REPORT

(2)

1. Agency Use Only *(Leave blank).*	2. Report Date. 1992	3. Report Type and Dates Covered. Final - Journal Article

4. Title and Subtitle.
The Assimilation of Marine Surface Data into the Navy Operational Global Atmospheric Prediciton System

5. Funding Numbers.
Contract
Program Element No. 0603207N
Project No. X2008
Task No.
Accession No. DN650751
Work Unit No. 94311B

6. Author(s).
Patricia A. Phoebus and James S. Goerss

7. Performing Organization Name(s) and Address(es).
Naval Oceanographic and Atmospheric Research Laboratory
Atmospheric Directorate
Stennis Space Center, MS 39529-5004

8. Performing Organization Report Number.
JA 431:040:92

9. Sponsoring/Monitoring Agency Name(s) and Address(es).
Space and Naval Warfare Systems Command (PMW-141)
Washington, DC 20363-5100

10. Sponsoring/Monitoring Agency Report Number.
JA 431:040:92

11. Supplementary Notes.
Published in Marine Technology Society Journal

12a. Distribution/Availability Statement.
Approved for public release; distribution is unlimited.

12b. Distribution Code.

13. Abstract *(Maximum 200 words).*

As numerical models of the ocean and the atmosphere become more sophisticated, it becomes necessary to account for the interchange of energy at the air/sea interface. Meteorological models require accurate estimates of sea surface temperature and ice fields. Oceanographic models require accurate estimates of the surface winds and heat fluxes. Thus, it is beneficial for scientists in each discipline to have some level of understanding and appreciation for how these various parameters are obtained.

Withe that in mind, we will describe how a meteorological analysis/forecast system uses marine data to estimate the surface wind field, with particular emphasis on the analysis process and the resulting nowcast product. The Navy's data analysis system uses the multivariate optimum interpolation technique to produce updates to background fields provided by the forecast model. The advantage of this technique is that the relative weights given the background and the observations are determined by taking into account the estimated error properties of each. As new data sources become available, the appropriate statistics defining the errors of the observations must be determined. Since we have recently added several new data sources to the assimilation cycle, new values of the observation errors have been determined for all marine surface data. The improvements made by these new data sources, including special sensor microwave/imager wind speeds and synthetic tropical observations, will also be illustrated.

14. Subject Terms. Data assimilation, quality control			15. Number of Pages. 15
			16. Price Code.
17. Security Classification of Report. Unclassified	18. Security Classification of This Page. Unclassified	19. Security Classification of Abstract. Unclassified	20. Limitation of Abstract. SAR

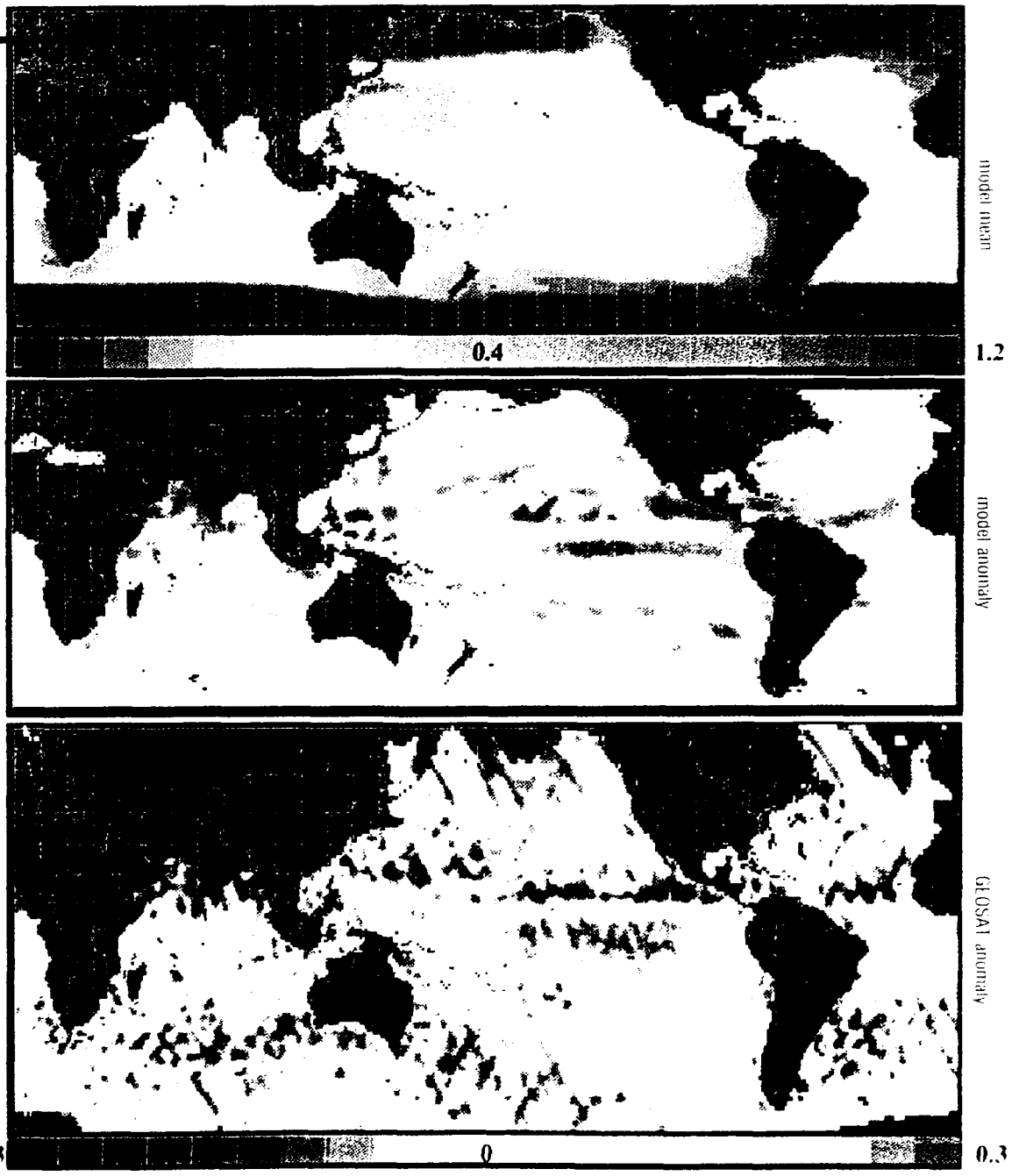
Marine Technology Society

JOURNAL

VOLUME 26
NUMBER 2
SUMMER 1992

THE INTERNATIONAL, INTERDISCIPLINARY SOCIETY DEVOTED TO OCEAN AND MARINE ENGINEERING, SCIENCE, AND POLICY

Oceanic and Atmospheric Nowcasting and Forecasting



The Assimilation of Marine Surface Data into the Navy Operational Global Atmospheric Prediction System

PAPER

Patricia A. Phoebus
Naval Research Laboratory—Monterey
Monterey, California

James S. Goerss
Naval Research Laboratory—Monterey
Monterey, California

ABSTRACT

As numerical models of the ocean and the atmosphere become more sophisticated, it becomes necessary to account for the interchange of energy at the air/sea interface. Meteorological models require accurate estimates of sea surface temperature and ice fields. Oceanographic models require accurate estimates of the surface winds and heat fluxes. Thus, it is beneficial for scientists in each discipline to have some level of understanding and appreciation for how these various parameters are obtained.

With that in mind, we will describe how a meteorological analysis/forecast system uses marine data to estimate the surface wind field, with particular emphasis on the analysis process and the resulting nowcast product. The Navy's data analysis system uses the multivariate optimum interpolation technique to produce updates to background fields provided by the forecast model. The advantage of this technique is that the relative weights given the background and the observations are determined by taking into account the estimated error properties of each. As new data sources become available, the appropriate statistics defining the errors of the observations must be determined. Since we have recently added several new data sources to the assimilation cycle, new values of the observation errors have been determined for all marine surface data. The improvements made by these new data sources, including special sensor microwave/imager wind speeds and synthetic tropical observations, will also be illustrated.

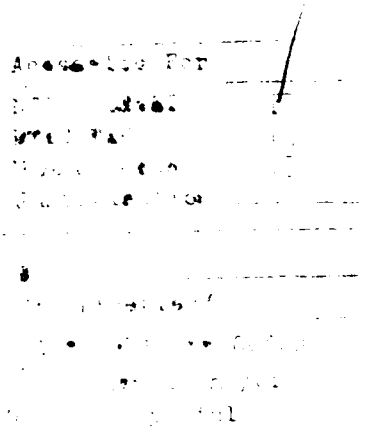
INTRODUCTION

Given the current emphasis on global climate change, the scientific community has begun to look more closely at the total global system and the relationships among its various components. As a result, research interest in coupled global ocean/atmosphere models has increased dramatically. As we study and define the interchange of energy at the air/ocean interface, scientists from each discipline must also develop a better understanding of how the parameters of the complementary system are derived. Although the Navy is not directly involved in the climate effort, they have long been at the forefront of air/sea research to support the requirements of their operations in the marine boundary layer. These efforts include but are not limited to numerical modeling of the atmosphere and ocean.

