UCAR

NOARL

Institute for
Naval Oceanography

AD A239801

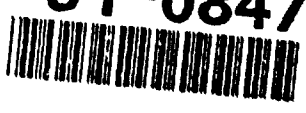
VERMOD CAPABILITIES VERMOD 1.0 AND FUTURE

HARSH ANAND
MICHAEL S. FOSTER
RAMESH KRISHNAMAGARU
RANJIT M. PASSI

INSTITUTE FOR NAVAL OCEANOGRAPHY

JUNE 1991

91-08470



91 8 21 046

The Institute for Naval Oceanography is operated by the University Corporation for Atmospheric Research (UCAR) under sponsorship of the Naval Oceanographic and Atmospheric Research Laboratory (NOARL). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of NOARL.

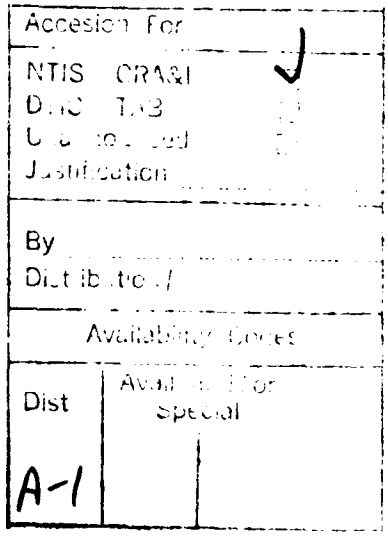


TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

The Institute for Naval Oceanography (INO) has established the Experimental Center for Mesoscale Ocean Prediction (ECMOP) to facilitate research and development of ocean models. ECMOP aims to take advantage of existing infrastructures to provide support to Navy and academic researchers in an end-to-end model evaluation effort. In addition, if and when necessary, ECMOP performs internal research and development as an adjunct to this support. ECMOP is composed of three functional subsystems, or modules: the Verification Module (VERMOD), the Visualization Module (VISMOD), and the Data Sub-System (DASS). This technical memorandum describes the current state of VERMOD, in particular, its statistical and physical criteria for model verification, and the software structural characteristics of the module. Finally, we discuss several of the verification concepts anticipated as future enhancements of VERMOD.

ACKNOWLEDGMENTS

We would like to thank Mr. James Corbin for his comments on an earlier draft and Ms. Lydia Harper for her editorial help.

1. INTRODUCTION

The Institute for Naval Oceanography (INO) has established the Experimental Center for Mesoscale Ocean Prediction (ECMOP) to provide capabilities that:

- permit development, demonstration, and evaluation of mesoscale ocean prediction models;
- furnish information pertinent to the transition of ocean modeling systems to Navy sponsors (and "decision makers") in an objective, orderly, and efficient manner; and
- are modular in nature and can be transitioned as appropriate to operational Navy.

While adhering to operational concepts developed by the Naval Oceanography Command, these capabilities are being realized through a modular structure that emphasizes a standardized interface among data, numerical models, and processing systems. ECMOP is comprised of three functional sub-systems: the Verification Module (VERMOD) provides objective evaluation routines, the Visualization Module (VISMOD) defines user definable graphic routines, and the Data Sub-System (DASS) provides data acquisition and management. In all areas of ocean modeling, whether the effort is R&D or operational, there is a need for basic technical support capabilities for common functions to acquire, manage and analyze data; in situ, remotely sensed, or model generated. All these functions incorporate some aspects of three basic technical capabilities: (1) data base management; (2) graphical and/or visualization schemes; and objective evaluation routines. The role of ECMOP is to focus on these three basic technical capabilities commensurate with the needs of the R&D community by way of building them into a state-of-the-art facility using leading edge technology in software and hardware. At the same time, ECMOP is to evaluate these state-of-the-art capabilities for applicability to the operational community and when identified, make them available for transition. The conceptual details of ECMOP were initially given by Leese (1988,a,b). The present ECMOP configuration, in its modular form, is shown in Fig. 1, wherein the three modules communicate externally using the network common data form (netCDF) jackets.

This technical memorandum documents the current status of VERMOD. VERMOD was originally developed with a minimal number of statistical and physical tools for model verification and evaluation in order to substantiate the software design characteristics. These tools are accessed through an intuitive graphical user interface (GUI) that requires little user indoctrination or training. The GUI functions within the X-windows environment as implemented under the UNIX operating system. Its features are accessible to heterogeneous environments and across computer networks and platforms. Enhancements and modifications will be made to VERMOD according to user requirements. Guidance for crucial enhancements has been

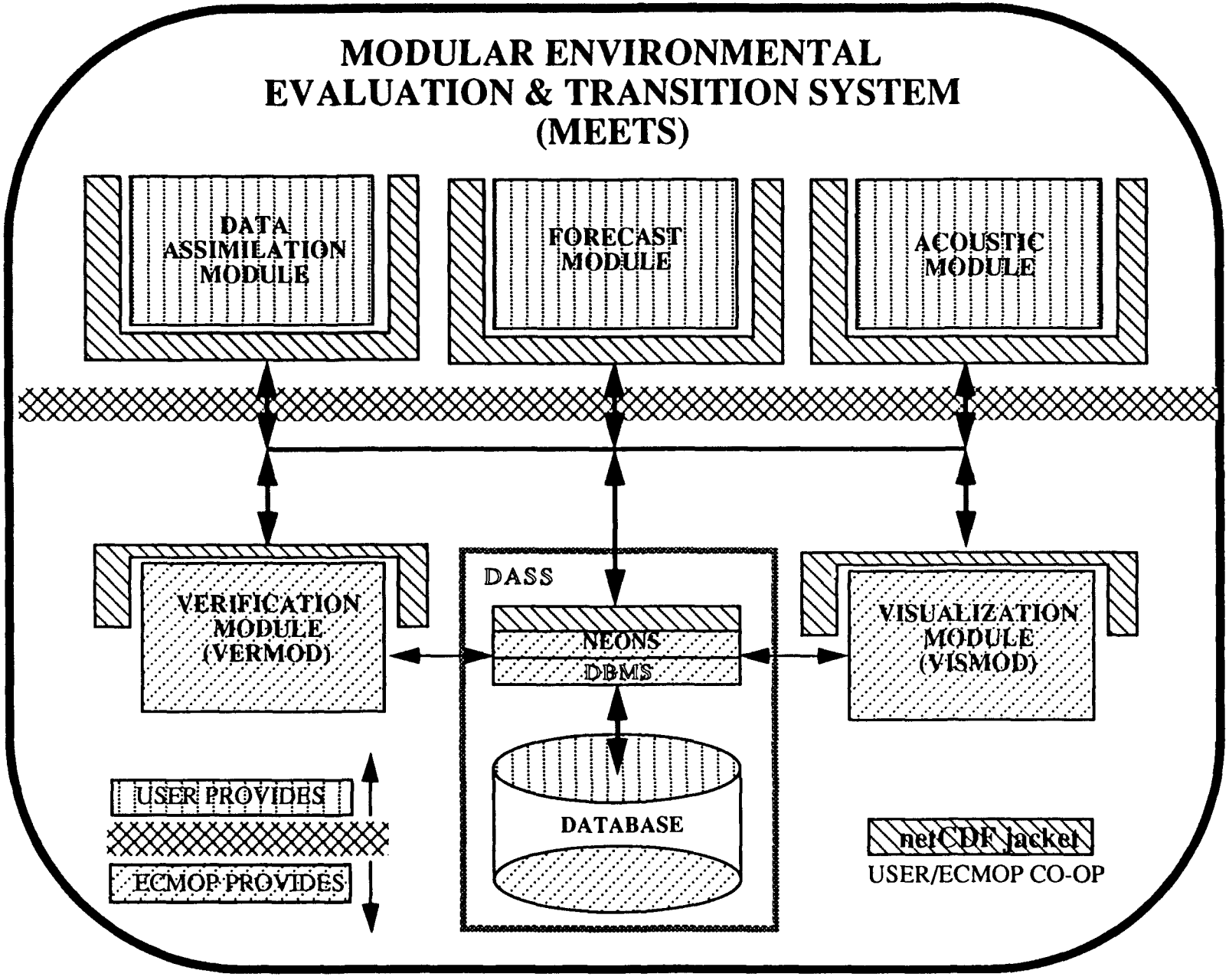


Fig. 1: ECMOP Generic System Configuration

derived from the INO Summer Colloquium on Model Evaluation (INO, 1989), as well as input from the community of scientific users.

The statistical measures included in VERMOD 1.0 are the Root Mean Square (RMS) error with both systematic and unsystematic decompositions (Wilmot et al., 1985), pattern correlation, and two criteria for gauging the departure from normality, skewness and kurtosis. These techniques are described in the text and illustrated with computer simulated examples.

Section 2 gives a brief description of the software architecture. Description of the statistical measures is presented in Section 3, which is followed, in Section 4, by three examples that illustrate the concepts. Finally, Section 5 briefly examines possible future enhancements to VERMOD.

