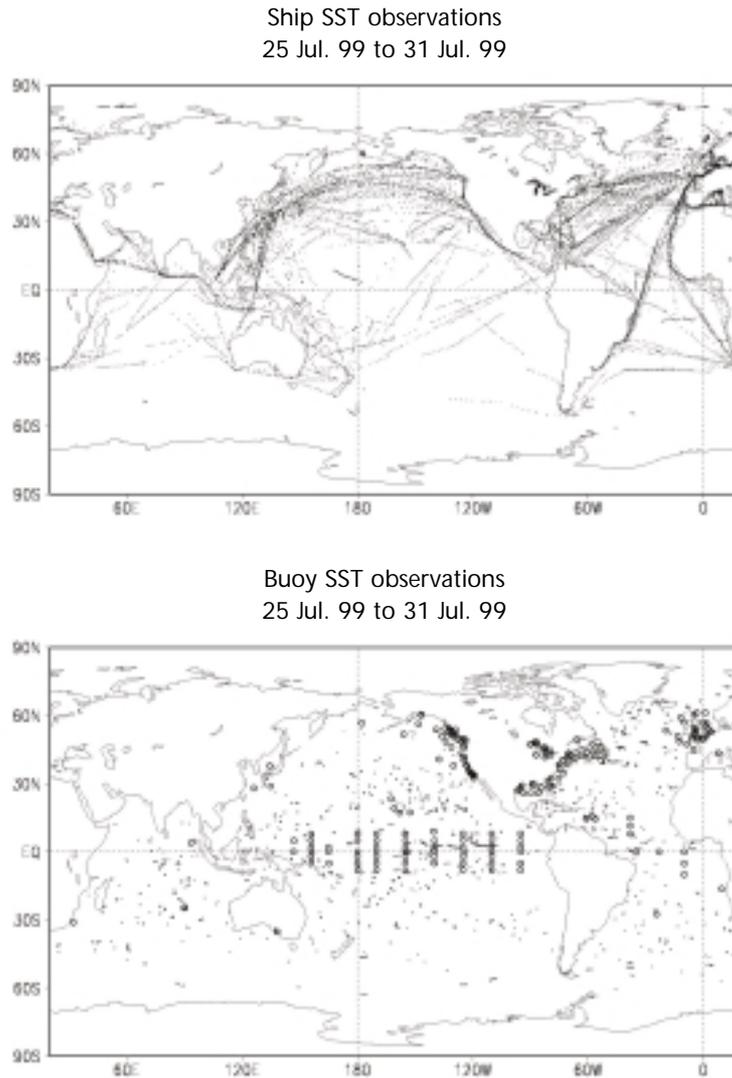
The longest data set of SST observations is based on observations made from ships. These observations include measurements of SST alone as well as temperature profiles with depth. However, the observations of SST alone dominate the data sets and account for more than 90 per cent of the observations. Although the earliest observations were taken in the first half of the 19th century, sufficient observations to produce a global SST analysis were not available until about 1870. From 1870 to present, the number of observations generally increased except for noticeable dips during the First and Second World Wars. In addition to the changes in the number of observations, the method of measuring surface marine observations changed over the period from temperatures measured from uninsulated buckets to temperatures measured from insulated buckets and engine intakes. These instrument changes resulted in biases in the data set, the corrections for which are discussed by Folland and Parker (1995) and incorporated into UK Met Office SST analyses. Although, as discussed in Kent *et al.* (1993), selected SST observations can be very accurate, typical RMS errors from ships are larger than 1°C and may have daytime biases of a few tenths of a degree C (Kent *et al.*, 1999).

SST observations from drifting and moored buoys began to be plentiful in the late 1970s. These observations are typically made by a thermistor or hull contact sensor and are usually relayed in real time by satellites. Biases in the SSTs from buoys can occur in some designs; for example, significant diurnal heating of the hull may occur under low wind conditions with some hull configurations. Although the accuracy of the buoy SST observations varies, the accuracies are usually better than 0.5°C, which is better than ships. In addition, typical depths of the observations are roughly 0.5 m rather than the 1 m and deeper depths from ships. The distribution of ship and buoy in situ SST observations (see Figure 1) shows that the deployment of the buoys has partially been designed to fill in some areas with few ship observations. This process has been most successful in the tropical Pacific and southern hemisphere.

In late 1981, accurate SST retrievals became available from the Advanced Very High Resolution Radiometer (AVHRR) instrument which has been carried on many NOAA polar-orbiting satellites. These retrievals improved the data coverage over that from in situ observations alone. The satellite retrievals allowed better resolution of small-scale features such as Gulf Stream eddies. In addition,

* Corresponding author's address: Mr Richard W. Reynolds, National Climatic Data Center, NESDIS/NOAA, 5200 Auth Road, Room 807, Camp Springs, MD 20746; e-mail: rreynolds@ncep.noaa.gov

Figure 1 — Distribution of SST in situ observations from ships (top panel) and buoys (lower panel) for the week of 25–31 July 1999.



especially in the southern hemisphere, SSTs could now be observed on a regular basis in many locations. These data are produced operationally by NOAA's National Environmental Satellite, Data and Information Service (NESDIS) and also, during the last few years, by the US Navy.

Because the AVHRR cannot retrieve SSTs in cloud-covered regions, the most important problem in retrieving SST is to eliminate clouds. The cloud clearing algorithms are different during the day and at night because the AVHRR visible channels can only be used during the day. After clouds have been eliminated, the SST algorithm is derived to minimize the effects of atmospheric water vapour. The satellite SST retrieval algorithms are 'tuned' by regression against quality-controlled drifting buoy data using the multichannel SST technique of McClain *et al.* (1985). This procedure converts the satellite measurement of the 'skin' SST (roughly a micron in depth) to a buoy 'bulk' SST (roughly 0.5 m). The tuning is carried out when a new satellite becomes operational or when verification with the buoy data shows increasing errors. The AVHRR instrument has three infrared (IR) channels. However, because of noise from sun glint, only two channels can be used during the day. Thus, the algorithm is usually tuned separately during the day and at night and typically uses three channels at night and two during the day (Walton *et al.*, 1998). The algorithms are computed globally and are not a function of position or time.

If the retrievals are partially contaminated by clouds, they have a negative bias. Negative biases can also be caused by aerosols, especially stratospheric aerosols from large volcanic eruptions (for example, see Reynolds, 1993). The ratio of the number of daytime to night-time satellite retrievals is now roughly one to

one. However, the ratio was roughly five to one prior to 1988. From 1989 to present, the night-time satellite algorithm was gradually modified to increase the number of night-time observations, while the daytime observations remained roughly constant. A reanalysis of the satellite data (now being completed by the Pathfinder project) would correct these differences and should be a better product for climate.

Future improvements in the SST observing system will primarily be due to new satellite data. A significant change occurred in 1999 when SSTs from a second polar-orbiting NOAA satellite were operationally processed for the first time. In addition, data from other satellites, including microwave satellites, which can see through clouds, and geostationary satellites, which can resolve the diurnal cycle, are now becoming available. This will make it easier to carry out high resolution SST analyses, as discussed later.