Accurate nowcasts and forecasts of atmospheric conditions are essential to the Navy for several reasons. Most obviously, the atmosphere affects ship and aircraft routing, aircraft carrier operations, and weapons systems performance. Not to be overlooked, however, is the impact of the atmosphere on the oceanographic products produced by the Navy. Currently, atmospheric winds and heat fluxes provide forcing for ocean thermodynamic, hydrodynamic, wave, and ice models (Clancy, 1987), while sea-surface temperature and ice analyses improve the boundary layer parameterizations within the meteorological forecast model. Through this interactive coupling, even the traditional oceanographic domains of search and rescue, acoustics, and arctic submarine operations are directly linked to the atmosphere.

Various meteorological and oceanographic products, such as those mentioned above, are produced and disseminated by the Navy's Fleet Numerical Oceanography Center (FNOC), while key research and development tasks are performed by divisions of the Naval Research Laboratory (NRL) in Monterey, California, co-located with FNOC, and at Stennis Space Center in Mississippi. While achievements in air/ocean model coupling are driven by Fleet requirements, these efforts are supported by the most complete global database in the world. In addition to conventional radiosonde and surface data, FNOC's meteorological database includes satellite-derived temperature soundings and satellite-derived wind information, along with synthetic observations in the vicinity of extratropical and tropical storms. Although much of this data is available over the Global Telecommunications System (GTS), FNOC was the first operational center to have real-time access to the marine surface wind data from the special sensor microwave/imager (SSM/I) flown onboard the Defense Meteorological Satellite Program spacecraft. While the SSM/I data have recently been made available to the National Meteorological Center, their distribution and operational use are still limited at this time. Furthermore, classified observations and locally generated synthetic observations are not available to outside users.

To take full advantage of this database for improved meteorological analysis (nowcast) and forecast products, significant upgrades were made to the Navy Operational Global Atmospheric Prediction System (NOGAPS) in January



A-1

of 1988. Like most global data assimilation systems, NOGAPS consists of four components that are inextricably connected—data quality control, data analysis, model initialization, and model forecast. The newer generation of NOGAPS included major upgrades to all of these components. The grid-point forecast model was replaced with a global spectral model (Hogan and Rosmond, 1991), the successive corrections analysis scheme was replaced by a three-dimensional multivariate optimum interpolation (MVOI) analysis (Barker, 1992; Goerss and Phoebus, 1992a), the balance equation was replaced by a nonlinear normal-mode initialization (Hogan et al., 1991), and significant improvements were made to the data quality control (Baker, 1992). Since 1988, modifications such as higher resolution, better physical parameterizations, and improved data handling have continued to increase the skill of the analysis and forecast products. As a result, NOGAPS is now competitive with the global data assimilation systems run at other operational centers.

This paper will concentrate on the near-surface analyses over the oceans—an area of obvious interest to the Navy, and an area where FNOC's unique database can be fully exploited. The following sections describe this database and the MVOI analysis technique in more detail. The final section will also discuss the impact of new data sources on the analyzed products. We hope to provide the oceanographer with an overview of the process that takes meteorological data from a number of sources that are randomly distributed in space and time and from them produces spatially coherent gridded wind fields in the marine boundary layer.

DATA ANALYSIS

The global analysis product is produced for sixteen pressure levels (the standard levels from 1000 millibars [mb] to 10 mb) on the Gaussian grid of the NOGAPS spectral forecast model. The Gaussian grid points are separated by approximately 1.5°, but are thinned near the poles for the data analysis. The analyzed meteorological variables are geopotential height h and the zonal and meridional wind components, u and v, respectively. Observations must be in the form of one of these three variables or in the form of geopotential thickness (the difference in height between two adjacent pressure surfaces). For example, observations of surface pressure are converted to 1000 mb heights, while satellite temperature soundings are converted to geopotential thickness. Wind direction and speed measurements are decomposed into u and v components.

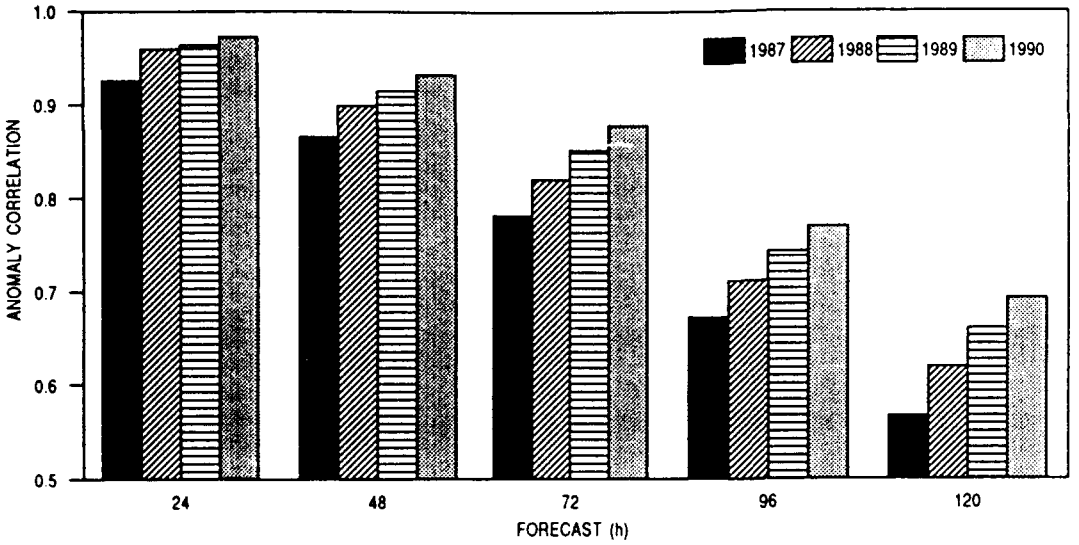
In a data assimilation system like NOGAPS, the 6-hr model forecast valid at the analysis time is used as the background, or first guess, for the analysis. The analysis, in turn, combines information from the background with the quality-controlled observations to provide updated fields that are used to initialize the forecast model. Thus, the analysis is an *incremental* analysis. Observation increments are computed for all data by interpolating the appropriate background field to the observation locations and subtracting. The *increments* are then analyzed and added to the background field to produce the latest three-dimensional representation of the atmospheric height and wind fields.

The MVOI technique utilizes data from all available platforms to produce analyzed fields that possess minimum error in a statistical sense, while taking into account the error properties of the model forecast used for the background fields and the error properties of the observational data used to update the background. The MVOI also constrains the results so that the analyzed height and wind increments are in geostropic balance outside of the tropics, and the vertical consistency of the model background fields are maintained. Thus, unlike earlier objective analysis schemes that essentially averaged the nearby data, the MVOI makes some very intelligent decisions about how to treat the various, often conflicting, sources of information it has available.

Four major factors influence the quality of the analyzed product. One is the quality of the forecast used as the analysis background; another is the quality of the observations themselves, which requires the application of sophisticated quality control algorithms. The third factor is the selection, during the analysis, of the appropriate mix of mass and wind observations from the various sensors. The last factor is the proper specification of the statistical parameters required by the optimum interpolation methodology.

Since the skill of the forecast model is dependent upon the quality of the other system components, it is impossible to completely isolate and define model skill. However, the skill of the data assimilation system as a whole is reflected in Figure 1. The 500 mb anomaly correlation is commonly used as a measure of forecast skill by the operational centers. A correlation of 0.6 is considered to be the minimum value for defining a useful forecast. As can be seen from Figure 1, the skill scores for NOGAPS show steady improvement over the four year period from 1987 to 1990. It is also interesting to compare the relative skill of, for example, a 48-hr forecast in 1987 with a 72-hr forecast in 1990. Clearly, there has been a one-day improvement, with a three-day forecast now having the same skill that a two-day forecast had four years ago.

FIGURE 1. NOGAPS 500 mb height anomaly correlation in the Northern Hemisphere extratropics for the years 1987 to 1990.



To ensure that erroneous observations are identified and rejected or corrected, data quality control is performed in several steps. First, except for locally generated synthetic data, all observations are checked before they are passed to the analysis (Baker, 1991). These tests include checks against climatological limits and checks for internal consistency between various parameters within each report. For example, if the wind direction is reported as variable, the wind speed must be under 7 m sec^{-1} or the wind observation is rejected; if both temperature and dew-point temperature are reported, the dew-point temperature cannot be greater than the air temperature, and over water, the difference between them is not allowed to exceed 30°C. In the case of sounding data, checks for vertical consistency are also applied, including lapse-rate checks, hydrostatic checks, and checks for excessive vertical wind shears. Erroneous values are flagged for rejection, while questionable values are flagged as suspect.

The suspicious observations are checked further within the analysis, after the observations have been converted to observed increments by interpolating and subtracting the appropriate background variable at each observation location. The analysis gross-error check compares the increments to predefined tolerances that are functions of the location and type of observation, the observation error assigned to the data, and the prediction error associated with the background at that location (Goerss and Phoebus, 1992b). Increments exceeding these limits may be rejected outright or flagged as suspect. All questionable increments are further checked against surrounding data during the optimum interpolation analysis. If an increment cannot be corroborated by surrounding data, it is excluded from the final analyzed product.