2. VERMOD ARCHITECTURE

The VERMOD application embodies characteristics drawn from the other modules within ECMOP. The software, its design and the supporting hardware are briefly discussed below.

2.1 Software

All ECMOP software, including VERMOD, is designed to execute under the UNIX operating system. To function properly, VERMOD must utilize services provided by the remaining ECMOP modules. Therefore, an understanding of VERMOD requires a description of ECMOP, both in its present configuration and that configuration to which it is evolving as a complete software package.

The GUI, which controls all modules, is implemented within X-windows. The core of DASS, the data support module, is "Empress", a proprietary relational database application that implements the structured query language (SQL) and features a capability to handle bulk data. Access to the data is via the Navy Environmental Operational Nowcast System (NEONS), an interface to "Empress" written in C to support high level access to the data stored within the data base. NEONS is described by Computer Sciences Corporation (1991). The VISMOD is currently built upon the National Center for Atmospheric Research (NCAR) application NCAR Graphics. Model verification and evaluation within the VERMOD are performed using established and widely accepted quantitative and qualitative evaluation techniques as implemented through FORTRAN subroutines and C functions. As a complete package, ECMOP software is written in "C" and "FORTRAN" with the support of SQL and C-shell scripts. The X-windows GUI uses the Open Software Foundation's "Motif" ("OSF/Motif") widget set. These features give the ECMOP software increased transparency across networks and portability across a variety of platforms that support X-windows.

2.2 Hardware

The ECMOP software has been developed on Sun Microsystems hardware. It is easily portable to other Unix-compatible platforms such as those manufactured by Hewlett-Packard, Silicon Graphics and Cray. This portability is made possible because "Empress", NCAR graphics and X-windows are available on these and other platforms.

2.3 Conceptualization

The ECMOP design is evolving into a modular structure. Currently, ECMOP modules are loosely linked; however, redesign efforts will require user control through the GUI, which operates as the front end to the ECMOP System Control Module (SCM). The rationale underlying ECMOP redesign is the need to support interaction among the several modules. The ECMOP SCM will integrate functional control of the modules while enabling communications among them. All existing modules are to be enhanced. In addition, a User Help Module (UHM) will provide context sensitive assistance to the user at any point within the application. The DASS is being redesigned into a Data Management Module (DAMM) that will include such features as on-line directories, conversion utilities and NEONS data ingestion. A Model Control Module (MCM) is to be added which will allow ocean models to execute within the ECMOP environment. The VISMOD will be upgraded to include capabilities that allow three-dimensional visual analysis. Enhancements to VERMOD are discussed in this document.

3. STATISTICAL MEASURES

In this section we present the derivations of the statistical measures that are included in VERMOD 1.0, viz, the decomposition of root mean square error (RMSE) into systematic and unsystematic components, the pattern correlation and the two simple measures of departure from normal distribution.

3.1 Preliminaries

We would like to compare two fields, m and y. To compare the model output with observed data, we may denote the model output by the symbol m and the observational data by the symbol y. For a proper comparison, the two fields must correspond to each other in some respect. This correspondence may be in terms of a time series for fixed spatial coordinates so that one may compare over time to obtain a spatial comparison. Another simple example

would be to compare model output fields at two different points in time. In this case the two symbols represent the same field, i.e., the model output, and the comparison yields a temporal difference. The more common need will be to compare model output with observed data. This requires that the two fields correspond in spatial distribution to provide a model-data comparison at a given time. However, this required correspondence between model output and data usually does not exist, since the data are not usually available at the model grid. Some numerical manipulation may be necessary to achieve this correspondence. When the resolution of observational data is finer than the model grid, as in the case of Multi-Channel Sea Surface Temperature (MCSST), some (optimum) interpolation of the data may be performed to align the two fields to the model grid. Otherwise, which is usually the case, the model output data are interpolated to observation locations to enable a valid comparison. This realignment introduces an error in addition to those present in either the model output or the observational data.

Note: VERMOD 1.0 provides capability of RMSE computations when the two fielded variables (model output and observations) are already properly aligned to the same grid frame.

We use the following notation. The mean values of m and y are denoted by μ_m and μ_y such that $\mathcal{E}(m) = \mu_m$ and $\mathcal{E}(y) = \mu_y$, where $\mathcal{E}$ is the statistical expectation operator. Corresponding variances are defined by

$$\sigma_m^2 = \mathcal{E}(m - \mu_m)^2 \text{ and } \sigma_y^2 = \mathcal{E}(y - \mu_y)^2. \tag{3.1}$$

3.2 The RMS Criterion

The RMS error criterion is easily understood because its square, the mean square error (MSE), is associated with the variance and the RMSE provides a direct measure of the difference between the two fields being compared. We first define the MSE between two variables, m and y, to be

$$MSE = \mathcal{E}(m - y)^2 = \int_\Omega (m - y)^2 f(\omega) d\omega \tag{3.2}$$

where $f(\omega)$ is the probability density function, or a weighting function defined on the space $\{\omega \in \Omega\}$ over which $m(\omega)$ and $y(\omega)$ are defined. For discrete grid matched variables m and y, we can approximate the mean square error as

$$MSE \approx \sum_{i=1}^{N} w_i (m_i - y_i)^2 \tag{3.3a}$$

where w_i are the weights such that $\sum_{i=1}^{N} w_i = 1$. Weights can be varied to emphasize subsets within regions of interest. For example, in the North Atlantic basin one may choose to emphasize the Gulf Stream by specifying larger weights for that region. For equal weighting $w_i = \frac{1}{N}$:

$$MSE \approx (1/N)\sum_{i=1}^{N}(m_i - y_i)^2. \tag{3.3b}$$

3.2.1 Systematic and unsystematic RMSE

Decomposition into systematic and unsystematic RMS errors, as discussed by Wilmot et al. (1985), has been quite popular in oceanography, and was recommended for implementation in VERMCD by the Working Group on Statistical Procedures during the INO Summer Colloquium on Model Evaluation (1989). The systematic error component identifies any bias between the two fields being compared, and the unsystematic component measures the goodness of fit.

The motivation for this decomposition derives from two common comparison scenarios: (a) the model output and the data are comparable within some random error, and (b) the model output is within a fixed bias plus random error. For the case (a) we will write

$$m_i = y_i + e_i$$

where e_i is the random error. For the case (b) we can express the relationship

$$m_i = \beta_0 + y_i + e_i$$

where β_0 is a fixed bias. These two conditions can be generalized to the following linear relationship:

$$m_i = \beta_0 + \beta_1 y + e_i. \tag{3.4}$$

From this the RMSE decomposition is obtained as follows. Let $\hat{m}$ be the ordinary least squares estimate of m computed from the regression of m on y. Then, the coefficients β_j, $j = 1, 2$ are estimated by minimizing the sum of squares $\sum_{i=1}^{N}(m_i - \beta_0 - \beta_1 y)^2$. Let b_0 and b_1 be the minimizing estimates of these coefficients. The predicted value is given by:

$$\hat{m}_i = b_0 + b_1 y_i. \tag{3.5}$$

The sum of squares of the differences between the two fields can then be written as:

$$\begin{aligned}\sum_{i=1}^{N} w_i(m_i - y_i)^2 &= \sum_{i=1}^{N} w_i[(m_i - \hat{m}_i) + (\hat{m}_i - y_i)]^2 \\ &= \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)^2 + \sum_{i=1}^{N} w_i(\hat{m}_i - y_i)^2 + \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)(\hat{m}_i - y_i).\end{aligned}$$

As shown in the Appendix, the cross-product term $\sum_{i=1}^{N} w_i(m_i - \hat{m}_i)(\hat{m}_i - y_i) = 0$. Thus, the above equation can be written as a sum of two components, the systematic and the unsystematic components, given by:

$$\sum_{i=1}^{N} w_i(m_i - y_i)^2 = \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)^2 + (\hat{m}_i - y_i)^2. \tag{3.6}$$

The unsystematic component is given by

$$\mathrm{RMSE}_u = \left[\sum_{i=1}^{N} w_i(m_i - \hat{m}_i)^2\right]^{\frac{1}{2}} \tag{3.7}$$

and the systematic component is given by

$$\mathrm{RMSE}_s = \left[\sum_{i=1}^{N} w_i(\hat{m}_i - y_i)^2\right]^{\frac{1}{2}}. \tag{3.8}$$

Note that

$$\mathrm{RMSE}^2 = \mathrm{RMSE}_u^2 + \mathrm{RMSE}_s^2. \tag{3.9}$$

Wilmot et al. define RMSE_s as the linear bias and RMSE_u as a precision criterion. The above decomposition can be performed at any depth, Z. Suppose the model output data m is available in the σ coordinate system while the other field, y, is available at the Z level. Here, the model data at each gridpoint are interpolated in the vertical at the Z level using a spline function, thus leading to grid-to-grid RMSE evaluation using (3.6).

Note: In their formulation, Wilmot et al. assume there is no measurement error in y. Without this assumption, the least squares estimation leads to a complicated statistical analysis, and the final solution is not that clean.