3.
CLIMATE SCALE SST
ANALYSES

For the purpose of this discussion, SST analyses have been divided into two groups: climate and high resolution. The climate scale analysis typically has temporal resolutions from weekly to monthly and spatial resolutions from 1° to 5°. These analyses use in situ SST data and may, or may not, use satellite SST data when available. As mentioned below, sea-ice concentrations may also be used to augment the SST data at high latitudes. These analyses are often used on seasonal and interannual scales for the monitoring and prediction of El Niño events, and on decadal and centennial scales for climate trend detection. In addition, the SSTs are used as the ocean boundary condition for atmospheric general circulation models. For these purposes, it is important that analysis methods be constant with time and not influenced by temporal changes in SST data. The latter is particularly difficult because not only did the number of in situ data generally increase with time, but additional data sources were added when observations from buoys, and of course satellites, became available.

To better understand the problems of climate scale SSTs, different SST analyses have been compared. Two studies will be discussed here. Hurrell and Trenberth (1999) compared four analyses: the National Center for Environmental Prediction (NCEP) optimum interpolation analysis, henceforth OI, of Reynolds and Smith (1994); the NCEP empirical orthogonal functions analysis, henceforth EOF, of Smith *et al.* (1996); the UK Meteorological Global Sea-ice SST analysis, version 2.3b, henceforth (GISST), of Rayner *et al.* (1996); and the Lamont-Doherty Earth Observatory analyses, henceforth LDEO, of Kaplan *et al.* (1998). A description of the data and analysis methods can be found in Hurrell and Trenberth (1999). The second study was presented at a Global Climate Observing System (GCOS) Workshop on Global Sea Surface Temperature Data Sets, held at the Lamont-Doherty Earth Observatory on 2-4 November 1998, and is updated here. This workshop study focused on the 1982-1997 period and added four additional analyses: the UK Met Office Historical SST analysis, version 6, henceforth MOHSST, of Parker *et al.* (1994); the Japan Meteorological Agency (T. Manabe, 1999, personal communication), henceforth JMA; the Naval Research Laboratory (J. Cummings, 1999, personal communication), henceforth NRL; and the Australian Bureau of Meteorology Research Centre (N. Smith, 1999, personal communication), henceforth BMRC. The resolution, period, and type of SST data used for each analysis are summarized in Table 1.

Table 1—SST analyses with analysis periods and resolution. All analyses used in situ (ship and buoy) data. Analyses using sea-ice data converted to SSTs are indicated by “yes” in the ice column. Analyses using satellite data are indicated by “yes” if used, or “corrected” if used with additional bias corrections. Months are noted under the “period” column if the analysis did not start in January.

Acronym	Period	Resolution	Satellite data	Ice data
BMRC	Jul.-93 to present	1°	Corrected	Yes
GISST	1871 to present	1°	Corrected	Yes
JMA	1982 to present	2°	No	No
LDEO	1856 to present	5°	No	No
MOHSST	1856 to present	5°	No	No
EOF	1950 to 1998	2°	No	No
OI	Nov.-81 to present	1°	Corrected	Yes
NRL	1995 to present	1/4°	Yes	Yes

Sea-ice information is used to generate additional SST data to augment other SST data in four of the analyses. The generation methods vary along with the accuracy of the sea-ice information. In the OI, BMRC and NRL analyses, an SST value representing the freezing point is added at locations where a specified sea-ice concentration is exceeded. The GISST method of generating SST from the sea-ice concentration, I , is more complicated and probably more realistic. In this method, a relation between SST and I is defined by a quadratic equation: $SST = aI^2 + bI + c$, where a , b and c are constants determined by climatological collocated match-ups between SST and sea-ice concentration, with the constraint that $SST = -1.8^\circ\text{C}$ or 0°C when $I = 1$ over the ocean or fresh water lakes, respectively. In addition to uncertainties in these methods, the analysed value of ice concentration as defined in different analyses is not accurately known, especially in summer. The climatological sea-ice concentrations are shown for July in Figure 2 for two analyses. The first, combined from Nomura (1995) and Grumbine (1996), the Nomura/Grumbine analysis, is an objective analysis of microwave satellite observations (SMMR and SSM/I); the second, the National Ice Center analysis (Knight, 1984), is a subjective analysis of in situ and satellite microwave and infrared observations. The concentrations of the Nomura/Grumbine analysis are much lower because the microwave satellite instrument interprets melt water on top of the sea ice as open water.

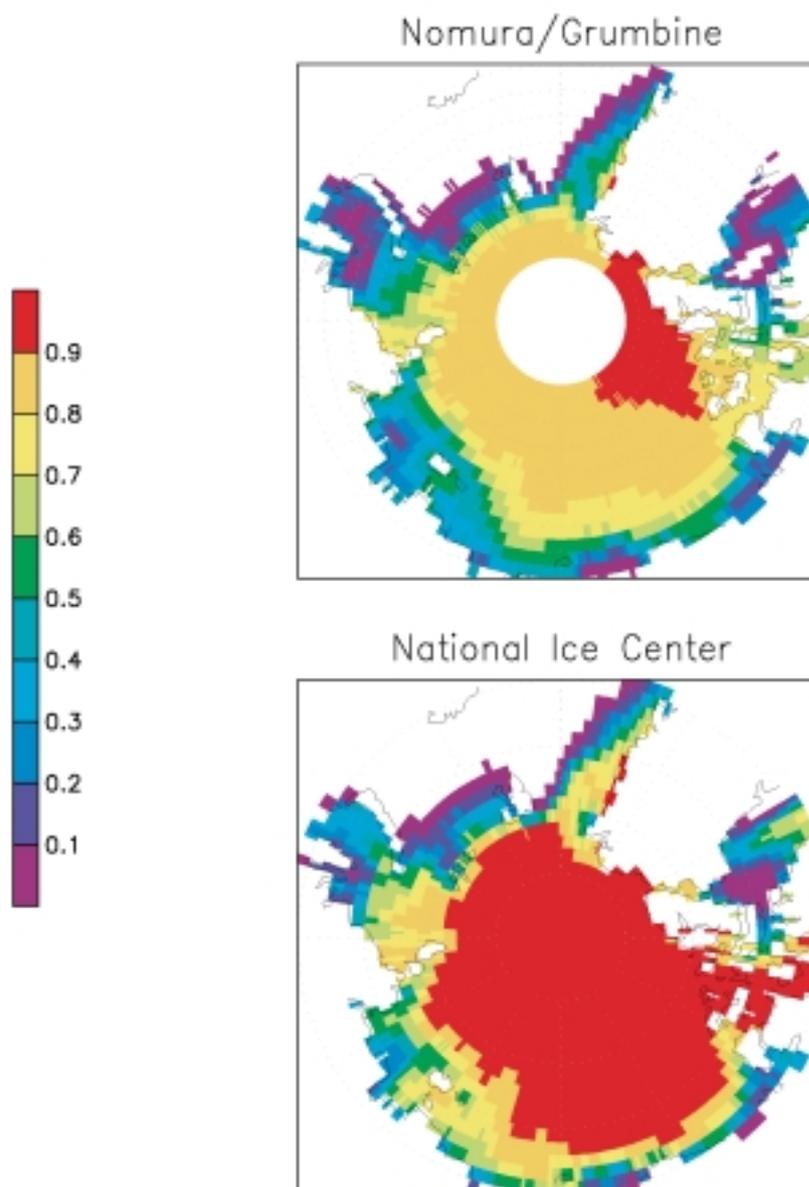


Figure 2— Climatological sea-ice concentrations for the Arctic for July for the 1979-1992 period.