The appropriate mix of different types of observations from the various sensors is guaranteed in several ways. By selecting data and performing the analysis within large volumes that contain many grid points (Lorenc, 1981), we avoid the problems associated with trying to select the few *best* observations for each grid point; we essentially allow all the data over a large volume to influence the outcome at a given point. Furthermore, the volume size is varied so that volumes in data-sparse regions (over the oceans) are larger than those in data-rich areas (over the Northern Hemisphere continents). The smallest volumes typically cover an area approximately 600 km on a side, with the largest around 2,400 km on a side, and all volumes extend throughout the depth of the atmosphere. There are limits set on the number of observations of one type that will be included in a particular volume. Thus, in data-rich areas, we ensure that no one data type will be selected at the expense of another.

All of the available data within a specified volume—both upper-air and surface data—are differenced from the background field and used to compute the analyzed increments at each grid point in the volume. Because the analysis is multivariate and three-dimensional, the entire set of height, thickness, and wind increments influence both the height and wind fields. The extent of this influence is controlled by the various statistical parameters that are part of the optimum interpolation formulation.

The determination of the statistical parameters required by the analysis is accomplished using a special database created for the purpose of monitoring the performance of the

MVOI. This database contains a history file of observation minus background increments and their associated quality flags, so only increments that passed all analysis quality checks are used in the calculations. Since the MVOI is an incremental analysis, the required statistics must pertain to the *increments*. The statistics that must be specified for the MVOI are the errors in the background fields (prediction error), the prediction error correlation functions for all possible combinations of data-type pairs (e.g., *hh, hu and hv*), and the observational errors for each data source. The data increments in the historical database reflect both the prediction error and the observation error, but this total error can be partitioned into the two separate error components following Hollingsworth and Lönnberg (1986). The prediction error is normally determined from the North American radiosonde increment database. Once the prediction error is known, the observation errors can be estimated from similar databases for each data source. The details of how these quantities are derived are discussed in Goerss and Phoebus (1992b).

For the purposes of this paper, we note that the horizontal and vertical prediction error correlation functions are determined from actual data increments in such a way that the analyzed height and wind increments will be geostrophically consistent outside of the tropics. The prediction errors specified for height and wind data are also geostrophically consistent, thus the height errors vary somewhat with latitude. Both height and wind errors vary with volume size to account for the fact that the background field is somewhat less accurate in data sparse areas. Table 1 shows sample values of the prediction errors at 1000 mb, as derived from actual data increments. Volume sizes range from 600 to 2,400 km along a horizontal side, depending upon volume location.

The observational error accounts for both the instrument error of the sensor and the sub-grid scale error, that is, how representative the observation is of the scales being analyzed. When a new data source becomes available, its observation error must be determined so that it receives the proper weight relative to the other data types. The weight a particular observation receives at a particular grid point is a function of its observation error, the prediction error at the observation's location, and the observation's correlation with the grid point and with the other data in the volume. For example, if two observations are co-located, the observation with the larger observation error will receive the smaller weight and will have less impact on the outcome of the analyzed product.

In the past, the error characteristics of the conventional marine surface observations were treated as equal, mainly due to the fact that the marine surface data were stored and processed without regard to data source. However, given the importance of marine surface forecasts to Navy operations and to the success of coupled air/ocean forecast models, recent research efforts have focused on better specification of the error properties of the marine observations and of the analysis background over the oceans.

TABLE 1. Prediction error standard deviations.

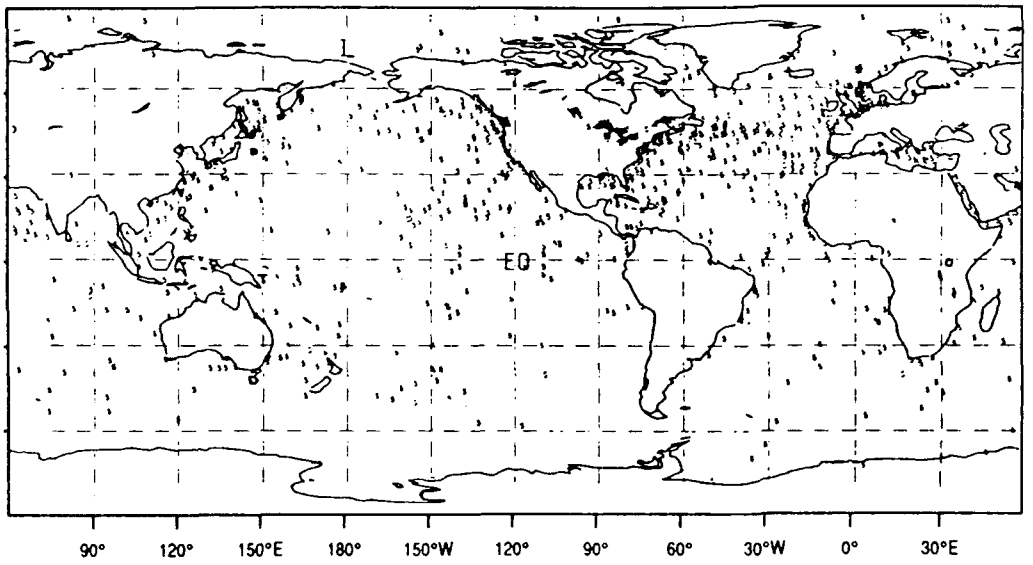
Pressure mb	Latitude	Vol. Size	Wind m s^{-1}	Height m
1000	60°	Small	2.0	10.5
1000		Average	2.3	12.0
1000		Large	2.6	13.6
1000	45°	Small	2.0	8.5
1000		Average	2.3	9.8
1000		Large	2.6	11.1
1000	30°	Small	2.0	6.0
1000		Average	2.3	6.9
1000		Large	2.6	7.8

Originally, we assumed that the height and wind observation errors were the same for all surface data. However, with the recent introduction of several new types of observations, we decided to revisit this problem and compute the statistics separately for each different marine data source. Monitoring the surface (observation minus background) increments by data type and by region indicated that measurements made by some platforms were better than others. Thus, we now identify the source of each observation and treat their statistics independently. The changes in the specified observation errors are given in Table 2, which reflects, for example, that the errors in the moored buoys reports are smaller than those from the ship reports, as are the island wind reports. Examination of Table 1 and Table 2 together illustrates how the errors in the model forecast compare to the observation errors of the data. The accuracy of global models today is such that the observation errors in the data may actually exceed the prediction error in the 6-hr forecast used for the analysis background.

TABLE 2. Observation error standard deviations.

	Height m		*u* or *v* Wind m s^{-1}	
	Old	New	Old	New
Island	6.0	7.0	—	1.9
Ship	6.0	7.0	2.2	2.4
USN ship	6.0	7.0	2.2	2.2
Moored Buoy	6.0	4.0	2.2	1.8
Drifting Buoy	6.0	7.0	2.2	2.2
C-MAN	6.0	7.0	2.2	2.2
PAOB	24.0	20.0	N.A.	N.A.
Sfc Bogus	6.0	7.0	2.2	2.4
SSM/I	N.A.	N.A.	2.2	2.4

IRE 2. Typical distribution of surface ship, buoy, and C-MAN stations within a 6-hr data analysis window.



ARINE DATABASE

While the radiosonde is the backbone of the observational database over continental as, the number of radiosondes launched over oceans is limited to a few from the more ulated islands and an occasional ship of ortunity. These isolated upper air stations supplemented with temperature soundings n the polar-orbiting satellites, cloud-tracked d observations from the geostationary satel-s, and aircraft wind reports. However, at the ocean interface, additional observations are vided by military, research, and commercial s; moored and drifting buoys; coastal marine omated network (C-MAN) stations; islands; high-resolution surface wind speed esti-es from the SSM/I. Because the MVOI analy-s a three-dimensional analysis, the upper-air ervations impact the resulting fields at the est analysis level, but this influence is limited rface data are available in the same area. s, we will only describe the surface data in paper. Additional information about the re database is given by Goerss and Phoebus 2a).

