NOAA Technical Report NOS CS 17

NOS STANDARDS FOR EVALUATING OPERATIONAL NOWCAST AND FORECAST HYDRODYNAMIC MODEL SYSTEMS

Silver Spring, Maryland
October 2003

noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE
National Ocean Service
Coast Survey Development Laboratory

Office of Coast Survey
National Ocean Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The Office of Coast Survey (CS) is the Nation's only official chartmaker. As the oldest United States scientific organization, dating from 1807, this office has a long history. Today it promotes safe navigation by managing the National Oceanic and Atmospheric Administration's (NOAA) nautical chart and oceanographic data collection and information programs.

There are four components of CS:

The Coast Survey Development Laboratory develops new and efficient techniques to accomplish Coast Survey missions and to produce new and improved products and services for the maritime community and other coastal users.

The Marine Chart Division collects marine navigational data to construct and maintain nautical charts, Coast Pilots, and related marine products for the United States.

The Hydrographic Surveys Division directs programs for ship and shore-based hydrographic survey units and conducts general hydrographic survey operations.

The Navigation Services Division is the focal point for Coast Survey customer service activities, concentrating predominantly on charting issues, fast-response hydrographic surveys and Coast Pilot updates.

NOAA Technical Report NOS CS 17

NOS STANDARDS FOR EVALUATING OPERATIONAL NOWCAST AND FORECAST HYDRODYNAMIC MODEL SYSTEMS

Kurt W. Hess
Thomas F. Gross
Richard A. Schmalz
John G. W. Kelley
Frank Aikman III
Eugene Wei
Office of Coast Survey

Mark S. Vincent
Center for Operational Oceanographic Products and Services

October 2003

noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE Donald Evans, Secretary	National Oceanic and Atmospheric Administration Conrad C . Lautenbacher, Jr., VADM USN (Ret.), Under Secretary	National Ocean Service Richard W. Spinrad, Ph.D. Assistant Administrator
Office of Coast Survey Captain Roger L. Parsons		Coast Survey Development Laboratory Bruce B. Parker

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The National Ocean Service (NOS) is developing and implementing nowcast and forecast models to support navigational and environmental applications in U.S. coastal waters. NOS aims to ensure that these models have been developed and implemented in a scientifically sound and operationally robust way. This report discusses the policies and procedures for the evaluation of nowcast/forecast models. Since the primary user group is the navigational community, and their concerns are under-keel clearance and maneuvering in port areas, the primary variables to be evaluated are water levels, currents, and water density.

This report focuses on skill assessment, although other components of evaluation include standardization, periodic review, and documentation. The components of skill assessment include: (1) the quantities relevant to navigation, (2) the time series of observed and predicted variables, (3) data processing techniques, (4) the model run scenarios, (5) the comparison statistics or quantities, (6) the target values, (7) comparison of forecast method, and (8) acceptance criteria.

The skill assessment statistics that can quantify model performance are easily calculated quantities that provide relevant information on the important categories of model behavior. The Standard Suite of statistics gives a global assessment of errors, and includes the series mean (SM) and the frequency with which errors lie within specified limits (herein termed the central frequency, CF). The Root Mean Square Error (RMSE) and Standard Deviation (SD) of error are also calculated. The frequency of times of poor performance is determined by analyzing the outliers, which are values that exceed specified limits. The Positive Outlier Frequency (POF) measures how often the nowcast/forecast is higher than the observed. The maximum duration of positive outliers (MDPO) indicates whether there are long periods when the model does poorly. The Negative Outlier Frequency (NOF) and Maximum Duration of Negative Outliers (MDNO) are analogous.

The three requirements for evaluating tidal water levels and tidal currents at each location where data are available are (1) comparison of tidal harmonic constants, (2) computation of the Standard Suite of statistics, and (3) comparison of forecast methods (e.g., astronomical tide only, tide plus a persisted offset, and model-based prediction). For water densities and water levels in areas without significant tidal variations (e.g., Great Lakes), the Standard Suite of statistics and comparison of forecast method are required. For the case of locations where no historical data exist, the model-generated data (time series or a field) are analyzed and a professional judgment is made as to its realism and the extent to which it captures actual features.

For a nowcast or forecast at a particular station to be approved for release to the public, the statistics related to model performance at that station must (a) meet or exceed all target frequencies or durations, or (b) meet or exceed most of the target frequencies or durations and be granted a waiver by NOS' Technical Review Team.