The upper panel shows the analysis from Nomura and Grumbine; the missing data near the pole occurs because of a lack of satellite observations. The lower panel shows the analysis from the National Ice Center (see text). The range of ice concentration is 0 (0%) to 1 (100%).

Both Hurrell and Trenberth (1999) and the workshop comparisons showed that differences among analyses were smaller within the tropics than the extratropics. This can be seen in the zonal averages shown for the four analyses with ice information in Figure 3. The figure shows that northern hemisphere middle latitude differences are smaller than middle and high latitudes differences in the southern hemisphere. However, the differences above 60°N are the largest due to uncertainties near, and within, the Arctic sea ice. The workshop comparisons found that the monthly RMS differences among analyses were within the range 0.2°C to 0.5°C between roughly 40°S and 60°N, except in coastal areas; they were larger outside this latitude belt. In particular, in situ only analyses had differences greater than 1°C south of 40°S. Hurrell and Trenberth (1999) showed that monthly lag one autocorrelations appeared to be depressed in the GISST analysis during 1982-1997 compared to the other analyses. In addition, they found differences in the regional trends between the GISST and LDEO. LDEO used MOHSST, version 5, and GISST used MOHSST, version 6, as in situ input data. Thus, the differences may be due to changes in MOHSST or differences in the analysis methods.

The comparisons have shown that analyses using satellite data without careful bias correction should not be used for climate studies because of large potential biases in satellite retrievals. Satellite data can improve the coverage and spatial resolution of SST analyses and should be used with bias corrections. The results also suggested that although real-time bias corrections were successful, a small persistent negative residual satellite bias of approximately 0.1°C often remained. These biases occurred primarily in the mid-latitude southern

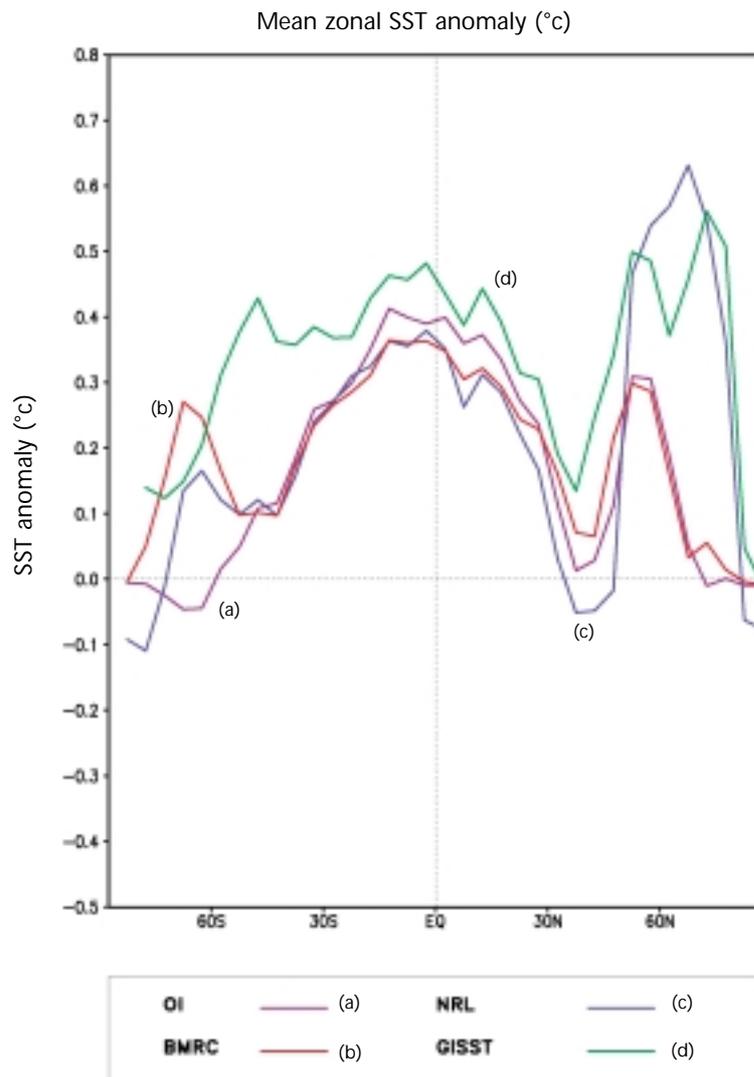


Figure 3— Mean zonally averaged SST anomalies from four analyses for the period January 1995 to December 1997. All analyses used in situ and satellite SST plus SSTs generated from sea-ice concentrations.

hemisphere where in situ observations were sparse. However, there were also large-scale differences among the in situ analyses of this magnitude which could persist for several months. These differences are most likely due to the nonlinear data procedures used to eliminate bad data rather than differences in the in situ data sets themselves. The largest differences among analyses with sea-ice data occurred near the sea-ice margins. The differences were due both to uncertainties in the ice analyses, as well as uncertainties in the method of converting from ice to SST.

4. HIGH RESOLUTION SST ANALYSES

High resolution SST analyses have spatial scales of 1° or higher and temporal scales of 24 hours or less. They have the same potential problems as those discussed for the climate SST analyses. However, the high resolution analyses have additional problems because the data are now relatively sparser, primarily because of shorter analysis periods. Satellite data are essential for these analyses.

In regions with light winds and strong net heat fluxes into the ocean, diurnal SST signals of several degrees C can occur. This signal may be very close to the surface and may not reach typical in situ observation depths. This problem is further complicated by satellite SSTs which measure a skin temperature which is typically 0.3°C colder than the layer immediately below the skin (see Webster *et al.*, 1996 for details). The tuning of the MCSST algorithm is based on assumed correlations of the skin and the bulk SST. This assumption begins to break down during the daytime when a diurnal signal is present in the SSTs. This problem is illustrated in Figure 4, which shows skin and bulk SSTs at a buoy deployed in light winds of the western tropical Pacific (Weller and Anderson, 1996). The upper panel shows the diurnal average; the lower panel shows a sample of the day-to-day variability. The differences caused by the potential decoupling of skin and