3.2.2 Grid-to-observations (option currently not available)

In this case the observations $y_i, i = 1, \ldots, N$ at a particular Z level are not coincident with the model grid. To compute RMSE the model output are first interpolated in the vertical at the Z level. A further interpolation is then performed to derive model output values at the observation points prior to RMSE calculations.

3.3 Pattern Correlation

Pattern correlation is used to estimate how closely the pattern of one field resembles that of another. Again, if m_i and $y_i, i = 1, \ldots, N$ are the values of the two fields, then the coefficient of pattern correlation is given by

$$\rho = \frac{\sum_{i=1}^{N}(m_i - \bar{m})(y_i - \bar{y})}{[\sum_{i=1}^{N}(m_i - \bar{m})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2]^{\frac{1}{2}}}. \tag{3.10}$$

This concept tacitly identifies if there is a linear relationship between the two fields. A high positive value of ρ indicates that the patterns of highs and lows of the two fields match and that the two fields are almost linear transformations of each other. A high negative value, on the other hand, indicates that the highs of the one field correspond to the lows of the other.

The pattern correlation measure is connected with the Wilmot type decomposition by the following relationship:

$$\begin{aligned} \mathrm{RMSE}_u^2 &= \sum_{i=1}^{N} w_i (m_i - \hat{m}_i)^2 \\ &= \sigma_p^2 (1 - \rho^2). \end{aligned} \tag{3.11}$$

Equation (3.11) says that the model prediction m can be described by $\hat{m}$ more precisely if the pattern correlation is large.

3.4 Departure from Normality Measures

During the Trial Ocean Prediction Experiment (TROPE), Waters et al. (1990) recommended to include in VERMOD measures of skewness and kurtosis that provide information on how closely the differences between two fields resemble normality. Skewness is defined as

$$\gamma_1 = \frac{\mu_3}{\mu_2}$$

and kurtosis is defined as

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

where μ_j is the jth central moment computed from the field differences. Skewness tells whether the differences are asymmetrical; a positive/negative value indicates that the distribution is skewed to the right/left. The Kurtosis measure indicates the flatness of the distribution. For a normal distribution $\gamma_1 = 0$ indicating symmetrical distribution, and $\gamma_2 = 1$. A negative value of γ_2 shows that the distribution is more peaked than the normal distribution, i.e., the distribution is concentrated more around the smaller differences.

4. EXAMPLES

To illustrate the use of the RMSE decomposition and pattern correlation we will use sea surface height data from a nowcast for the Gulf of Mexico as the model output. The observation data will be simulated using three sets of β_0 and β_1 by

$$y_i = (m_i - \beta_0)/\beta_1$$

followed by generation of the model output by using (3.4). The noise added in (3.4) is Gaussian with zero mean and standard deviation σ_n= 1 cm. The three sets of (β_0, β_1) will provide interpretation for three different situations, one corresponding to zero systematic bias and two situations corresponding to when it is non-zero.

4.1 The Three Cases

Case A: $RMSE_s$=0

This situation arises when, in (3.4), $\beta_0 = 0$ and $\beta_1 = 1$. For the ideal estimation situation, $b_0 \approx 0$ and $b_1 \approx 1$. Then the predicted model values are given by

$$\hat{m}_i \approx y_i.$$

Thus:

$$\text{RMSE}_s^2 = \sum_{i=1}^{N}(\hat{m}_i - y_i)^2 \approx \sum_{i=1}^{N} w_i(y_i - y_i)^2 = 0.$$

The unsystematic component in this case is given by:

$$\text{RMSE}_u^2 = \sum_{i=1}^{N}(\hat{m}_i - m_i)^2 \approx \sum_{i=1}^{N} w_i(m_i - y_i)^2.$$

Simulations for this case were performed using $\beta_0 = 0$ and $\beta_1 = 1$. The estimated values of these parameters (Table 1) are $b_0 = -0.008$ and $b_1 = 0.996$, which are close to the actual parameter values. The simulated model output and observations are presented in Figs. 2a and 2b. The sea surface height patterns in the two are almost identical but for the fact that, due to the simulation effect, the model output appears ragged.

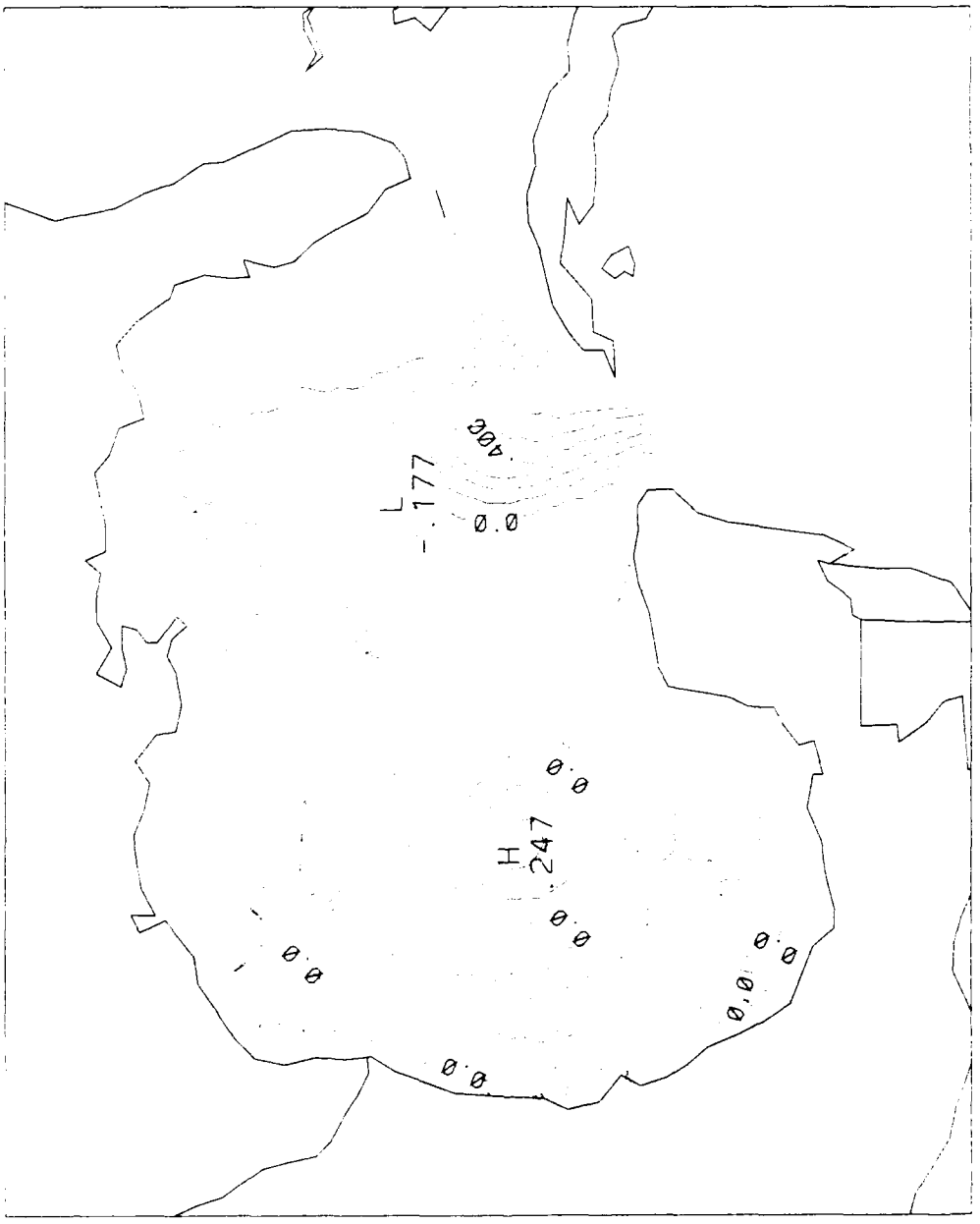


Fig. 2a: Sea Surface Height contours for simulated model output; Case A: $\beta_0 = 0, \beta_1 = 1$

Fig. 2b: Sea Surface Height contours for simulated observations; Case A: $\beta_0 = 0.10, \beta_1 = 1.00$

Table 1: Results of Simulation Examples

	(β_0, β_1)	(b_0, b_1)	(σ_u, σ_s)	Total RMSE	ρ	Skewness	Kurtosis
A	(0.00, 1.00)	(-0.001, 1.000)	(0.010, 0.001)	0.011	0.997	-0.016	3.457
B1	(0.10, 1.00)	(0.099, 1.000)	(0.010, 0.099)	0.100	0.997	-0.011	3.430
B2	(0.20, 0.80)	(0.199, 0.800)	(0.010, 0.252)	0.252	0.997	-2.272	9.352

The computed values of RMSE_u is 0.010, the same as σ_n, while RMSE_s is 0.001, which is quite close to zero. The RMSE decomposition is in keeping with our expectation.

The computed value for pattern correlation in this case is fairly high, 0.997. The magnitude of the correlation is a function of the noise level (standard deviation). Since the noise level is the same for the three cases, the computed pattern correlation has a constant value.