Key words: Nowcast, Forecast, Skill Assessment, Hydrodynamic Models, Statistics, Tides, Water levels, Currents, Water Density

1. INTRODUCTION

In order to meet its operational oceanographic mission responsibilities, the National Ocean Service (NOS) is developing and implementing nowcast and forecast models to support Physical Oceanographic Real Time Systems (PORTS) and other navigational and environmental applications in U.S. coastal waters. These models are designed to enhance the navigational guidance supplied by NOS' real-time observations by providing information regarding both the present (nowcast) and future (forecast) oceanographic conditions at many locations within an estuary, bay, lake, or coastal ocean. These models will be developed by the Coast Survey Development Laboratory (CSDL) and by other groups within and outside of NOS.

NOS must ensure that these models have been developed and implemented in a scientifically sound and operationally robust way; that the model's shortcomings are understood; that the products are clear, understandable, and useful; and that all products and procedures are authoritative in the face of potential legal challenges. It is imperative that the nowcast and forecast systems are developed consistent with user needs and with the operational environment in which they will be run. All models (including statistical models) that produce nowcasts and forecasts in support of safe navigation, whether developed within or outside NOS, will be developed and implemented in adherence to the procedures contained in this document.

This report discusses the specific policies and procedures for the evaluation of NOS' nowcast/forecast models for navigation. The evaluation focuses on the performance of the model system during the development phases, and the accuracy of the system and its products during the operational phase. Since the suitability of the predictions is ultimately determined by the user (the navigational community), the primary variable discussed here are water levels, currents, and water density.

The main components of model evaluation are standardization, periodic review, skill assessment, product quality control, and documentation.

Standardization - Standardization means that NOS model output and products, documentation, skill assessment, and review will be uniform. Although standardization is a goal, we must recognize that since each region and forecast system will be custom-tailored to some extent, we cannot expect total uniformity. However, the System Design and Implementation Team, or SDIT (Vincent et al., 2003), has the responsibility of being aware of the features of all other nowcast/forecast systems and must justify any significant departures.

Periodic Review - During the model development, the SDIT should schedule and lead periodic (at least twice a year) technical meetings to present progress to the Technical Review Committee (Vincent et al., 2003). As part of the process, reviewers are to objectively assess progress and offer constructive written comments; the SDIT is required to respond to written comments.

Skill Assessment - Skill assessment is an objective measurement of how well the model nowcast or forecast does when compared to observations. The approach here is to measure the performance of the model in (1) simulating astronomical tidal variability, (2) simulating total (tide and non-tidal effects) variability in both the model development stage and the operational environment, and (3) giving a more accurate forecast than the tide tables and/or persistence. Skill assessment is discussed in depth in Section 3. If data are minimal or lacking, other procedures are needed (Section 8).

Product Quality Control - Once the system begins operation, its products must be continually assessed for accuracy. To assist the SDIT in judging the probable accuracy of the forecast, the forecast system will have standard graphical display and data exchange products, and save important information in a system status file for the Continuous Operational Realtime Monitoring System, or CORMS (Gill et al., 1997).

Documentation - Written documentation of the model and model system is necessary for the communication of model structure and performance. Documentation includes clear explanation of equations, boundary conditions, inputs, and constants, as well as sample (graphical) outputs and skill statistics. It will include copies of internally documented computer code for all processes. It will also cover real-time and forecast input data streams and the telecommunications that provide them.

This document is a revision of an earlier report (NOS, 1999). The major difference is that this document focuses on system evaluation, and the sections of the earlier document that describe the procedures for implementation of nowcast/forecast systems now appear in a separate report (Vincent et al., 2003). Also, the new approach to implementation (Vincent et al., 2003) prescribes that when model systems are created within NOS, there will be a Technical Review Team (TRT), consisting of members of CSDL and the Center for Operational Oceanographic Products and Services (CO-OPS), to guide the process. The actual model development and skill testing will be carried out by the System Design and Implementation Team (SDIT).

The following sections of this report focuses on skill assessment and the development of acceptance criteria (Section 2). Following that will be a discussion of the required performance measures for the evaluation for water levels (Section 3) and for currents (Section 4) when comparing predicted and observational data. Requirements for salinity and water temperature are discussed in Section 5, and evaluation in non-tidal regions is covered in Section 6. If data are minimal or lacking altogether, or the product is an entire spatial field, other procedures are needed; these are discussed in Section 7. Topics for future research are discussed in Section 8. Section 9 presents a summary of the skill assessment criteria.

2. SKILL ASSESSMENT

For NOS' purposes, skill assessment is the primarily objective judgment of a model's performance (i.e., its ability to reproduce observed variability and to predict future variability) using both objective standards and measures against other prediction methods. Some skill assessment statistics were designed to show how the prediction method could be improved. The standards described here are considered to be 'user based', that is, established to show how well a model meets user needs (for the navigation community, they are under-keel clearance and maneuvering in port areas), and are not influenced by the model's capabilities.