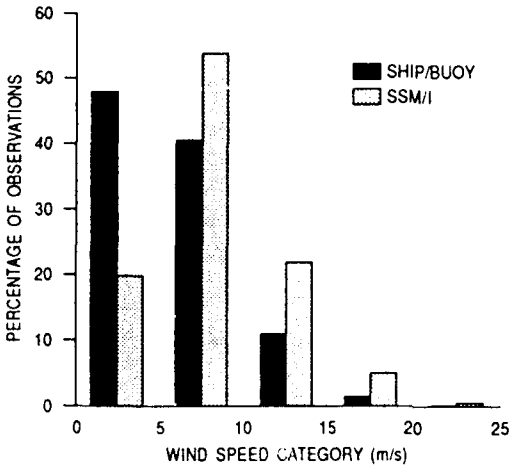
A marine surface observation may be sured or considered valid at heights that from a few to several tens of meters above actual air/ocean interface, depending on the and location of the instrument making the surement. Since the vertical resolution of global model does not support these distinc-s, all marine surface observations are con-ed to observations of 1000 mb height and 0 mb u and v wind components. We do, ever, make use of boundary layer informa-when we do this assignment. The model produces a 10 m wind field that is stability dependent and consistent with the model's boundary layer physics, and we have found that these winds agree more closely with the marine data than the 1000 mb winds do. Thus, to convert the surface wind measurements to 1000 mb increments, we actually compute boundary layer wind increments by interpolating and subtracting the model's 10 m winds from the observed surface wind components. These 10 m u and v wind increments are then assigned to the 1000 mb level. By assuming that the wind increments do not vary vertically over the small distance between 10 m and 1000 mb, we are preserving the vertical structure defined by the model's boundary layer.

Conventional Data

The conventional marine surface data include reports from ships, moored and drifting buoys, and C-MAN stations. A typical 1200 universal time clock (UTC) analysis product might include 400 or 500 ships, slightly less than 100 moored and drifting buoys (each), and 50 or 60 C-MAN reports. Figure 2 illustrates the global distribution of these observations. Approximately 90 percent of the reports are in the Northern Hemisphere. The variables normally reported by these marine platforms include sea-level pressure, temperature, moisture, wind speed, and wind direction. Drifting buoys are the exception since, for the most part, the only atmospheric variable they measure is sea-level pressure.

Many of the marine platforms report frequently. As the data are processed, duplicate or redundant reports are identified by

FIGURE 3. Distribution of wind observations by wind speed category. Percentage of SSM/I observations in each category compared to percentage of ship and buoy observations in the same categories.



comparing latitudes, longitudes, and block/station names. Reports at the same location or with the same identifier are removed. Only one observation per platform is used in the analysis, even if that platform reports every hour. The observation saved is the one closest to the analysis time.

Island Data

Like other surface observations over land, islands typically report wind speed and direction and either sea-level pressure, station pressure, or geopotential height at a standard pressure surface. If pressure is reported, it is converted to a height (Goerss and Phoebus, 1992b). Zonal and meridional wind components are computed from the reported wind speed and direction. Until recently, surface wind observations over land were *not* used in the analysis. Such observations are often not representative of the wind in the free atmosphere because of the effects of terrain or shallow radiation inversions. However, many of the island instrument stations have been relocated to reduce these problems. Personnel from the Joint Typhoon Warning Center (JTWC) in Guam, who routinely monitor and analyze the Pacific island wind reports, indicated that JTWC generally accepts the island wind directions as accurate, but makes corrections to the reported speed. So, these observations were re-evaluated for use in NOGAPS. The observation minus background differences indicated that the Pacific island wind observations were as accurate as the ship wind reports. Thus, these particular island wind observations are now included in NOGAPS, but *only if* the reported wind speed is greater than 4 m/s. This caveat rejects the light and variable or calm wind reports that are so common during the night but includes reports that are more representative of a well-mixed boundary layer.

The Pacific islands cover an area that extends from 30°S to 30°N and includes slightly less than 100 island stations. Islands outside this region are still excluded, since they are more isolated and it is extremely difficult to determine the quality of their observations. Surface wind reports from continental stations are also ignored.

SSM/I Wind Data

FNOC uses the algorithm of Goodberlet et al. (1990) to estimate the ocean surface (19.5 m) wind speeds from the SSM/I measured brightness temperatures. These speeds range up to 30 m s^{-1} and meet the specified accuracy of ±2 m s^{-1} in rain-free conditions. Since heavy concentrations of water vapor and rain effectively mask microwave emission from the ocean surface at the SSM/I frequencies, wind speed retrievals under these conditions are more likely in error. FNOC uses four rain-flags to indicate accuracy, ranging from 0 (no rain, error < 2 m s^{-1}) to 3 (heavy rain, error > 10 m s^{-1}). Currently, the NOGAPS data pre-processor screens the SSM/I data so that only wind observations with a rain-flag of 0 or 1 are used in the MVOI.

The SSM/I data are valued not only for their accuracy but also for their distribution. Figure 3 shows the distribution of the SSM/I wind speeds compared to those from the ships and buoys. The majority of the SSM/I wind speed data (> 50 percent) fall in the 5 to 10 m s^{-1} category, while the 0 to 5 m s^{-1} category has the largest percentage of ship and buoy observations. What is especially noticeable is that, overall, the higher wind speed categories contain a larger percentage of SSM/I observations than ship observations. Since ships are deliberately routed away from high wind speed areas, this difference most likely reflects the indiscriminate global coverage of the SSM/I data.

Reasonable global coverage is provided by the SSM/I within a twenty-four hour period, thus providing an important supplement to the other relatively sparse observations over the ocean. On the other hand, a typical six-hour coverage (the data assimilation cycle) produces in excess of 200,000 observations at approximately 25 km resolution, a resolution much higher than the 1.5° resolution of NOGAPS. Thus, it is reasonable to filter the information provided by the SSM/I observations. Since the SSM/I observations are extremely consistent, only every sixth SSM/I observation is passed to the analysis. Furthermore, any wind speed reported outside the range of the SSM/I instrument is ignored.

Next, the wind speed data are subjected to quality control procedures that compare

ABLE 3. SSM/I quality control.

Preliminary Analysis	Rejection Criteria	Percentage Rejected
< 4 m s^{-1}	SSM/I > 7 m s^{-1}	~ 5.0%
4-10 m s^{-1}	| Diff | > 7.5 m s^{-1}	~ 1.5%
> 10 m s^{-1}	| Diff | > 10 m s^{-1}	~ 0.1%

the observed wind speed to a preliminary analysis of wind speed valid at 10 m. The 10 m wind analysis product is obtained from a preliminary analysis that is executed from the same software and the same background fields as the NOGAPS analysis, but it is executed at only two hours past the synoptic time, so there are fewer observations available than there are during the normal assimilation cycle. The preliminary analysis uses almost all of the available data, except for aircraft wind reports, satellite temperature soundings, and the SSM/I wind speed data, and it is performed only at the lowest four analysis levels (1000 to 700 mb). By the same reasoning that is used to assign surface data increments to 1000 mb, we create the 10 m wind analysis product by adding the analyzed 1000 mb increments to the 10 m background fields.

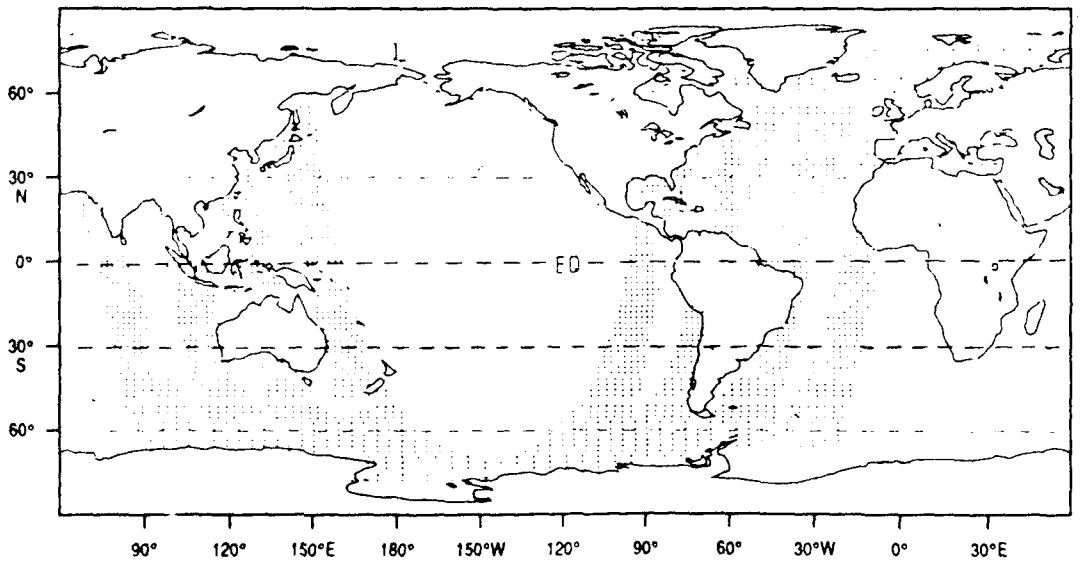
Whether or not an observation is rejected during the comparison to the preliminary 10 m wind product depends upon both the observed wind speed and the analyzed wind speed valid at the observation location. The rejection criteria are given in Table 3, where | Diff | represents the absolute value of the observed minus analyzed wind speed, and the percentage rejected is in reference to the total number of SSM/I observations processed. The strictest criteria are applied and the most data are rejected when the analyzed wind speed is very light coupled with an SSM/I wind speed that is substantially stronger. Since the preliminary analysis product will also be used to assign wind directions to the SSM/I observations, this decision reflects the fact that it would be inappropriate to assign a wind direction from a light and variable condition to a stronger wind speed. Doing so could create unrealistic low-level wind shears and convergence/divergence zones. The other categories of comparison exhibit more tolerance for differences between the preliminary analysis fields and the SSM/I data and, as a result, reject fewer observations.