Note that the RMSE_s is 0 only if both conditions, $\beta_0 = 0$ and $\beta_1 = 1$, are satisfied. To see this we note that if $\beta_0 = 0$ but $\beta_1 \neq 1$, then

$$\hat{m}_i \approx \beta_1 y_i,$$

which leads to $\text{RMSE}_s = (\beta_1 - 1)^2 \sum_{i=1}^{N} y_i^2 \neq 0$.

On the other hand, if $\beta_0 \neq 0$ and $\beta_1 = 1$, then $\hat{m}_i \approx \beta_0 + y_i$ which yields

$$\text{RMSE}_s^2 = \sum_{i=1}^{N} w_i(\beta_0 + y_i - y_i)^2 = \sum_{i=1}^{N} w_i \beta_0^2 = \beta_0^2 \sum_{i=1}^{N} w_i = \beta_0^2 \neq 0.$$

Case B: $\text{RMSE}_s \neq 0$

The non-zero systematic RMSE component is marked with either $\beta_0 \neq 0$ or $\beta_1 \neq 1$. When $\beta_0 \neq 0$, the case of interest is $\beta_1 = 1$. All other cases of $\text{RMSE}_s \neq 0$ are covered under $\beta_1 \neq 1$. These two situations are described below.

Case B1: $\text{RMSE}_s \neq 0$ with $\beta_1 = 1$

This requires $\beta_0 \neq 0$. This is the pure bias situation wherein the model output has a fixed bias, β_0, and a random error, ϵ_i. In this case (3.4) is written as:

$$m_i = \beta_0 + y_i + \epsilon_i. \tag{3.14}$$

This case was discussed above and yields $\text{RMSE}_s^2 \approx \sum_{i=1}^{N} w_i \beta_0^2 = \beta_0^2$. Thus, $\text{RMSE}_s \approx \beta_0$, as one would expect. The simulations for this case were performed using $\beta_0 = 0.10$ and $\beta_1 = 1$. The computed values for their estimates are $b_0 = 0.099$ and $b_1 = 1.00$. As indicated earlier, the computed value of the pattern correlation is 0.997, showing a very good fit between the simulated model output and the observations; again, as expected. The simulated model output and data are shown in Figures 3a and 3b. The two show similar patterns. However, a close examination reveals that the contour values are displaced by $\approx \beta_0 = 0.10$. Note that $\text{RMSE}_u = \sigma_n$, while $\text{RMSE}_s = 0.099 \simeq \beta_0 = 1$.

Case B2: $\text{RMSE}_s \neq 0$ with $\beta_1 \neq 1$

In this case β_0 may or may not be zero. We simulated this case with $\beta_0 = 0.20$ and $\beta_1 = 0.80$. The computed estimates come out to be $b_0 = 0.199$ and $b_1 = 0.80$. Again, we expect the RMSE_u to be close to σ_n. In this case the computed value matches exactly. However, since

$$\text{RMSE}_s^2 = \sum_{i=1}^{N} w_i (\hat{m}_i - y_i)^2 = \sum_{i=1}^{N} w_i [b_0 + (b_1 - 1) y_i]^2,$$

depends on both b_0 and b_1, unlike the previous two cases, it is difficult to accurately predict the computed value of RMSE_s when $\beta_1 \neq 1$. In our simulations, its computed value is found to be 0.252.

Sea surface height patterns for the simulated model output and observations are presented in Figs. 4a and 4b. Even though the two patterns are linearly related, it is difficult to discern the relationship from such a graphical presentation. The RMSE decomposition into systematic and unsystematic components, along with analysis of the regression coefficients, b_0 and b_1, is required to obtain the necessary insight.

4.2 Skewness and Kurtosis

Skewness and kurtosis measures are computed for the differences $d_i = m_i - y_i$. Whereas the kurtosis measu e is difficult to interpret, the skewness can be quite revealing. Skewness provides a measure of symmetry of the differences in the distribution, d_i. Note that this distribution is symmetrical only if $\beta_1 = 1$, as in the case

$$m_i = \beta_0 + y_i + e_i.$$

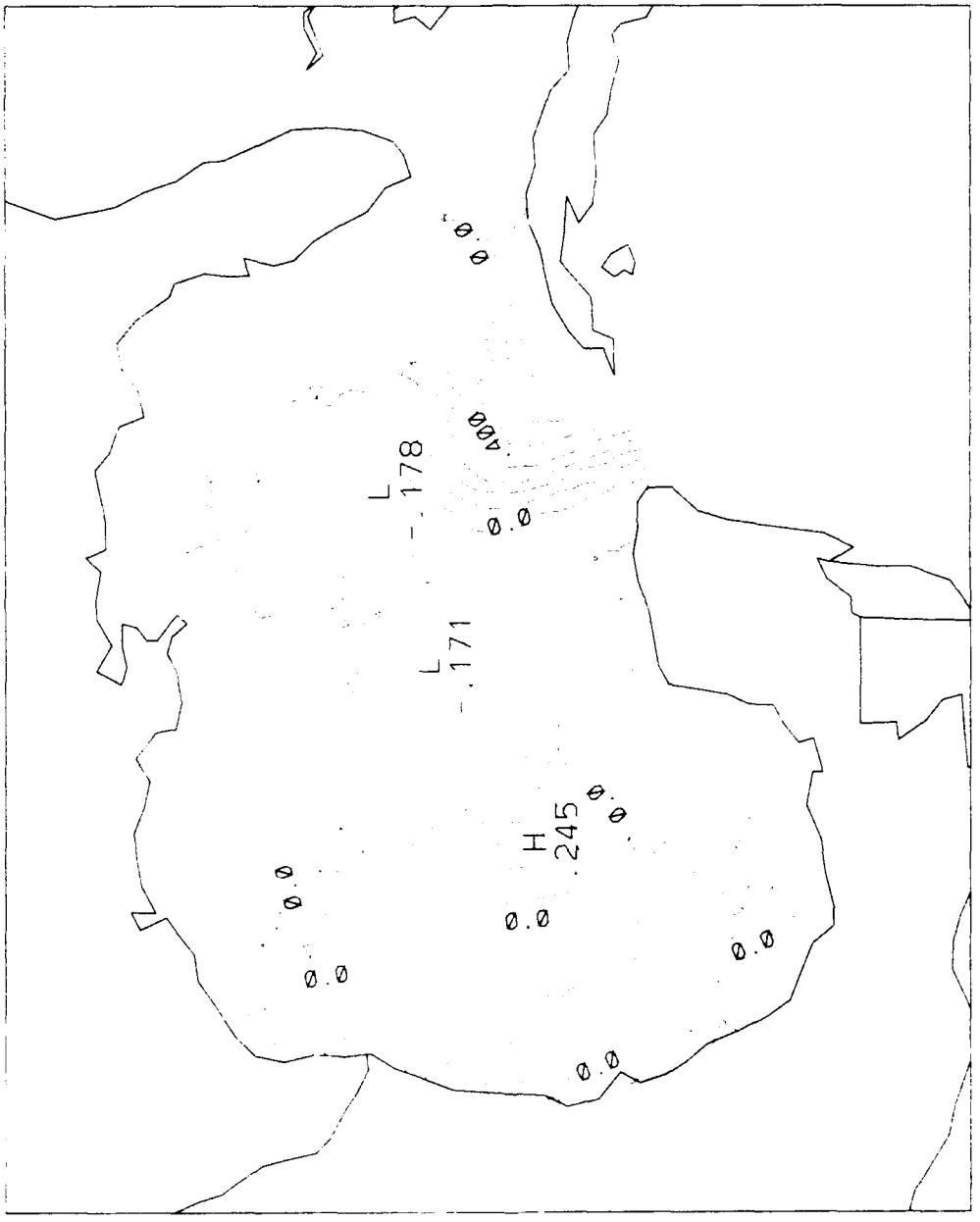


Fig. 3a: Sea Surface Height contours for simulated model output; Case B1: $\beta_0 = 0.10, \beta_1 = 1.00$