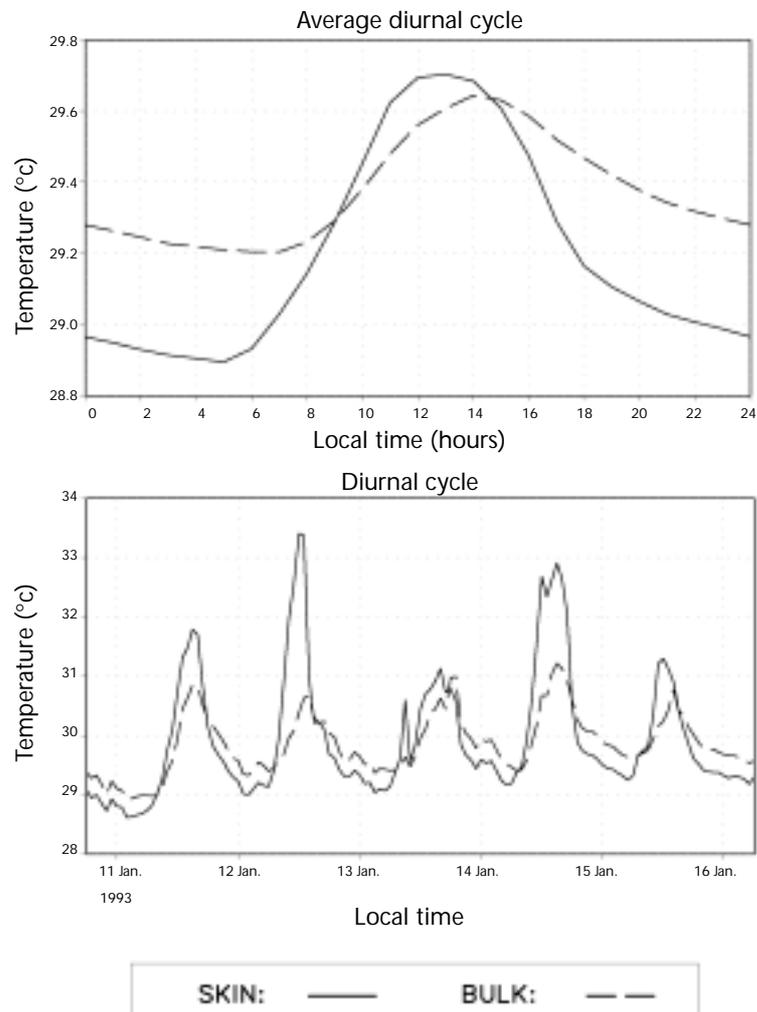


Figure 4— Skin and bulk SSTs (see text) from a buoy at 1.8°S and 156°E . The top panel shows the average diurnal cycle for the period 22 October 1992 to 3 March 1993. The bottom panel shows the variability in the diurnal cycle. In the bottom panel, the data labels indicate local midnight.

bulk SSTs are minimized by smoothing and by increasing the error statistics of day satellite SSTs relative to night. However, for high resolution SSTs, the vertical structure of the depth of the different observations must be properly resolved.

The satellite data used in the SST analyses listed in Table 1 are derived from the AVHRR instrument. Although there were two polar-orbiting satellites for most of the 1982-99 period, data were operationally processed from only one satellite until late spring 1999. Because of swath width limits, one satellite cannot see the entire globe twice a day. This problem is made worse by cloud cover, which further degrades the coverage. Thus, only analyses with a dynamical component may be able to properly interpolate the analysis in space and time.

This data coverage problem will become less critical when more satellite data become available. Accurate SSTs from a microwave instrument, for example, Tropical Rainfall Measuring Mission (TRMM), would produce SSTs which are unaffected by cloud cover. In addition, SSTs from US Geosynchronous Operational Environmental Satellites (GOES) are now available (Wu, 1999). The GOES instrument is similar to the AVHRR and can resolve the diurnal cycle in cloud-free areas. However, further research is needed to improve the retrievals, as discussed by Wick (1999). In addition, future GOES SST retrievals will be degraded because of instrument changes which make the correction for atmospheric water vapour more difficult.

Some improvements in the in situ data must also be made. Many of the open ocean buoys do not report SSTs at six-hour intervals so as to save on satellite transmission costs. For example, the TAO network of moored buoys in the tropical Pacific (McPhaden, 1995) would be ideal for determining the diurnal cycle if all the data collected by the buoys were available in real time. Metadata information on the characteristics of both ship and buoy SSTs is also needed to better define error characteristics so that better use can be made of the in situ data. In addition, more ship and buoy data are required south of 45°S where there are currently insufficient in situ data to completely correct any satellite biases.

5. CONCLUSIONS

For both climate and high resolution SST analyses, satellite data should be used with care. These data can greatly improve the coverage and spatial resolution of SST analyses. However, because of large potential biases in satellite retrievals, accurate bias corrections are needed, particularly for climate studies. For climate purposes, reliance on in situ data alone does not eliminate SST analysis differences. A careful intercomparison of the in situ data processing methods is needed to develop more uniform procedures. Because of large uncertainties in present ice analyses and the methods of converting from ice to SST, in situ observations of both SSTs and sea-ice concentrations are urgently needed near the ice.

For high resolution SST analyses, the use of accurate satellite data from multiple sensors, including microwave and geostationary instruments, is critical. In addition, dynamic models are needed to interpolate in both space and time in regions where SST data are missing. These models must include the resolution of vertical scales so that the differences in SST measurements from ships, buoys and satellites can be assimilated at the depths where the observations are made.

Intercomparisons of different SST products have shown important differences. It is important that SST intercomparisons continue so that analysis and data differences can be better quantified and methods can be developed to minimize these differences. Because analyses continue to change, a continued reevaluation of the differences is required. An international GCOS working group has been established by the Atmospheric Observation Panel for Climate (AOPC) and the Ocean Observations Panel for Climate (OOPC) to evaluate climate SST products. This effort should be extended to include high resolution SSTs analyses. A parallel effort may be needed to include comparisons of high resolution SSTs analyses.

REFERENCES

- Folland, C.K. and D.E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, 121, 319-367.
- Grumbine, R.W., 1996: Automated passive microwave sea ice concentration analysis at NCEP, unreviewed manuscript, 13 pp. (NCEP/NWS/NOAA, Camp Springs, MD, USA).

- Hurrell, J.W. and K.E. Trenberth, 1999: Global sea surface temperature analyses: multiple problems and their implications for climate analysis, modeling and reanalysis. *Bull. Amer. Met. Soc.*, 80, 2661-1678.
- Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856-1991. *J. Geophys. Res.*, 103, 18567-18589.
- Kent, E.C., P.K. Taylor, B.S. Truscott and J.A. Hopkins, 1993: The accuracy of Voluntary Observing Ship's meteorological observations, *J. Atmos. & Oceanic Tech.*, 10, 591 - 608.
- Kent, E.C., P.G. Challenor and P.K. Taylor, 1999: A statistical determination of the random observational errors present in Voluntary Observing Ships meteorological reports, *J. Atmos. & Oceanic Tech.*, 16, 905-914.
- Knight, R.W., 1984: Introduction to a new sea-ice database. *Annals of Glaciology*, 5, 81-84.
- McClain, E.P., W.G. Pichel and C.C. Walton, 1985: Comparative performance of AVHRR-based multichannel sea surface temperatures. *J. Geophys. Res.*, 90, 11587-11601.
- McPhaden, M.J., 1995: The tropical atmosphere-ocean array completed. *Bull. Amer. Met. Soc.*, 76, 739-741.
- Nomura, A. 1995: Global sea ice concentration data set for use in the ECMWF Re-analysis system, re-analysis project, No. 76 (ECMWF, Reading, Berkshire, UK).
- Parker, D.E., P.D. Jones, C.K. Folland and A. Bevan, 1994: Interdecadal changes of surface temperature since the late 19th century. *J. Geophys. Res.*, 99, 14377-14399.
- Rayner, N.A., E.B. Horton, D.E. Parker, C.K. Folland and R.B. Hackett, 1996: Version 2.2 of the global sea-ice and sea surface temperature data set, 1903-1994, Climate Research Technical Note CRTN 74, 43 pp. (The Met Office, London Road, Bracknell, UK).
- Reynolds, R.W., 1993: Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate*, 6, 768-774.
- Reynolds, R.W. and T.M. Smith, 1994: Improved global sea surface temperature analyses. *J. Climate*, 7, 929-948.
- Smith, T.M., R.W. Reynolds, R.E. Livezey and D.C. Stokes 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, 9, 1403-1420.
- Walton, C.C., W.G. Pichel, J.F. Sapper and D.A. May, 1998: The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. *J. Geophys. Res.*, 103, 27,999-28,012.
- Webster, P.J., C.A. Clayson and J.A. Curry, 1996: Clouds, radiation and the diurnal cycle of sea surface temperature in the tropical western Pacific Ocean. *J. Climate*, 9, 1712-1730.
- Weller, R.A. and S.P. Anderson, 1996: Surface meteorology and air-sea fluxes in the western equatorial Pacific warm pool during the TOGA Coupled Ocean-Atmosphere Response Experiment. *J. Climate*, 9, 1959-1990.
- Wick, G.A., J.J. Bates and D.J. Scott, 1999: Accurate measurement of sea surface temperature in the Americas from geostationary satellites. Third Symposium on Integrated Observing Systems, American Meteorological Society Annual Meeting, Dallas, TX, pp. 135-138.
- Wu, X., P. Menzel, G.S. Wade, 1999: Estimation of sea surface temperatures using GOES-8/9 radiance measurements. *Bull. Amer. Met. Soc.*, 80, 1127-1138.