Computation of Superobs

Since the small scales sampled by the SSM/I are of little importance to a global data assimilation system, it is reasonable to further filter the information provided by the SSM/I observations. To reduce the number of data points without losing information on scales relevant to NOGAPS, the SSM/I data are combined into *superobs*, which are essentially local averages of groups of observations. A typical distribution for the superobs available for a single analysis product is shown in Figure 4, which was plotted using a six-hour data window centered at 1200 UTC.

Before the superobs are computed, the 10 m wind speeds from the preliminary analysis fields are interpolated to the location of each individual observation, and the interpolated values are subtracted from the observed wind speeds to form SSM/I wind speed increments. Next, the globe is divided into 200 km boxes, and the SSM/I wind speed *increments* within

FIGURE 4. Typical distribution of SSM/I superobs within a 6-hr data analysis window.



each box are averaged. The increments rather than the original values are used because, in general, the increment field is more horizontally homogeneous than the raw data. Once the superob increments are computed, they are assumed to be valid at the centers of the boxes. Finally, the superob wind speeds are determined by adding each superob increment to the 10 m analyzed wind speed interpolated to the superob location.

Statistical comparisons of observed minus analyzed increments for the original SSM/I wind speed data and the SSM/I wind speed superobs (Table 4) illustrate that the differences in the root-mean-square (rms) errors are small, which indicates that the superob process is not introducing significant errors in the data. If anything, the averaging process tends to reduce the random noise that is present in individual observations. Similar statistics were produced for superobs classified by the number of individual SSM/I observations composing the superobs. The errors were larger for those cases where only a few increments were averaged—most likely near coastlines or along the edges of swaths, where the SSM/I data may not be as reliable. Furthermore, the error introduced by assigning the superob to a location at the center of the box will be greater when there are fewer data points in the box. Thus, superobs are computed only for boxes with at least five observations.

Assignment of Wind Direction

The analysis variables are geopotential height and the u and v wind components—variables that have well-known dynamical relationships. Thus, the MVOI analysis does not utilize scalar wind speed information directly. So, once the SSM/I superob increments have been computed, each superob wind speed must be assigned a wind direction. There have been several techniques proposed for independently determining appropriate directions for the SSM/I wind speeds. Atlas et al. (1991) combine conventional wind data and SSM/I winds with background analyzed surface wind fields from the European Centre for Medium Range Weather Forecasts, using variational techniques to constrain the degree to which the SSM/I data can modify the background. Yu (1987) bases his technique on Ekman boundary layer dynamics, using the National Meteorological Center's surface pressure analysis product and the SSM/I wind speed to compute the surface drag coefficient, which then allows the computation of the vector wind at some height z^* within the boundary layer (typically 10 m).

In our case, we use the power of a global data assimilation system to produce the 10 m preliminary fields using the most recent data, independent of the SSM/I. This analyzed wind field is interpolated to the location of each superob wind speed, and the resulting wind directions are assigned to the superob locations. From this point on, the SSM/I superob components are treated like any other surface wind observations.

TABLE 4. SSM/I wind speed statistics. May 24–28, 1991.

	Original SSM/I Wind Speeds Ob — Prelim. Anal. (m s^{-1})		
SSM/I (m s^{-1})	Mean	St Dev	rms
0– 5	–1.2	1.8	2.1
5–10	0.4	1.7	1.7
10–15	1.9	2.1	2.8
15–20	3.6	2.1	4.2
20–25	6.4	2.2	6.9
	SSM/I Superob Wind Speeds *Ob — Prelim. Anal. (m s^{-1})*		
Superob (m s^{-1})	Mean	St Dev	rms
0– 5	–0.4	1.5	1.5
5–10	0.6	1.5	1.6
10–15	1.9	2.0	2.7
15–20	3.5	1.9	4.0
20–25	6.0	2.3	6.3

Synthetic Data

Finally, the global database is supplemented by the creation of synthetic observations for specific applications. The most commonly used synthetic observations, which are available to all the operational forecast centers, are the subjectively derived estimates of sea-level pressure (PAOBS) that are produced by the Australian Bureau of Meteorology from satellite imagery. These observations are evenly distributed over both land and sea, but are confined to the Southern Hemisphere. NOGAPS uses this data only south of 20°S. Approximately 150 PAOBS are available at FNOC for the 0000 UTC and 1200 UTC analysis products.

Locally, FNOC personnel generate synthetic, or *bogus*, surface observations in the vicinity of oceanic extra-tropical cyclones. Typically, such observations are created in relatively data-sparse areas like the North Pacific, in situations where satellite imagery indicates that the cyclone is deeper than its depiction by NOGAPS. This technique is particularly useful for capturing explosively deepening lows in the Gulf of Alaska, since in the absence of other data, the synthetic observations can provide a surface baseline for the layer thicknesses derived from the satellite soundings.

Only the central location and sea-level pressure of the low is subjectively specified. The difference between this synthetic observation and the analyzed sea-level pressure at the same location is then used to automatically generate eight additional synthetic observations, centrally

located 325 km from the storm's center on radii separated by 45° (Goerss, 1989). These outer observations include both sea-level pressures and winds. The pressures are determined by assuming that the pressure difference changes uniformly from the central difference to zero as distance from the center increases from 0 to 650 km along a given radius. Thus, half the central pressure difference is added to the analyzed sea-level pressure at each of the eight observation locations. The synthetic wind observations are determined by adding the geostrophic components implied by the pressure difference distribution to the analyzed wind components at the observation locations.

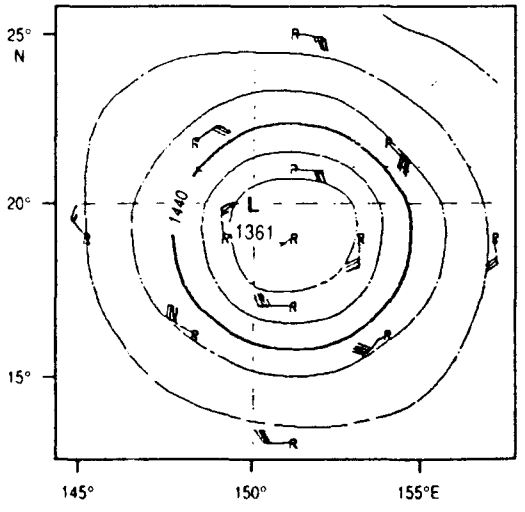
Once the synthetic surface observations have been generated, they are temporarily entered into the NOGAPS operational database. A preliminary analysis is then performed with the MVOI for a limited area surrounding the cyclone of interest. If the FNOC personnel are satisfied with the resulting changes in the low, the synthetic observations are entered into the database permanently, and they are accessed for the next analysis/forecast update cycle.

After changes to the physical parameterizations and resolution of the forecast model in 1989, NOGAPS began actively producing and developing tropical circulations. To ensure that the circulations were maintained in the proper location and at the correct strength throughout the data assimilation cycle, NRL developed a technique for assimilating synthetic *wind soundings* in the vicinity of these tropical storms (Goerss et al., 1991). In this case, the generation of the synthetic observations is completely automated, using the warning messages issued by JTWC and by the National Hurricane Center in Miami for tropical cyclones whose maximum wind speeds are greater than 13 m/s (depression stage). The synthetic soundings are centered on the storm's location and reflect the maximum wind speed and radius of maximum wind information contained in the warning message.

Each synthetic tropical sounding contains wind information from 1000 mb to 400 mb, along with a 1000 mb height observation. If the circulation has reached tropical storm strength or greater, soundings are generated at thirteen points: one at the center; four located 220 km north, south, east, and west of the center; four located 440 km northeast, southeast, southwest, and northwest of the center; and four located 660 km north, south, east, and west of the center (Figure 5). For tropical depressions, soundings are generated only at the five most interior points. All positions are updated and new observations are generated every 6 hrs.

Each sounding is composed of two parts. The first is the contribution of the large-scale environmental flow, determined by spectrally truncating the NOGAPS background field so that only the first twenty wave numbers are retained. The other is the vortex component, derived from a symmetric Rankine vortex in gradient balance (Andersson and Hollingsworth, 1988). The structure of the Rankine vortex is controlled by the maximum wind speed, the radius of maximum wind, and an exponential factor the governs the flow beyond this radius. The contribution of the symmetric vortex component is reduced in the vertical to simulate the warm-core structure of the tropical cyclone. Furthermore, the Rankine vortex wind speeds do not drop off quickly enough with distance in the horizontal. Thus, the vortex wind speeds in the outer two circles are reduced by 40 percent at 660 km and by 20 percent at 440 km to compensate. The resulting set of synthetic observations produces a tropical cyclone that is consistent with how the global model depicts such features rather than how tropical storms actually appear in nature.

FIGURE 5. Location of synthetic tropical soundings relative to storm center, overlayed on an 850 mb height analysis of typhoon Pat.