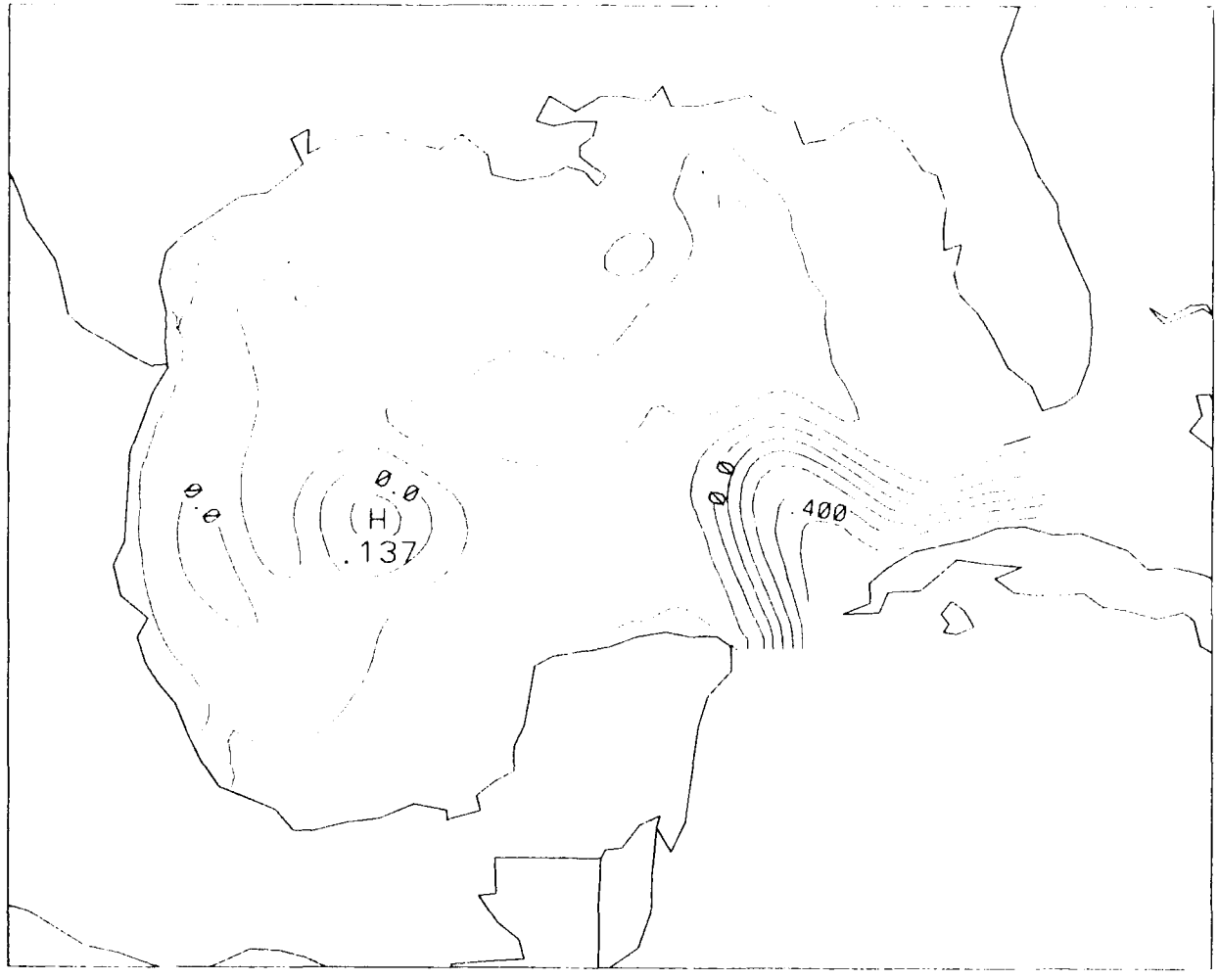


Fig. 3b: Sea Surface Height contours for simulated observations; Case B1: $\beta_0 = 0.10$, $\beta_1 = 1.00$

Fig. 4a: Sea Surface Height contours for simulated model output; Case B2: $\beta_0 = 0.20$, $\beta_1 = 0.80$

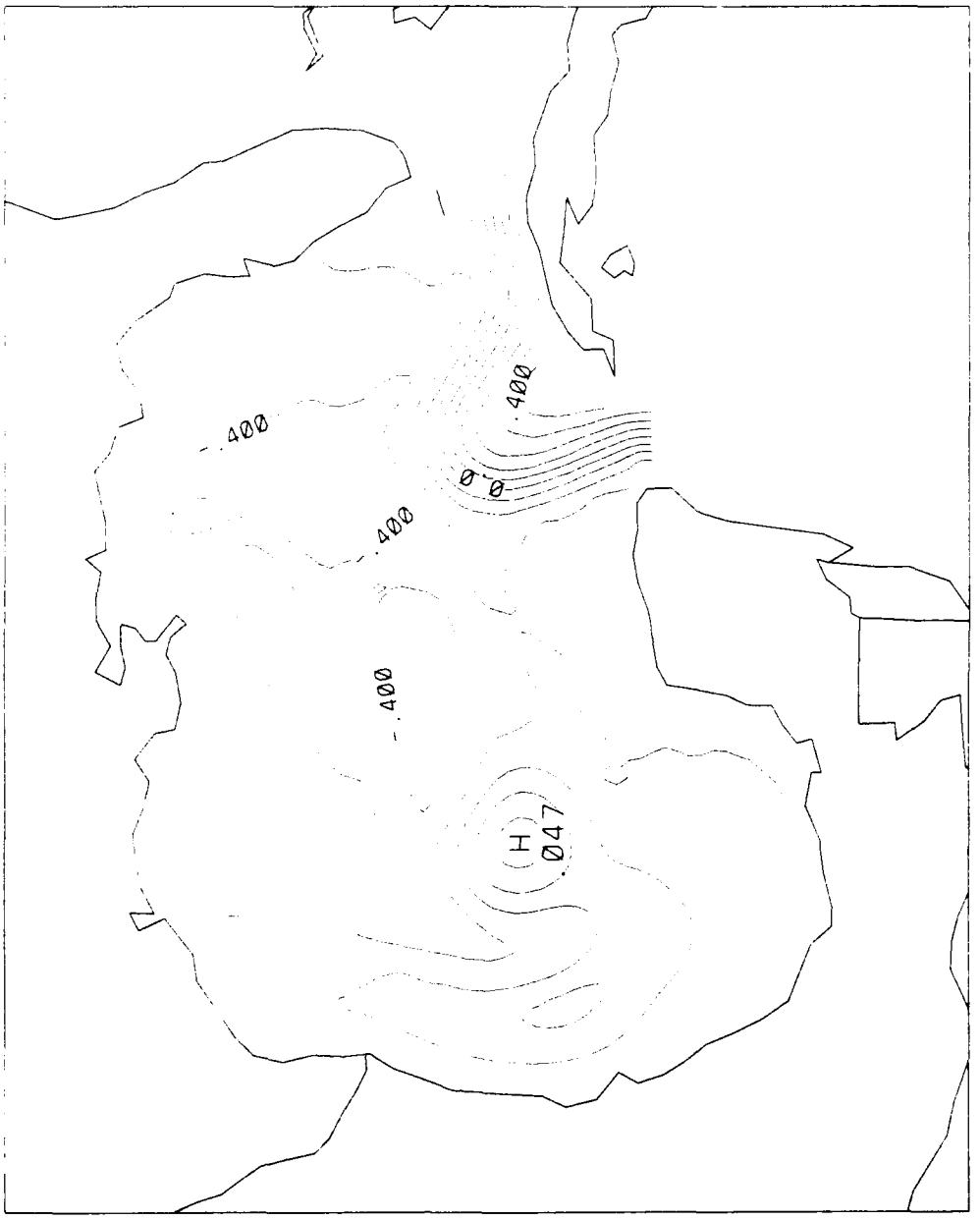


Fig. 4b: Sea Surface Height contours for simulated observations; Case B2: $\beta_0 = 0.20, \beta_1 = 0.80$

This covers the two scenarios of interest: (1) when there is no bias, i.e., $\beta_0 = 0$, and (2) when there is a fixed bias, β_0. This is evident from Figures 5a and 5b which present the histograms for the simulation cases A and B1. When $\beta 1 \neq 1$, we have

$$d_i = \beta_0 + (\beta_1 - 1)y_i + e_i.$$

Here, the distribution of d_i is centered at β_0; however, its symmetry is determined by the sign of $(\beta_1 - 1)$. The distribution is skewed to the left if the sign is negative and to the right when positive. In Case B2, the sign is negative leading to a distribution skewed to the left as confirmed by the histogram in Figure 5c. This is also apparent from the computed values of skewness in Table 1.

5. SOME FUTURE ENHANCEMENTS

The next version of VERMOD will incorporate several enhancements, including a test of hypothesis setup for the RMSE decomposition, space-time spectral decomposition, and empirical orthogonal functions (EOF's).

We will provide the capability to assign confidence bars to the linear coefficients β_0 and β_1 of RMSE decomposition. This will lead to an objective procedure of judging the systematic and unsystematic components.

We will also develop/assemble a methodology for space-time spectral decomposition of gridded fields. This will assist in determining a model's agreement with observed ocean features, component wise. Although spectral decomposition does not provide a direct model to data comparison like the RMSE measure, the component-wise verification can increase the validity of the model. Such applications require identification of well-established phenomenological features of ocean dynamics that a model incorporating proper physics and forcings must produce. Examples of such phenomenon often occur in the literature, e.g., (1) 26-day oscillations observed in the Western Indian Ocean which were reproduced by Kindle and Thompson (1989) using monthly averaged wind fields to drive the ocean, and (2) the eddy kinetic energy in certain frequency bands in the North Atlantic Gulf Stream region (Schmitz and Holland, 1982).

Empirical orthogonal function (EOF) representation of the time series of various oceanographic parameter fields is another useful procedure of reducing the dimensionality and extracting information. EOF representation has been used extensively in meteorological and oceanographic analyses. Recent examples from the literature include Hurlburt et al. (1990) and Carnes et al. (1991) that give useful applications to oceanography.