IMPORTANCE OF MARINE DATA TO SEASONAL FORECASTING IN AUSTRALIA

Scott Power*, Dean Collins and Grant Beard, National Climate Centre, Bureau of Meteorology, Australia

1. INTRODUCTION

Each month, Australia's Bureau of Meteorology issues a seasonal climate outlook of rainfall outlooks throughout the country for the coming three months (see <http://www.bom.gov.au/au/climate/ahead> for further details). This site depicts a small but important subset of the forecast information we provide. A much larger set is provided on a cost-recovery basis to hundreds of subscribers from around the country. Forecast information is also disseminated via the mass media through the radio, newspapers (national, state and rural), and more recently via satellite television to over 300,000 subscribers in rural Australia, on a weekly basis.

These outlooks, together with estimates of their reliability, are useful to a wide range of users, including farmers, water managers, banking groups, and scientists, and many other users connected with the rural sector.

The outlooks are based on the statistical relationships between rainfall and patterns of sea surface temperature (SST) anomalies in the region (Drosdowsky and Chambers, 1998). A similar scheme for Australian temperature forecasts has also been developed (Jones, 1998). Over the longer term, we also provide forecasts of NINO3 out to nine months using an intermediate coupled model (Kleeman, 1993). The Bureau is also moving towards forecasts based on coupled general circulation models (Power *et al.*, 1998; Wang *et al.*, 1999). In this paper we will outline the fundamental importance of marine data sets for both the statistical and numerical prediction schemes.

In addition, we will describe research which suggests that the predictability of seasonal rainfall anomalies over Australia associated with ENSO, waxes and wanes from generation to generation and this variability may be associated with the "Interdecadal Pacific Oscillation" (Power *et al.*, 1999). Additional marine observations may be required to ensure that variability of this kind is adequately represented in future coupled models used for seasonal forecasting.

2. STATISTICAL PREDICTION SCHEMES BASED ON SEA SURFACE TEMPERATURE

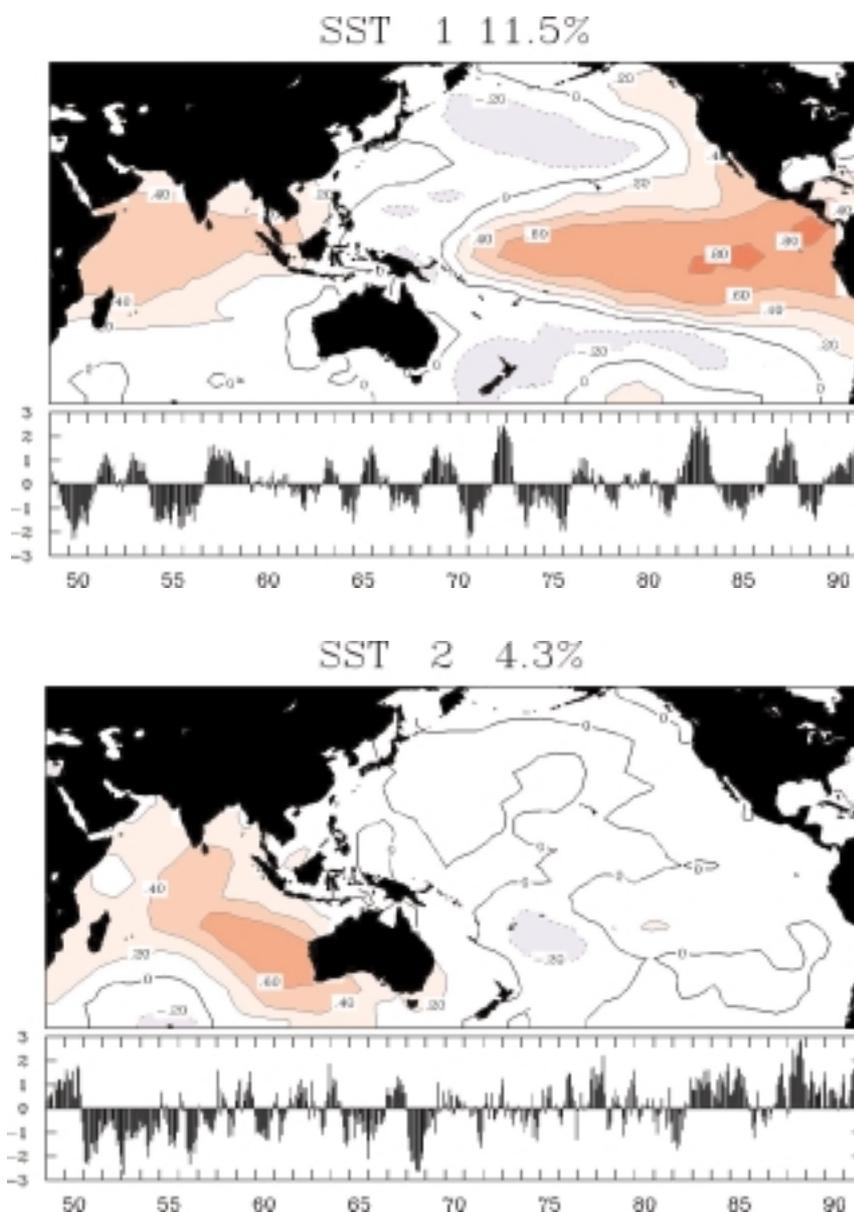
The operational scheme currently used by the Bureau's National Climate Centre for rainfall prediction (Drosdowsky and Chambers, 1998) uses the lagged relationship between SST and Australian rainfall to provide estimates of the probability of total rainfall in the following season being above median, for example. The scheme uses Indian Ocean and Pacific Ocean SST patterns as predictors (Figure 1, from Drosdowsky and Chambers, 1998) and displays greater skill than previous operational schemes based on the SOI.

The marine data used to develop the statistical relationships for the rainfall prediction scheme were provided by a number of sources. SST data from 1949 to 1991 were provided by the UK Met Office Global Ice and Sea Surface Temperature data set (GISST, version 1.1; Parker *et al.*, 1995). To extend the data to the present, the National Center for Environmental Prediction optimum interpolation analyses (Reynolds and Smith, 1994) and the operational analyses of the Bureau of Meteorology (Smith, 1995) were used.