DATA IMPACT

Impact of SSM/I Data

The performance of the data assimilation system is monitored by continually producing statistics of observation minus background and observation minus analysis differences. Just after the introduction of the SSM/I data into NOGAPS, global statistics were computed from the observation minus (1000 mb) background differences (Table 5). First, we notice that a significant bias is present at the higher wind speeds, which is also reflected in the rms error.

TABLE 5. SSM/I wind speed statistics. September 13–14, 1990.

	Ob — Background ($m\ s^{-1}$)		
SSM/I ($m\ s^{-1}$)	Mean	St Dev	rms
0– 5	–1.7	2.2	2.8
5–10	–0.3	2.3	2.3
10–15	0.8	3.3	3.4
15–20	2.7	3.1	4.1
20–25	6.6	2.4	7.0

TABLE 6. SSM/I wind speed statistics. September 23–24, 1990.

	Ob — Background ($m\ s^{-1}$)		
SSM/I ($m\ s^{-1}$)	Mean	St Dev	rms
0– 5	–1.5	2.0	2.5
5–10	–0.2	2.1	2.1
10–15	0.8	2.7	2.8
15–20	1.8	2.6	3.2
20–25	4.3	2.6	5.0

TABLE 7. Ship/buoy wind statistics.

	Ob — Background Differences	
rms	Before SSM/I	After SSM/I
N Hem	5.0 $m\ s^{-1}$	4.9 $m\ s^{-1}$
S Hem	5.3 $m\ s^{-1}$	5.1 $m\ s^{-1}$

TABLE 8. Marine wind speed statistics. May 24–28, 1991.

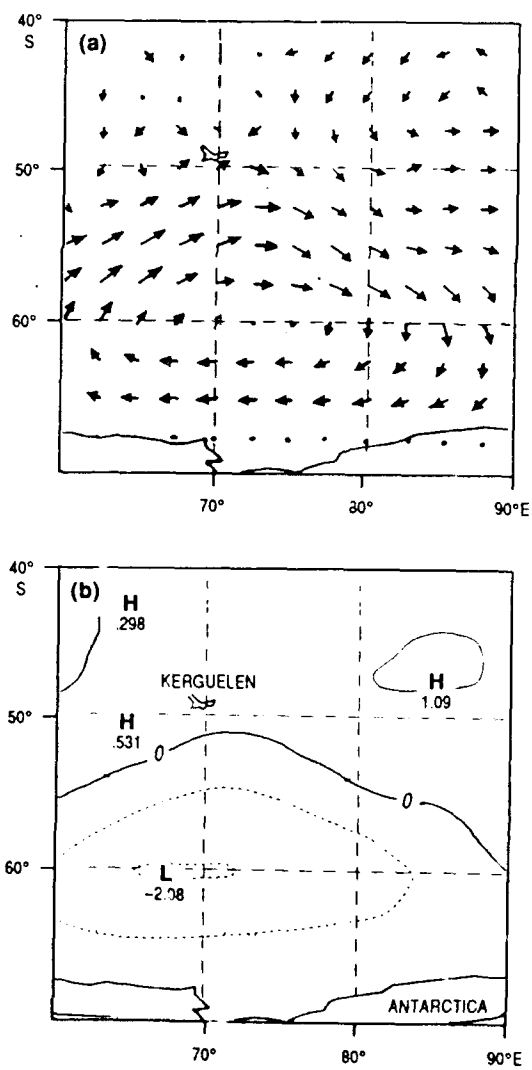
SSM/I Superob Wind Speeds			
	Ob — Background ($m\ s^{-1}$)		
SSM/I ($m\ s^{-1}$)	Mean	St Dev	rms
0– 5	–0.4	2.1	2.1
5–10	1.2	2.1	2.4
10–15	3.4	2.7	4.3
15–20	6.6	2.2	7.0
Ship/Buoy Wind Speeds			
	Ob — Background ($m\ s^{-1}$)		
SSM/I ($m\ s^{-1}$)	Mean	St Dev	rms
0– 5	–0.8	2.1	2.2
5–10	1.7	2.1	2.7
10–15	4.7	2.5	5.3
15–20	6.7	2.0	6.9

However, a similar bias is present in the statistics for other types of surface wind data increments, indicating that the model tends to underestimate the stronger wind speeds. This bias is most likely due to the fact that many of these high wind speed events cannot be well represented by the resolution of the current model. However, the source of this bias is under investigation.

Ten days after NOGAPS incorporated the SSM/I data, the same statistical computations were performed once again (Table 6). Comparison to Table 5 shows that, for all wind speeds, the observation minus background differences have decreased over time. Thus, we conclude that the forecast system is accepting the data. As a means of independent verification, we also computed the vector rms errors of observation minus background increments for the ship and buoy data before and after the inclusion of the SSM/I wind speeds. From Table 7, it is apparent that the (*1000 mb*) background also agrees more closely with these independent data sources after the SSM/I data were introduced. The largest change is observed in the Southern Hemisphere, as would be expected, since any new data set will have a more significant impact in data sparse areas.

To make the comparison to independent data sources more consistent, we compared the SSM/I superob wind speed statistics to

FIGURE 6. Difference fields of (*a*) 1000 mb wind vectors (maximum vector corresponds to 5 m/s) and (*b*) sea-level pressure, contoured every mb. Differences are for an analysis *with* SSM/I data minus an analysis *without* SSM/I wind data.



the same statistics for ship and buoy *wind speeds* for the 1200 UTC analyses from May 24 to 28, 1991. The mean, standard deviation, and rms error for the observation minus (10 m) background differences are shown in Table 8. The SSM/I data do not appear to be biased relative to the other marine surface wind observations, and their difference statistics are comparable. When compared to Table 4, it is also clear that the SSM/I wind speeds agree more closely with the preliminary analysis than with the analysis background field, further showing consistency between the SSM/I data and the other marine wind observations. While the statistics presented here are from a limited time period, there are only slight day-to-day variations in these global values.

Finally, Figure 6 shows the impact of the SSM/I wind speeds on the analysis of a low pressure system in the Southern Hemisphere. Inclusion of the SSM/I data increases the strength of the flow around the low (Figure 6a), and through the multivariate coupling of the analysis, consistently decreases the central pressure of the low by 2 mb (Figure 6b). While this may seem like a small difference, such differences in the analyzed fields can often make large differences in the model forecasts. As an example, SSM/I data were withheld from the data assimilation cycle for 48 hrs; then a 48-hr forecast was made from these initial fields. The forecast without the SSM/I data was differenced from the NOGAPS 48-hr forecast valid at the same time from the operational run. As can be

FIGURE 7. Difference fields of (*a*) 1000 mb wind vectors (maximum vector corresponds to 10 m/s) and (*b*) 1000 mb heights, contoured every 10 m. Differences are for a 48-hr forecast made *with* minus a 48-hr forecast made *without* SSM/I wind data. In the *without* SSM/I data case, the observations were withheld for 48 hrs of the data assimilation cycle prior to making the comparison forecast.