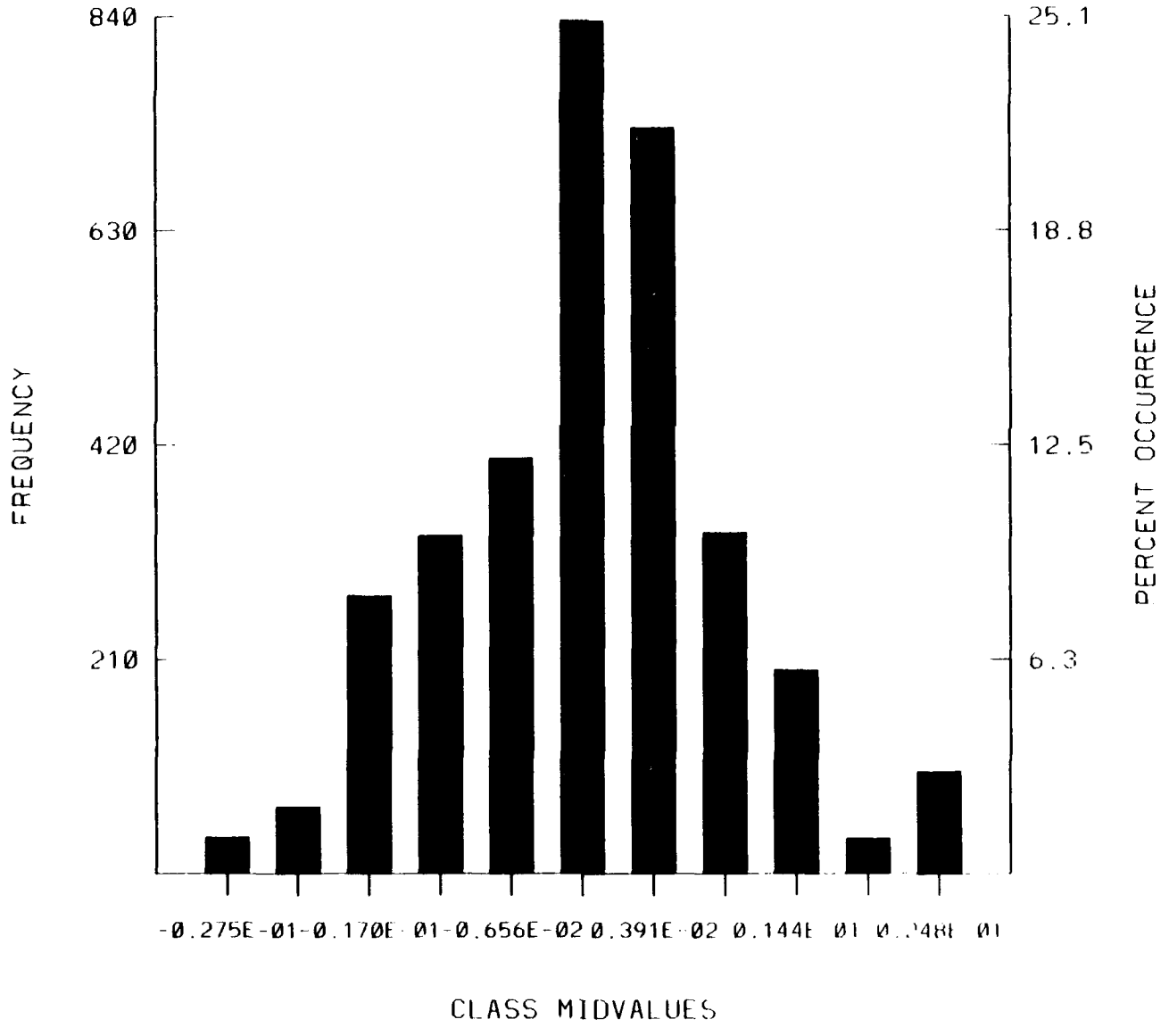


Fig. 5a: Histogram of (model output - observations) differences; Case A: $\beta_0 = 0$, $\beta_1 = 1$

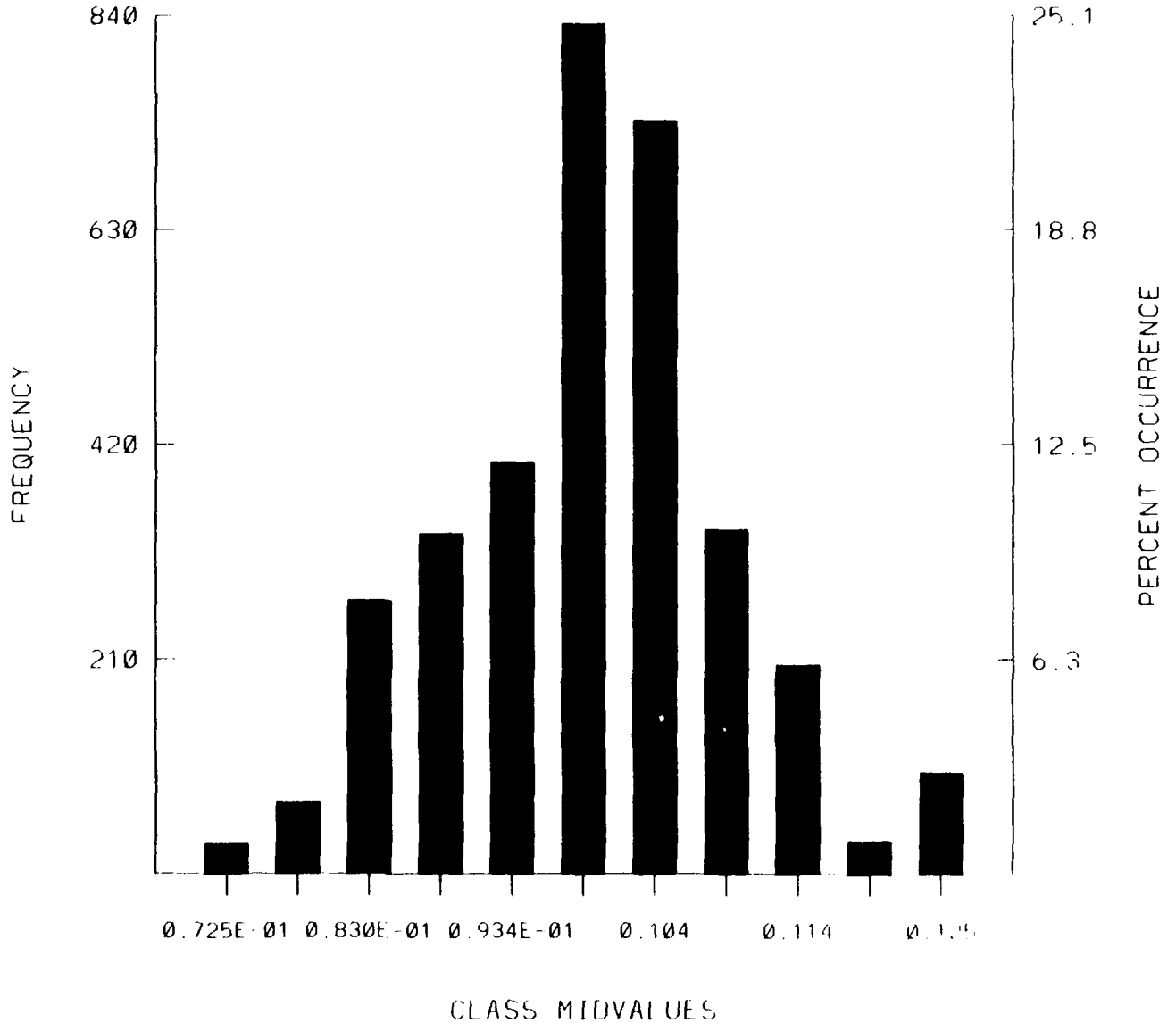


Fig. 5b: Histogram of (model output - observations) differences; Case B1: $\beta_0 = 0.10$, $\beta_1 = 1.00$