The Smith (1995) products are analyses of surface and subsurface tropical Pacific Ocean temperatures (Figure 2). The observational marine data used for these analyses primarily comes from the volunteer observing XBT programme (TWXXPPC, 1993) and the TOGA TOA array (Hayes *et al.*, 1991).

* Corresponding author's address: Scott Power, National Climate Centre, Bureau of Meteorology, GPO Box 1289K, Melbourne, AUSTRALIA, 3001; e-mail: s.power@bom.gov.au

Figure 1—The first two principal components of SST variability that are used for seasonal prediction of Australian rainfall and temperature.



Further details on the ocean temperature analysis system can be found at: <http://www.bom.gov.au/bmrc/mrlr/nrs/climocan.htm>.

The differences between one observational estimate of SST and another are large enough to produce clear differences in the seasonal forecasts of rainfall that we are able to provide, and, therefore, efforts aimed at improving the accuracy of SST analyses are strongly supported.

A similar prediction scheme has also been developed to forecast seasonal temperature anomalies (Jones, 1998), and this is expected to be routinely disseminated to the general public soon. These temperature forecasts have been found to be more skilful than those for rainfall.

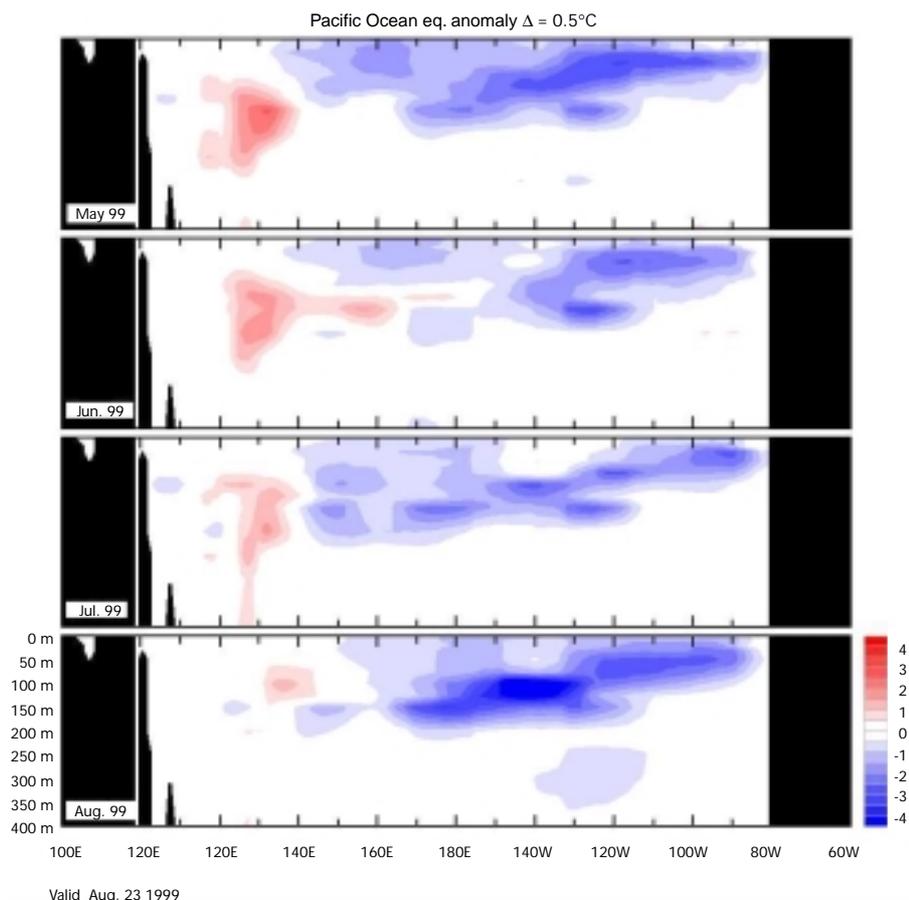
3. MARINE DATA AND SEASONAL FORECASTS BASED ON COUPLED OCEAN-ATMOSPHERE MODELS

3.1 INITIALIZATION

The use of subsurface analyses has been shown to significantly increase the hind-cast performance of the BMRC (Bureau of Meteorology Research Centre) intermediate coupled model (Kleeman, 1993; Kleeman *et al.*, 1995) used to provide guidance of ENSO development for the coming nine months.

The surface conditions of the ocean component during the assimilation phase were specified by the wind stress derived from FSU wind data (Goldenberg and O'Brien, 1981) and the SST data set of Reynolds and Smith (1994). Subsurface ocean data were derived from a number of real-time and near real-time sources as well as some ocean experiment data sets and archives.

Figure 2—Incorporating analyses of subsurface ocean temperatures has significantly increased the skill of the Bureau's coupled model predictions.



As well as the operational prediction schemes, the Bureau of Meteorology Research Centre is currently developing a forecast scheme based on a coupled ocean-atmosphere general circulation model (CGCM; Power *et al.*, 1998; Wang *et al.*, 1999). Marine data have played a critical role in the development and verification of this model and will also be important in the routine running of the model once it has been implemented as an operational scheme.

Recent sensitivity studies of the model by Wang *et al.* (1999) indicate that the inclusion of ocean data is crucial to achieve higher skill in hindcasts. The fact that reliable subsurface analyses prior to the early 1980s are not available necessarily restricts the hindcast period for coupled models and, therefore, impedes our ability to verify CGCM hindcasts. Given that estimates of forecast reliability are very important for potential users, this limitation in our climatic database represents a major shortcoming. Efforts aimed at expanding the range of climatic variables or lengthening the period over which relevant data are available are, therefore, eagerly awaited by seasonal forecasters.

- 3.2 VERIFICATION
- The verification of CGCMs requires a wide range of observational data sets. In fact, the WMO Commission for Basic Systems recently described a group of data sets which might be useful as part of an experimental long-range forecast verification project. These include :
- (i) Sea surface temperature
Reynolds OI, with option for additional use of GISST
 - (ii) Precipitation
Xie-Arkin; GPCP data; ECMWF Reanalysis and operational analysis data
 - (iii) Mean sea-level pressure
ECMWF Reanalysis and operational analysis data; own centre operational analysis data if available; UKMO GMSLP data set.

The ideal verification data set for seasonal prediction would, in general, have global spatial coverage, would extend back at least a few decades and would also be available in real time so that both hindcast and real-time forecasts could be

verified in a consistent fashion. In fact, the number of such data sets available is low and so many of the data sets suggested fall short of this ideal. This also represents a significant impediment to being able to provide assessments of seasonal forecasts that are as reliable and standardized as one might hope.

4. INTER-DECADAL VARIATIONS IN PREDICTIVE SKILL

It has been found that fluctuations in SST on inter-decadal time-scales seem to have a profound influence on the ability to predict ENSO-related seasonal rainfall anomalies in Australia (Power *et al.*, 1999). The success of an ENSO-based statistical scheme - and indeed the influence of ENSO on Australia in general - are shown to vary in association with a coherent, inter-decadal oscillation in SST over the Pacific Ocean.

When this Inter-decadal Pacific Oscillation (IPO) raises SSTs in the tropical Pacific Ocean, there appears to be no robust relationship between year-to-year Australian climate variations and ENSO. When the IPO lowers temperature in the same region, on the other hand, year-to-year ENSO variability is closely associated with year-to-year variability in rainfall, surface temperature, river flow and the domestic wheat crop yield. The contrast in ENSO's influence between the two phases of the IPO is quite remarkable and has serious implications for seasonal climate prediction in Australia, and possibly in other countries as well.