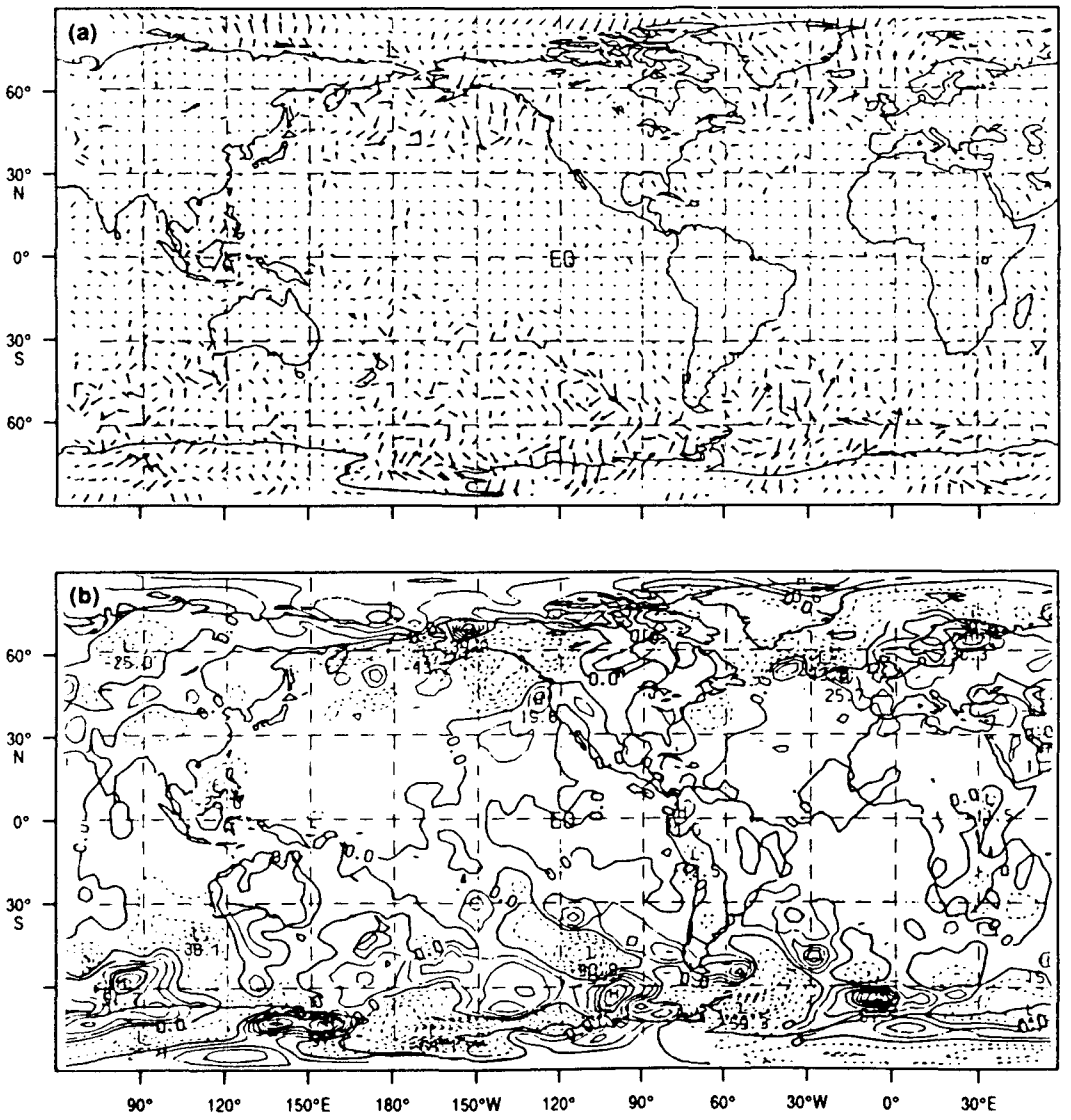
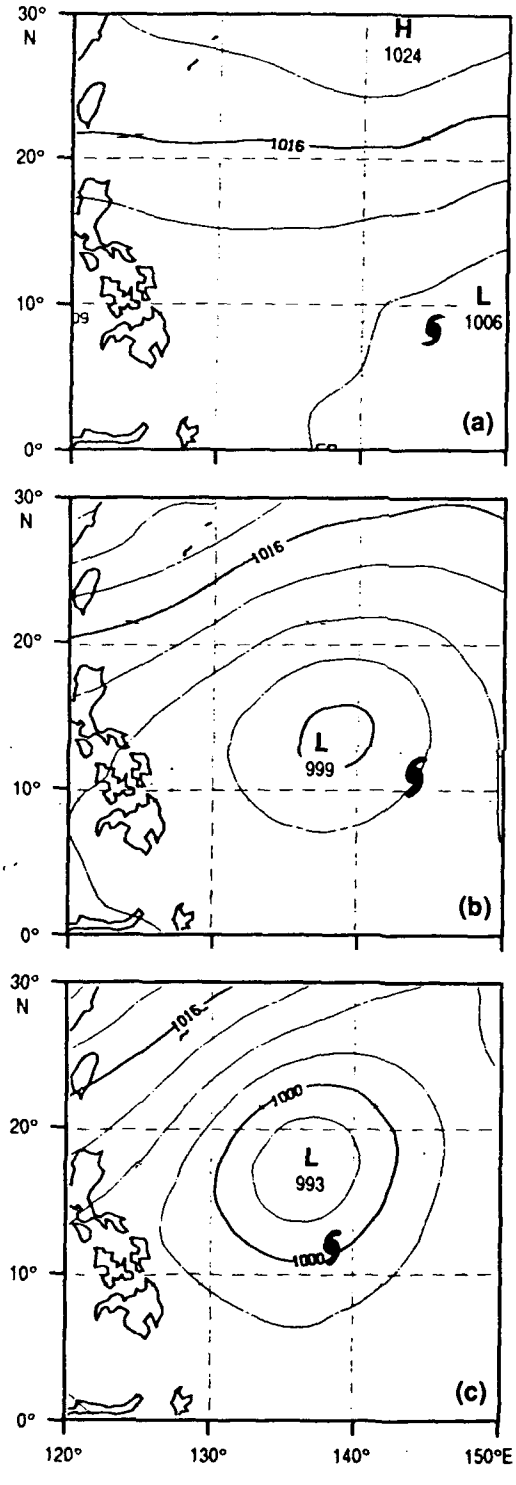


FIGURE 8. NOGAPS depiction of supertyphoon Page prior to the assimilation of any synthetic observations. The small hurricane symbol marks the actual storm position. (*a*) Analysis valid at 0000 UTC, November 19, 1990; (*b*) 72-hr forecast valid at 0000 UTC, November 22, 1990; (*c*) 120-hr forecast valid at 0000 UTC, November 24, 1990.



seen from Figure 7, the differences in the wind field are as large as 20 m s^{-1} and the 1000 mb height field exhibits differences as large as 82 m. This height difference would translate to about an 11 mb difference in surface pressure, just from excluding SSM/I data for 48 hrs. Although the impact of the SSM/I data appears to be the greatest near the poles and in the northern mid-latitudes, sizeable differences are also seen in the tropical winds in the western Pacific and Indian oceans.

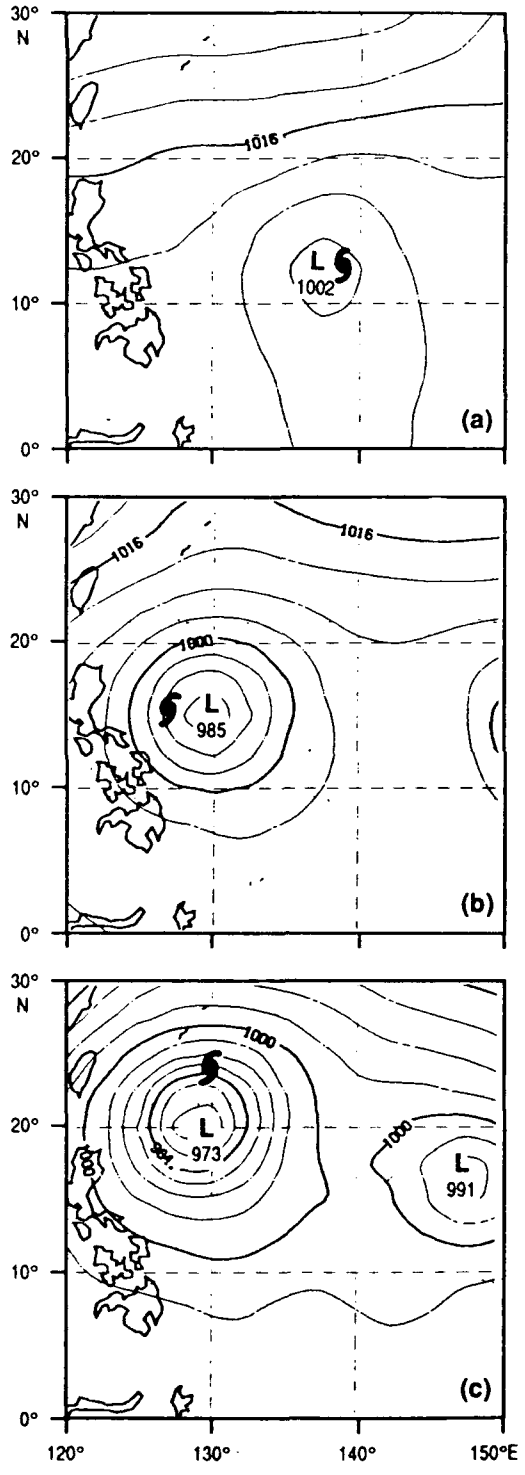
Impact of Synthetic Tropical Soundings

In 1989, the resolution of the forecast model was increased and improvements were made in the parameterization of the diabatic processes (Hogan and Rosmond, 1991). As a result of these changes, NOGAPS began actively developing and forecasting tropical storms. The ability of the forecast model to develop, intensify, and forecast the movement of tropical circulations accentuated the need for properly positioning the storms in the NOGAPS initial fields. It was also a necessary ingredient for successful tropical data assimilation. If the model could maintain these initial circulations throughout the forecast period, then synthetic observations could be used to insert the developing storms into the analyzed fields.

To illustrate the model's capabilities *prior to* the use of synthetic data, Figure 8 displays the NOGAPS sea-level pressure analysis product and the 72-hr and 120-hr forecast products made from this analysis. At 0000 UTC, November 19, 1990, a 1006 mb low is analyzed about 500 km from the JTWC position of a tropical depression that eventually developed into supertyphoon Page (winds exceeding 66 m/s). The 72-hr NOGAPS forecast valid at 0000 UTC on November 22 has developed a 999 mb low. By this time, in reality Page had reached tropical storm strength with maximum winds of 18 m/s, and the error in the forecast position is approximately 750 km. After 120 hrs, the storm has deepened to a 993 mb low (Figure 8c). At 0000 UTC, November 24, Page was actually located about 550 km from the forecast position and had maximum wind speeds of 31 m/s. Even though the error in the forecast position is fairly large, the model's development of the storm is notable.