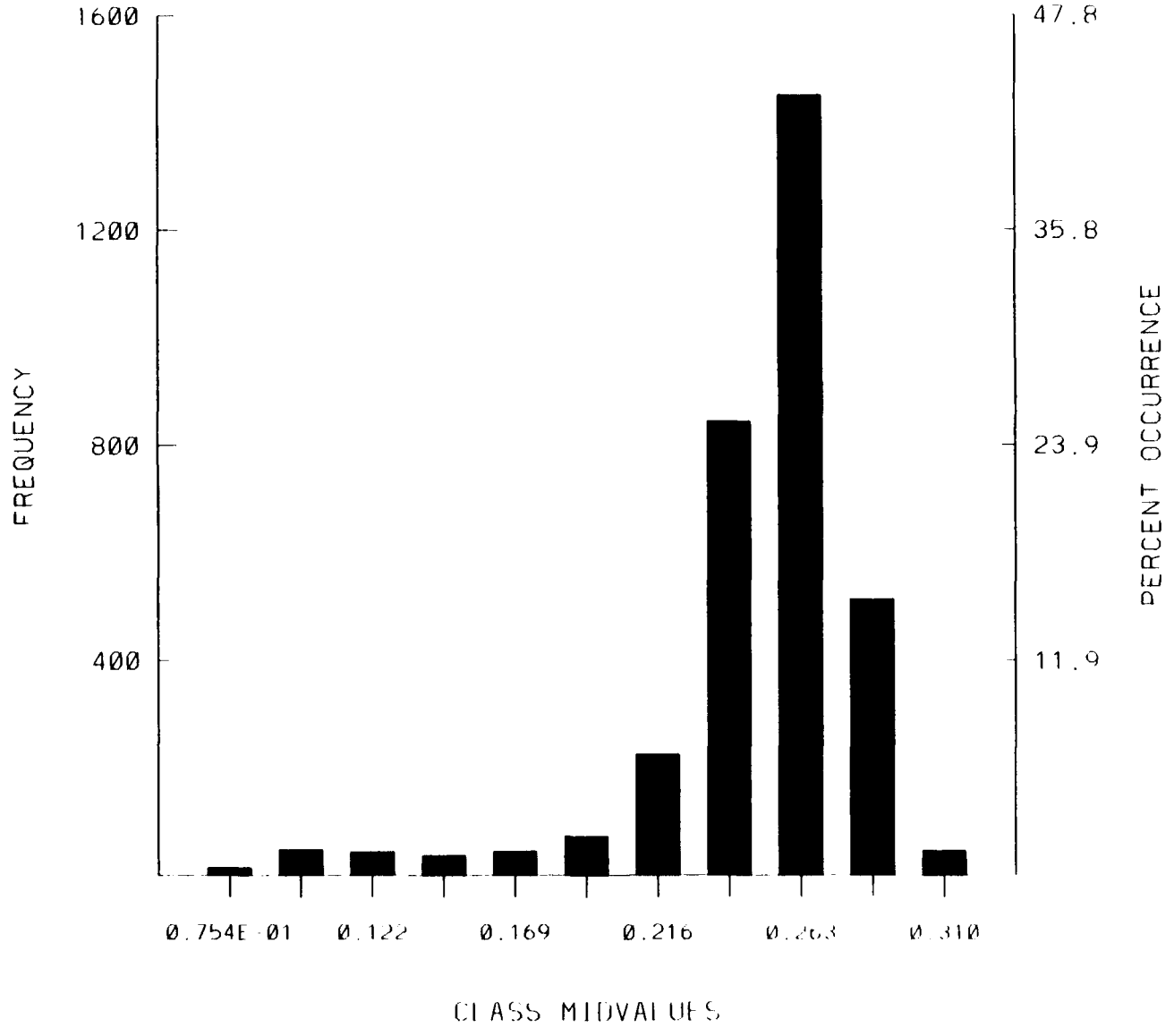


Fig. 5c: Histogram of (model output - observations) differences; Case B2: $\beta_0 = 0.20$, $\beta_1 = 0.80$

5.1 Confidence Intervals for RMSE Decomposition

Having decomposed the RMSE decomposition into systematic and unsystematic components, it is now necessary to anchor the methodology in an objective sense by specifying the $p-$values along with the estimated coefficients b_0 and b_1. Equivalently, we must provide confidence intervals (error bars). In this way, it is possible to determine the degree of conformation to one of the four classes: (1) no bias Case A, (2) simple linear bias as in Case B1, (3) a simple linear relationship in which $\beta_1 \neq 0$, and finally, (4) wherein the situation is not described by a linear relationship.

5.2 Spatial and Temporal Spectral Analysis

Spectral decomposition of a space-time series is an important analytical tool often used in analyzing statistical/dynamical behavior of the ocean. A few examples of the phenomena that can be studied using space-time spectral analysis are (William Johns, 1989 – personal communication): meander scales of the Gulf Stream, their propagation speeds and their growth rates. However, we must be cautious in applying it because a space-time series is defined in a two-dimensional space. It poses greater difficulties than if it were defined on a one-dimensional space. Space-time series data can be analyzed in three different ways: (1) analyze the data at a fixed point in space as a simple time series; (2) analyze the data at a single time point as a series in space, being cautious in this case to define the spatial axis; and (3) consider the data as a joint space-time series whose analysis provides wavenumber-frequency spectra from which we can analyze traveling and standing waves.

A spectral decomposition of a space-time series $a(x,t)$, with x as space index and t as the time index, is described following Pratt (1976) and Halliwell and Mooers (1979, 1983). Let $a(x_i,t_j)$ be discrete samples at points x_i at times t_j where x_i and t_j are equally spaced, such that $x_i = i\Delta x$ and $t_j = j\Delta t, i = 1,\ldots,M;\ j = 1,\ldots,N$. These series represent the amplitudes of the process being observed. We can then define three autocorrelation functions. Two of these corresponding to time and space, respectively, are:

$$R_{0m}(x_i,\tau_m) = \frac{1}{C(x_i,0)(N-m)}\sum_{j=0}^{N} -ma(x_i,t_j)a(x_i,t_j+\tau_m), \tag{5.1}$$

$$R_{l0}(\xi_l,t_j) = \frac{1}{C(0,t_j)(M-l)}\sum_{i=0}^{M} -la(x_i,t_j)a(x_i+\xi_l,t_j). \tag{5.2}$$

These autocorrelation functions are normalized respectively by the variances $C(x_i, 0)$ and $C(0, t_j)$. The third is the space-time autocorrelation function, defined by

$$R_m(\xi_l, \tau_m) = \frac{1}{(M-l)(N-m)} \sum_{i=0}^{M} -l \frac{1}{C(x_i, 0)} \sum_{j=0}^{N} -ma(x_i, t_j) a(x_i + \xi_l, t_j + \tau_m). \quad (5.3)$$

Frequency autospectra can be computed by Fourier transformation of the temporal autocovariance function (5.1). Assuming the spectrum is space-invariant, we can average it over its spatial coordinates. Similarly, the wavenumber autospectra can be computed by Fourier transforming the spatial autocovariance function (5.2). Averaging over the time domain can be performed using the assumption that the wavenumber spectra are time invariant.

To compute the wavenumber-frequency spectra, the space-time series $a(x_i, t_j)$ is first expanded into temporal Fourier harmonics as

$$a(x_i, t_j) = \sum_{n=1}^{N} [C_n(x_j) \cos(\sigma_n t_j) + S_n(x_i) \sin(\tau_n t_j)]. \quad (5.4)$$

The wavenumber autospectra and cross spectra of the functions $C_n(x_l)$ and $S_n(x_l)$, for the frequency σ_n and the wavenumber k_m, are given by $W_m(C_n)$, $W_m(S_n)$, $K_m(C_n, S_n)$ and $Q_m(C_N, S_N)$, where W_m is the the wavenumber autospectrum for C_m and S_m, and K_m and Q_m are the wavenumber cospectrum and quadrature for C_n and S_n. The ordinary wave number-frequency spectrum is:

$$E(\pm k_m, \sigma_n) = \frac{1}{4}[W_m(C_n) + W_{n\iota}(S_n)] \pm \frac{1}{2} Q_M(C_n, S_n), \quad (5.5)$$

and the total wavenumber-frequency autospectrum, without regard to the sign of k, is

$$T(k_m, \sigma_n) = \frac{1}{2}[W_m(C_n) + W_m(S_n)]. \quad (5.6)$$

The contribution from propagating waves is defined to be the difference between $E(+k_m, \sigma_n)$ and $E(-k_m, \sigma_n)$ given by

$$P(k_m, \sigma_n) = Q_m(C_n, S_n).$$

5.3 Empirical Orthogonal Functions

Empirical orthogonal functions (EOFs) are often used to reduce the dimensionality of the variables under consideration. For instance, consider the ocean model output field $\{a(t,x) : t = 1,\ldots,n; x = 1,\ldots,p\}$ such that t is the time index, x is the space (grid point) index, and p is large. In the Chervin-Schneider (1976) framework, we want to test the hypothesis for change in the mean sea surface temperature for the grid under consideration. Thus, the hypothesis pertains to the entire grid and not individual gridpoints. It is well known that testing point by point does not yield the proper significance at which we are testing; a few individual significance cases is not equivalent to the significance testing of the entire grid under consideration, and vice versa. In such situations we can take advantage of the unique decomposition of space-time data as follows. Let $\mathbf{a}(t) = [a(1,t),\ldots,a(p,t)]'$, where prime indicates matrix transposition, and

$$\mathbf{A} = \sum_{t=1}^{n} \mathbf{a}(t)\mathbf{a}'(t). \tag{5.7}$$

Let $\mathbf{\Phi}_k, k = 1,\ldots,p$ be the p eigenvectors corresponding to the eigenvalues λ_k. Then $a(t,x)$ can be expanded into a modal time series

$$b_k(t) = \mathbf{\Phi}_k' \, \mathbf{a}(t) = \sum_{x=1}^{p} \phi_k(x) a(x,t), \tag{5.8}$$

where $\phi_k(x)$ are the elements of the eigenvector P_k, and λ_k are the variances of $b_k(t)$. Usually, a few λ_k account for a large fraction of the variance and the remaining λ_k are negligible. Modal components, $b_k(t)$, for only nonnegligible λ_k are retained and the remaining are ignored, thus reducing the dimensionality. Preisendorfer et al. (1981) provide objective rules for selecting the number of EOFs. Relation (5.8) can be inverted to obtain the space-time decomposition of the field $\mathbf{a}(t,x)$ as:

$$a(t,x) = \sum_{k=1}^{p} b_k(t)\phi_k(x). \tag{5.9}$$

With this reduced dimensionality we can now perform all statistical tests of hypotheses that were possible with the original space-time series $a(t,x)$.