If subsequent research supports the need to ensure that variability of this kind is properly represented in future CGCMs, then additional data sets might be required, e.g. surface and sub-surface. This provides further evidence that marine data will continue to be vitally important for seasonal forecasting in the future.

5. SUMMARY

The success of the Bureau of Meteorology's forecast systems - both numerical and statistical, and operational and experimental - are critically dependent upon the ongoing availability of reliable marine data. The improvement in forecast skill of climate predictions depends, in part, on making greater use of existing and future marine data sets. From a seasonal forecasting perspective, therefore, efforts aimed at improving and extending existing data sets and expanding the range of appropriate climate data sets for use in statistical schemes, and for the verification and initialization of coupled atmosphere-ocean models, are highly valued and strongly supported.

REFERENCES

- Drosowsky, W. and L.E. Chambers, 1998: Near global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *BMRC Res. Rep. No. 65*, Bur. Met., Australia.
- Goldenberg, S.B. and J.J. O'Brien, 1981: Time and space variability of tropical Pacific wind stress. *Mon. Wea. Rev.*, 109, 1190-1207.
- Hayes, S.P., L.J. Mangum, J. Picaut, A. Sumi and K. Takeuchi, 1991: TOGA TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bull. Amer. Meteor. Soc.*, 72, 339-347.
- Jones, D.A. 1998: The prediction of Australian land surface temperatures using near global sea surface temperature patterns. *BMRC Res. Rep. No. 70*, Bur. Met., Australia.
- Kleeman R., 1993: On the dependence of hindcast skill on ocean thermodynamics in a coupled ocean-atmosphere model. *J. Clim.*, 6, 2012-2033.
- Kleeman, R., A.M. Moore and N.R. Smith, 1995: Assimilation of sub-surface thermal data into an intermediate tropical coupled ocean-atmosphere model. *Mon. Wea. Rev.*, 123, 3103-3113.
- Parker, D.E., C.K. Folland, A.C. Bevan, M.N. Ward, M. Jackson and K. Maskerll, 1995: Marine surface data for analysis of climatic fluctuations on interannual-to-century time scales. In: *Natural Climate Variability on Decade-to-Century Time Scales*. Climate Research Committee, National Academic Press, Washington D.C., U.S.A., 241-252.
- Power, S.B., F. Tseitkin, R.A. Colman and A. Sulaiman, 1998: A coupled general circulation model for seasonal prediction and climate change research. *BMRC Research Report No. 66*, Bur. Met., Australia.

- Power, S., T. Casey, C. Folland, A. Colman and V. Mehta, 1999: Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, 15, 319-324.
- Reynolds, R.W. and T.M. Smith, 1994: Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, 7, 929-948.
- Smith, N.R. 1995: The BMRC ocean thermal analysis system. *Aust. Met. Mag.*, 44, 93-110.
- TOGA/WOCEXBT/XCTD Programme Planning Committee 1993: Workshop on the use of sub-surface thermal data for climate studies (Brest, France). ITPO Rep. No.9/WOCE Rep. No. 110/93, WMO, 60 pp.
- Wang, G., R. Kleeman, N. Smith and F. Tseitkin, 1999: Seasonal predictions with a coupled global ocean-atmosphere system at BMRC. *BMRC Res. Rep.* Bur. Met., Australia (in preparation).

New JCOMM series replaces discontinued MMROA series

**JOINT WMO/IOC TECHNICAL COMMISSION
FOR OCEANOGRAPHY AND MARINE METEOROLOGY
TECHNICAL REPORT SERIES**

No.	Title	WMO/TD-No.	Issued
21	Workshop on Wind Wave and Storm Surge Analysis and Forecasting for Caribbean Countries (Dartmouth, Canada, 16-20 June 2003) - CD ROM only	WMO/TD-No. 1171	2003
20	JCOMM Ship Observations Team, Second Session - National Reports (London, United Kingdom, 28 July - 1 August 2003) - Website only	WMO/TD-No. 1170	2003
19	Automated Shipboard Aerological Programme (ASAP) - Annual Report for 2002	WMO/TD-No. 1169	2003
18	JCOMM Expert Team on Maritime Safety Services (ETMSS), First Session - National Reports (Lisbon, Portugal, 11-14 September 2002) - Website only	WMO/TD-No. 1135	2002
17	JCOMM Ship Observations Team, First Session - National Reports (Goa, India, 25 February - 2 March 2002) - Website only	WMO/TD-No. 1121	2002
16	Scientific and Technical Workshop of the JCOMM Ship Observations Team - Presentations at the first session of the Ship Observations Team (Goa, India, 26 February 2002) - CD ROM only	WMO/TD-No. 1118	2002
15	Automated Shipboard Aerological Programme (ASAP) - Annual Report for 2001	WMO/TD-No. 1112	2002
14	Operational Oceanography - Scientific Lectures at JCOMM-I (Akureyri, Iceland, June 2001)	WMO/TD-No. 1086	2001
13	Advances in the Applications of Marine Climatology - The Dynamic Part of the WMO Guide to the Applications of Marine Climatology	WMO/TD-No. 1081	2003
12	Automated Shipboard Aerological Programme (ASAP) – Annual Report for 2000	WMO/TD-No. 1069	2001
11	JCOMM Capacity Building Strategy	WMO/TD-No. 1063	2001
10	Proceedings of CLIMAR 99	WMO/TD-No. 1062	
9	Estimation of Extreme Wind Wave Heights	WMO/TD-No. 1041	2000
8	Oceanographic and Marine Meteorological Observations in the Polar Regions - A Report	WMO/TD-No. 1032	2000

*Out-of-print

No.	Title	WMO/TD-No.	Issued
	to the Joint WMO/IOC Technical Commission on Oceanography and Marine Meteorology		
7	Proceedings of a Workshop on Mapping and Archiving of Sea Ice Data – The Expanding Role of Radar, Ottawa, Canada, 2-4 May 2000	WMO/TD-No. 1027	2000
6	Automated Shipboard Aerological Programme (ASAP) – Annual Report for 1999	WMO/TD-No. 1011	2000
5	Voluntary Observing Ships (VOS Climate Subset Project (VOSCLIM) – Project Document)	WMO/TD-No. 1010	2000
	Revision 1	WMO/TD-No. 1042	2001
	Revision 2 - Website only	WMO/TD-No. 1122	2002
4	The Voluntary Observing Ships Scheme – A Framework Document	WMO/TD-No. 1009	2000
3*	JCOMM Ship-of-opportunity Programme Implementation Panel, Third Session, La Jolla, CA, USA, 28-31 March 2000 – SOOP Status Reports – SOOP Scientific and Technical Developments	WMO/TD-No. 1005	2000
2	Meeting of Experts on a JCOMM/GOOS Polar Region Strategy, Geneva, Switzerland, 6-8 December 1999 - Status Reports from Existing Polar Region Observing Systems	-	2000
1*	First Transition Planning Meeting - St Petersburg, Russian Federation, 19-23 July 1999 – Status Reports from JCOMM Component Bodies and Activities	-	1999