Consider the model's depiction of the same storm a few days later (Figure 9). In this case, synthetic observations have been generated for Page since it reached tropical storm strength at 0000 UTC, November 22. These observations were assimilated into NOGAPS beginning at 1200 UTC, November 22. The 0000 UTC analysis product on November 24 shows a 1002 mb low located approximately 100 km from the JTWC position of Page, which at this time is approaching typhoon strength with maximum

FIGURE 9. NOGAPS depiction of supertyphoon Page after the assimilation of synthetic observations. The small hurricane symbol marks the actual storm position. (*a*) Analysis valid at 0000 UTC, November 24, 1990; (*b*) 72-hr forecast valid at 0000 UTC, November 27, 1990; (*c*) 120-hr forecast valid at 0000 UTC, November 29, 1990.

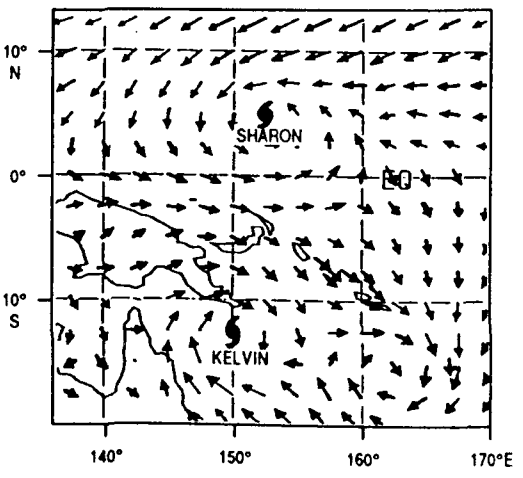


wind speeds of 31 m/s. The 72-hr forecast develops a 985 mb low located about 300 km from the actual storm position. By this time, 0000 UTC, November 27, Page was at its peak intensity with maximum winds of 72 m/s. The 120-hr forecast shows a 973 mb low whose center is located about 400 km from the verifying position of Page at 0000 UTC, November 29. The storm was still a damaging typhoon as it made landfall in Japan on November 30.

Comparing Figures 8 and 9, the analysis/forecast products with no synthetic data had position errors of 500 km, 750 km, and 550 km, while the later series had track errors of 100 km, 300 km, and 400 km. While the initial fields of surface pressure differed by only 4 mb, the series of forecast products generated from the analysis using synthetic data produced a storm whose central pressure was 20 mb lower at 120 hrs. Thus, the assimilation of the synthetic observations produced a more accurate representation of the storm in the analyzed field, resulting in more accurate forecasts of the storm's track and its strength. Even though the circulations generated by NOGAPS are larger than those observed in nature, clearly the forecast model develops and intensifies tropical circulations. This case illustrates a very dramatic deepening that is not typical of most tropical circulations. However, in the few cases where such a strong deepening has been observed in the NOGAPS forecasts prior to the assimilation of synthetic data, almost all were associated with storms that approached or reached supertyphoon intensity.

Another example of the quality of the NOGAPS analysis product in the tropics is illustrated in Figure 10, which shows a 1000 mb wind nowcast in the western equatorial Pacific, where

FIGURE 10. 1000 mb wind analysis valid at 0000 UTC, March 5, 1991, showing the coupled tropical circulations that are associated with the strong zonal westerly wind anomalies along the equator in the western Pacific.



a tropical cyclone couplet is developing. These couplets are of interest because the concentration of strong westerly winds along the equator between the two storms can generate ocean Kelvin waves that may play a role in the El Ninō warming in the central and eastern equatorial Pacific (Keen, 1982). The nowcast shown in Figure 10 is valid at 0000 UTC, March 5, 1991. In the Southern Hemisphere, Kelvin reached tropical storm strength on March 3. Since that time, the NOGAPS analysis used synthetic soundings to position and enhance Kelvin's circulation. However, at the time of this nowcast, the Northern Hemisphere storm was just beginning to develop. It did not become tropical storm Sharon until 0600 UTC on March 7. Accordingly, synthetic observations for Sharon were not used prior to that time. Even so, the equatorial westerly wind anomaly associated with these two storms is well depicted. This westerly wind event was verified by independent island reports that, at the time, were not included in the NOGAPS analysis product. The ability of an atmospheric model to depict such events in their formative stages may be extremely useful to climatologists, given the typical sparseness of data in the region of interest in the western Pacific (Harrison and Giese, 1991).

SUMMARY

The Navy's operational global data assimilation system has shown a notable improvement over the past few years, and each component of the data assimilation system has contributed to the overall performance. Recent improvements in the data analysis component include a better specification of the observation errors for marine surface data and the inclusion of new sources of data, including SSM/I and Pacific island winds and synthetic tropical soundings. As we continue to monitor the performance of the multivariate optimum interpolation analysis, we expect to further fine-tune the analysis statistics. We are also prepared to assimilate new data sources as they become available. Finally, improved coupling of the oceanographic and atmospheric data assimilation systems will continue to be at the forefront of future research efforts.

ACKNOWLEDGMENTS

These improvements to the operational system would not have been possible without the efforts of N. Baker, R. Brody, and R. Hodur of NRL, C. Mauck of FNOC, G. Bayler and B. Stone, formerly with FNOC, and H. Lewit of Computer Sciences Corporation. We gratefully acknowledge the support of our sponsors, the Space and Naval Warfare Systems Command (PMW-165), Capt. C. Hoffman, USN, Program Element 63207N, and the Office of Naval Technology, Cmdr. L. Bounds, USN, Program Element 62435N. Contribution Number 92:040:431. Approved for public release; distribution is unlimited.

REFERENCES

Andersson, E. and Hollingsworth, A. 1988. Typhoon Bogus Observations in the ECMWF Data Assimilation System. ECMWF Res. Dept. Techn. Memo #148. 25 pp.

Atlas R., Bloom, S., Hoffman, R., Ardizone, J. and Brin, G. 1991. Space-based surface wind vectors to aid understanding of air-sea interactions. *EOS, Trans., American Geophysical Union*, 72:201–208.

Baker, N. 1991. Quality control of meteorological observations for the Fleet Numerical Oceanography Center operational atmospheric database. NOARL Technical Note 124, Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, MS, 39529, 59 pp.

Baker, N. 1992. Data quality control for the Naval Operational Global Atmospheric Prediction System. *Wea. and Forecasting*, 7:250–261.

Barker, E. 1992. Design of the Navy's multivariate optimum interpolation analysis. *Wea. and Forecasting*, 7:220–231.

Clancy, R.M. 1987. Real-time applied oceanography at the Navy's global center. *Marine Technology Society Journal*, 21:33–46.

Goerss, J. 1989. The impact of bogus observations upon the Navy Operational Global Atmospheric Prediction System. In: *Preprints, Twelfth Conference on Weather Analysis and Forecasting*, pp. 41–45. Monterey, California: American Meteorological Society.

Goerss, J. and Phoebus, P. 1992(a). The Navy's operational atmospheric analysis. *Wea. and Forecasting*, 7:232–249.

Goerss, J. and Phoebus, P. 1992(b). The Multivariate Optimum Interpolation Analysis of Meteorological Data at FNOC. NOARL Report 31, Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, MS, 39529, 60 pp.

Goerss, J., Brody, L. and Jeffries, R. 1991. Assimilation of tropical cyclone observations into the Navy Operational Global Atmospheric Prediction System. In: *Proceedings, Ninth Conference on Numerical Weather Prediction*, pp. 638–641. Denver, Colorado: American Meteorological Society.

Goodberlet, M., Swift, C. and Wilkerson, J. 1990. Ocean surface wind speed measurements of the Special Sensor Microwave/Imager (SSM/I). *IEEE Trans. Geosci. Remote Sensing*, 28:823–828.

Harrison, D. and Giese, B. 1991. Episodes of surface westerly winds as observed from islands in the western tropical Pacific. *J. Geophys. Res.*, 96: 3221–3237.

Hogan, T. and Rosmond, T. 1991. The description of the Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.*, 119: 1786–1815.

Hogan, T., Rosmond, T. and Gelaro, R. 1991. The NOGAPS Forecast Model: A Technical Description. NOARL Report 13, Naval Oceanographic and

Atmospheric Research Laboratory, Stennis Space Center, MS, 39529, 223 pp.

Hollingsworth, A. and Lönnberg, P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, 38A:111–136.

Keen, R. 1982. The role of cross-equatorial tropical cyclone pairs in the Southern Oscillation. *Mon. Wea. Rev.*, 110:1405–1416.

Lorenc, A. 1981. A global three-dimensional multivariate statistical interpolation. *Mon. Wea. Rev.*, 109:701–721.

Phoebus, P. and Goerss, J. 1991. The operational assimilation of SSM/I wind speed data. In: *Proceedings, Ninth Conference on Numerical Weather Prediction*, pp. 569–572. Denver, Colorado: American Meteorological Society.

Yu, T. 1987. A technique for deducing wind direction from satellite microwave measurements of wind speed. *Mon. Wea. Rev.*, 115:1929–1939.