6. CONCLUDING REMARKS

The current capabilities of the Verification module VERMOD being developed in ECMOP/INO have been described. Included in VERMOD are several of the capabilities of VISMOD, the visualization module. The two modules combined together provide an effective tool to the scientific community for some aspects of model evaluation. Version 1.0 includes only rudimentary statistical measures to quantify model-data differences. The interpretation and use of these tools are illustrated by examples in which model output and observational data are simulated using computer simulations. This current version is being transitioned to the Naval Oceanographic Office.

The modules, VERMOD and VISMOD, by their nature, will be in a perpetual state of enhancement. As indicated in Section 5, several enhancement, e.g, space-time spectral decomposition and empirical orthogonal functions software, will be incorporated. The future versions will be developed as newer and more sophisticated softwares are included according to the scientists' requirements and availability/development of the advanced techniques. It is anticipated that the majority of future evaluation techniques will be developed within the modeling groups themselves, and integrated into VERMOD where applicable.

APPENDIX

We show that the cross-product term

$$\sum_{i=1}^{N} w_i(m_i - \hat{m}_i)(\hat{m}_i - y_i) = 0. \tag{A1}$$

Note that this cross-product term is a sum of two terms:

$$T_1 + T_2 = \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)\hat{m}_i - \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)y_i.$$

We will show that each of these two terms is zero. We start by obtaining the estimates b_0 and b_1 of β_0 and β_1 by minimizing the sum of the squares $\sum_{i=1}^{N}(m_i - \beta_0 - \beta_1 y_i)^2$. Then

$$\hat{m}_i = b_0 + b_1 y_i. \tag{A2}$$

The minimization yields the following normal equations:

$$\sum_{i=1}^{N} w_i(m_i - b_0 - b_1 y_i) = 0,$$

$$\sum_{i=1}^{N} w_i(m_i - b_0 - b_1 y_i)y_i = 0.$$

Using (A2) we can rewrite the normal equations as:

$$\sum_{i=1}^{N} w_i(m_i - \hat{m}_i) = 0, \tag{A3}$$

$$\sum_{i=1}^{N} w_i(m_i - \hat{m}_i)y_i = 0. \tag{A4}$$

Note that (A4) is the same as

$$T_1 = 0. \tag{A5}$$

Multiply (A3) by b_0 and (A4) by b_1 and add to obtain:

$$\begin{aligned} 0 &= \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)b_0 + \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)b_1 y_i \\ &= \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)(b_0 + b_1 y_i) \\ &= \sum_{i=1}^{N} w_i(m_i - \hat{m}_i)\hat{m}_i \\ &= T_2. \end{aligned} \tag{A5}$$

REFERENCES

Carnes, M.R., J.L. Mitchell, and P.B. deWitt, 1991: Synthetic temperature profiles derived form GEOSAT altimetry: Comparison with AXBT profiles. In press *J. Geophys. Res.*.

Chervin, R.M., and S.H. Schneider, 1976: On determining the statistical significance of climate experiments with general circulation models. *J. Atmos. Sci.*, 33, 405-412.

Computer Sciences Corporation, 1991: Database design and document for the Naval Environmental Operational Nowcasting System, Version 3.3, prepared for Naval Oceanographic and Atmospheric Research Laboratory, Monterey, CA by Computer Sciences Corporation under contract No. GS09K-90-BHD0001.

Halliwell, G.R., and C.N.K. Mooers, 1979: The space-time structure and variability of the shelf water-slope water and Gulf Stream surface temperature fronts and associated warm-core eddies, *J. Geophys. Res.*, 84, 7707-7725.

Halliwell, G.R., and C.N.K. Mooers, 1983: Meanders of the Gulf Stream down stream from Cape Hatteras 1975-1977. *J. Phys. Ocean.*, 13, 1275-1292.

Hurlburt, H.E., D.N. Fox, and E.J. Metzger, 1990: Statistical inference of weekly correlated subthermocline fields from satellite altimeter data. *J. Geophys. Res.*, 95, 11,375-11,409.

INO, 1989: Institute for Naval Oceanography Summer Colloquium on Mesoscale Ocean Science and Prediction (1989), Theme: Model Evaluation Criteria, *INO Technical Note, Proc-1.*

Kindle, J.C., and J.D. Thompson, 1989: The 26- and 50-day oscillations in the Western Indian Ocean: Model results, *J. Geophys. Res.*, 94, 4721-4736.

Leese, J.A., 1989a: ECMOP as a Concept, in Experimental Center for Mesoscale Ocean Prediction (ECMOP) *INO Report 89-2*, 59-62.

Leese, J.A., 1989b: ECMOP as a System, in Experimental Center for Mesoscale Ocean Prediction (ECMOP) *INO Report 89-2*, 63-66.

Pratt, R.W., 1976: The interpretation of space-time spectral quantities. *J. Atmos. Sci.*, 33, 1060-1066.

Preisendorfer, R.W., F.W. Zwiers and T.P. Barnett, 1981: Foundations of principal component rules. *SIO Reference Series No. 81-4*, Scripps Institution of Oceanography, La Jolla, CA, 192 pp.

Schmitz, Jr., W.J., and W.R. Holland, 1982: A preliminary comparison of selected numerical eddy-resolving general circulation experiments with observations, *J. of Marine Res.*, 40, 75-117.

Waters, M.P., R.G. Kelly, B.S. Carlson, and M.J. Roth, 1990: Trial Ocean Prediction Experiment of the Experimental Center for Mesoscale Ocean Prediction's Primitive Equation Data Assimilation Model of the Gulf of Mexico. An internal INO Report, 39pp.

Wilmot, C.J., S.G. Ackleson, R.E.Davis, J.J. Feddema, K.M. Klink, D.R. Legates, J. O'Donnell and C.M. Rowe, 1985: Statistics for the evaluation and comparison of models. *J. Geophys. Res.*, 90(C5), 8995-9005.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. Agency Use Only *(Leave blank).*	2. Report Date.	3. Report Type and Dates Covered.
	June 1991	Technical Memo -- June 1991

4. Title and Subtitle.

"VERMOD Capabilities; VERMOD 1.0 and Future"

5. Funding Numbers.

Program Element No. 0602435N

Project No. RM35G97

Task No. 801

Accession No. DN250023

6. Author(s). Harsh Anand
Michael S. Foster
Ramesh Krishnamagaru
Ranjit M. Passi

7. Performing Organization Name(s) and Address(es).

Institute for Naval Oceanography
Bldg. 1103, Rm. 233
Stennis Space Center, MS 39529-5005

8. Performing Organization Report Number.

TM-4

9. Sponsoring/Monitoring Agency Name(s) and Address(es).

Naval Oceanographic and Atmospheric Research Laboratory
Code 113
Stennis Space Center, MS 39529-5004

10. Sponsoring/Monitoring Agency Report Number.

11. Supplementary Notes.

12a. Distribution/Availability Statement.

Approved for public release; distribution is unlimited.

12b. Distribution Code.

13. Abstract *(Maximum 200 words).*

The Institute for Naval Oceanography (INO) has established the Experimental Center for Mesoscale Ocean Prediction (ECMOP) to facilitate research and development of ocean models. ECMOP aims to take advantage of existing infrastructures to provide support to Navy and academic researchers in an end-to-end model evaluation effort. In addition, if and when necessary, ECMOP performs internal research and development as an adjunct to this support. ECMOP is composed of three functional subsystems, or modules: the Verification Module (VERMOD), the Visualization Module (VISMOD), and the Data Sub-System (DASS). This technical memorandum describes the current state of VERMOD, in particular, its statistical and physical criteria for model verification, and the software structural characteristics of the module. Finally, we discuss several of the verification concepts anticipated as future enhancements of VERMOD.

14. Subject Terms.

(U) INO (U) ECMOP (U) HARVARD PE (U) DATA (U) VERMOD
(U) VISMOD (U) PEDAM (U) NOGUFS (U) DASS

15. Number of Pages. 28

16. Price Code.

17. Security Classification of Report.	18. Security Classification of This Page.	19. Security Classification of Abstract.	20. Limitation of Abstract.
Unclassified	Unclassified	Unclassified	SAR