MARINE METEOROLOGY AND RELATED OCEANOGRAPHIC ACTIVITIES REPORT SERIES

No.	Title	WMO/TD-No.	Issued
44	Marpolser 98 - Metocean Services for Marine Pollution emergency Response Operations, Townsville, Australia, 13-17 July 1998 – Proceedings - Volume 2 - Review and Information Papers	WMO/TD-No. 960	1999
	Marpolser 98 - Metocean Services for Marine Pollution emergency Response Operations, Townsville, Australia, 13-17 July 1998 – Proceedings - Volume 1 - Research Papers	WMO/TD-No. 959	1999
43	Proceedings of the International Workshop on Digitization and Preparation of Historical Marine Data and Metadata, Toledo, Spain, 15-17 September 1997	WMO/TD-No. 957	1999

*Out-of-print

No.	Title	WMO/TD-No.	Issued
-	Automated Shipboard Aerological Programme (ASAP) - Annual Report for 1998	WMO/TD-No. 951	1999
42	Provision and Engineering Operational Application of Ocean Wave Data - in French only	WMO/TD-No. 938	1998
41	The Climate of the Baltic Sea Basin	WMO/TD-No. 933	1998
40	Automatisation de l'observation en mer Automation of Observations at Sea	WMO/TD-No. 928	1998
-	Automated Shipboard Aerological Programme (ASAP) - Annual Report for 1997	WMO/TD-No. 900	1998
39	Proceedings of the Commission for Marine Meteorology Technical Conference on Marine Pollution	WMO/TD-No. 890	1998
38	Evaluation of the Highest Wave in a Storm (A.V. Boukhanovsky, L.J. Lopatoukhin, V.E. Ryabinin)	WMO/TD-No. 858	1998
37	Tropical Coastal Winds (W.L. Chang)	WMO/TD-No. 840	1997
36	Handbook of Offshore Forecasting Services (Offshore Weather Panel)	WMO/TD-No. 850	1997
-	Automated Shipboard Aerological Programme (ASAP) - Annual Report for 1996	WMO/TD-No. 819	1997
35	Ice Navigation Conditions in the Southern Ocean (A.A. Romanov)	WMO/TD-No. 783	1996
-	Automated Shipboard Aerological Programme (ASAP) - Annual Report for 1995	WMO/TD-No. 767	1996
34	Polar Orbiting Satellites and Applications to Marine Meteorology and Oceanography – Report of the CMM-IGOSS-IODE Sub-group on Ocean Satellites and Remote Sensing	WMO/TD-No. 763	1996
33	Storm Surges (Vladimir E. Ryabinin, Oleg I. Zilberstein and W. Seifert)	WMO/TD-No. 779	1996
32*	Proceedings of the WMO/IOC Workshop on Operational Ocean Monitoring Using Surface Based Radars, Geneva, 6-9 March, 1995	WMO/TD-No. 694	1995
-	Proceedings of the WMO/IOC Technical Conference on Space-based Ocean Observations, September 1993, Bergen, Norway (US \$15.00)	WMO/TD-No. 649	1994
31*	Proceedings of the International Workshop on Marine Meteorology	WMO/TD-No. 621	1994
30	Proceedings of the International Seminar for Port Meteorological Officers	WMO/TD-No. 584	1993
29	Meteorological Requirements for Wave	WMO/TD-No. 583	1993

*Out-of-print

No.	Title	WMO/TD-No.	Issued
	Modelling (Luigi Cavaleri)		
28*	Proceedings of the Commission for Marine Meteorology Technical Conference on Ocean Remote Sensing	WMO/TD-No. 604	1993
27	A Survey on Multidisciplinary Ocean Modelling and Forecasting (Johannes Guddal)	WMO/TD-No. 516	1992
26	The Accuracy of Ship's Meteorological Observations - Results of the VSOP-NA		1991
25	Ships Observing Marine Climate - A Catalogue of the Voluntary Observing Ships Participating in the VSOP-NA	WMO/TD-No. 456	1991
24	Proceedings of the Commission for Marine Meteorology Technical Conference on Ocean Waves	WMO/TD-No. 350	1990
23	Summary Report on National Sea-ice Forecasting Techniques	WMO/TD-No. 329	1989
22	Wind Measurements Reduction to a Standard Level (R.J. Shearman and A.A. Zelenko)	WMO/TD-No. 311	1989
21	Coastal Winds (E.P. Veselov)	WMO/TD-No. 275	1988
20*	La Prévision du Brouillard en Mer (M. Trémant) Forecasting of Fog at Sea - in French only	WMO/TD-No. 211	1987
19	A Global Survey on the Need for and Application of Directional Wave Information (S. Barstow and J. Guddal)	WMO/TD-No. 209	1987
18	Baltic Multilingual List of Sea-ice Terms (Jan Malicki, Alex N. Turchin, Hans H. Valeur)	WMO/TD-No. 160	1987
17	Processing of Marine Data (G.D. Hamilton)	WMO/TD-No. 150	1986
16*	Field Workshop on Intercalibration of Conventional and Remote-sensed Sea Surface Temperature Data (A.E. Strong and E.P. McClain)		1985
15*	Forecast Techniques for Ice Accretion on Different Types of Marine Structures, Including Ships, Platforms and Coastal Facilities	WMO/TD-No. 70	1985
14	Scientific Lectures at CMM-IX	WMO/TD-No. 41	1985
13	User's Guide to the Data and Summaries of the Historical Sea Surface Temperature Data Project	WMO/TD-No. 36	1985
12	WMO Wave Programme: National Reports for 1984 on Wave Measuring Techniques, Numerical Wave Models and Intercomparisons	WMO/TD-No. 35	1985
*	Supplement No. 4 - Reports for 1991 to 1994 on Wave Measuring Techniques, Numerical	WMO/TD-No. 35	1994

*Out-of-print

No.	Title	WMO/TD-No.	Issued
	Wave Models and Intercomparisons		
	Supplement No. 3 - Reports for 1989 to 1990 on Wave Measuring Techniques, Numerical Wave Models and Intercomparisons	WMO/TD-No. 35	1990
	Supplement No. 2 - Reports for 1986 to 1988 on Wave Measuring Techniques, Numerical Wave Models and Intercomparisons	WMO/TD-No. 35	1989
	Supplement No. 1 - Reports for 1985 on Wave Measuring Techniques, Numerical Wave Models and Intercomparisons	WMO/TD-No. 35	1986
11	Drifting Buoys in Support of Marine Meteorological Services (Glenn D. Hamilton)	-	1983
10*	Guide to Data Collection and Services Using Service Argos (revised version)	-	1989
9	Intercalibrations of Directly-measured and Remotely Sensed Marine Observations (Alan E. Strong)	-	1983
8	Summary WMO Technical Conference on Automation of Marine Observations and Data Collection	-	1981
7*	Proceedings of the WMO Technical Conference on the Automation of Marine Observations and Data Collection	-	1981
6	Report on the Results of An Enquiry on Marine Meteorological Services Products	-	1981
5	The Automation of Observational Methods on Board Ship (M. Yasui)	-	1981
4	Intercalibration of Surface-based and Remotely Sensed Data (except Sea Surface Temperature Data) to Be Used in Marine Applications (E.P. McClain)	-	1981
3	Review of Reference Height for and Averaging Time of Surface Wind Measurements at Sea (F.W. Dobson)	-	1981
2	Investigation of Contemporary Methods of Measuring Sea Surface and Surface-layer Temperatures (F.S. Terziev)	-	1981
1	Precipitation Measurement at Sea (G. Olbrück)	-	1981

*Out-of-print