Skill assessment applies both to the model and to the entire nowcast/forecast system, since the availability, quality, and timeliness of input data (from observations and from other models) affects the quality of the nowcast/forecast. The methods discussed are to be applied to a model (either a numerical circulation model or statistical prediction scheme) that has been previously developed; therefore, basic questions about methodology (in the use of numerical models) about mass conservation, etc., will have been settled.

2.1. Overview

The general approach to creating the skill assessment procedures described here was as follows.

- The first step was to identify the specific variables that are required by the user, i.e., the coastal navigational community.

- The next step was to determine how models could be run to produce the relevant variables. This included (a) defining the important time series, (b) selecting analysis procedures, and (c) determining model run scenarios.

- The third step was, depending on the scenario, to select measures of performance. This included (a) defining statistics that quantify model performance, (b) selecting the target values of the statistics that define model success, and (c) assessing of the systems' ability to give a more accurate forecast than other methods (e.g., the tide tables and/or persistence).

2.2. Relevant Variables

In terms of importance to navigation in U.S. coastal waters and ports, the primary variables are:

- the <u>magnitude of the water level</u> at all times and locations for under-keel clearance,
- the <u>times and amplitudes of high and low water</u> for under-keel clearance,
- the <u>speed and direction of the currents</u> at all times and locations, but especially at channel junctions, for maneuvering (the direction is computed only for current speeds above ½ knot),
- the <u>times, amplitudes, and directions of the maximum flood and ebb currents</u> (for maneuvering),

- the start times and end times of slack water (slack water is defined as by a current speed of less than ½ knot) before flood and ebb at all locations, but especially at channel junctions for planning turns in confined areas, and
- water density, since it contributes to buoyancy, for under-keel clearance and cargo loading capacity. Density is usually defined in terms of salinity and temperature.

Skill assessment will be focused on the model system's accuracy in simulating the above variables.

2.3. Time Series Data

The predicted and observed data must be processed to extract the variables important to navigation. In particular, the data must be organized into sets showing consecutive values of a variable. There are three types of data sets: Group 1, a time series of values at uniform time intervals; Group 2, a set of values representing the consecutive occurrences of an event (such as high water or slack water); and Group 3, a set of values representing a forecast valid at a given projection time (Table 1). See Sec. 2.5 for a discussion of time series length and intervals.

Table 1. Data series Groups and the variables in each. Note that upper case letters indicates a prediction series (e.g., H), and lower case (e.g., h) indicates a reference series (observation or astronomical prediction). Slack water is defined as a current speed less than ½ knot. The direction is computed only for current speeds above ½ knot.

Group	Variable	Symbol
Group 1 (Time Series)	Water level	H, h
	Current speed	U, u
	Current direction	D,d
	Salinity	S, s
	Water temperature	T,t
Group 2 (Values at a Tidal Stage)	Amplitude of high water	AHW,ahw
	Amplitude of low water	ALW,ahw
	Time of high water	THW,thw
	Time of low water	TLW,tlw
	Amplitude of maximum flood current	AFC,afc
	Amplitude of maximum ebb current	AEC,aec
	Time of maximum flood current	TFC,tfc
	Time of maximum ebb current	TEC,tec
	Direction of current at maximum flood	DFC,dfc
	Direction of current at maximum ebb	DEC,dec
	Time of start of current slack before flood	TSF,tsf
	Time of end of current slack before flood	TEF, tef
	Time of start of current slack before ebb	TSE, tse
	Time of end of current slack before ebb	TEE, tee
Group 3 (Values from a Forecast)	Water level at forecast projection time of nn hrs	Hnn, hnn
	Current speed at forecast projection time of nn hrs	Unn, unn
	Current direction at forecast projection time of nn hrs	Dnn, dnn
	Salinity at forecast projection time of nn hrs	Snn, snn
	Water temperature at forecast projection time of nn hrs	Tnn, tnn

For Group 1, the data can be either (1) a time series of values (such as observations at a location) or (2) a series of values from concatenated segments (such as a set of 24-hr nowcasts or forecasts starting at one time in the day). For currents, the time series will need to have speed and direction; the direction error is computed only for current speeds above ½ knot.

For Group 2, values are created from a Group 1 series by selecting a sub-set of values such as the time and amplitude of high water or the time of the start and end of slack water (defined as having a current speed less than ½ knot).

For Group 3, data consist of the values of the forecast variable that are valid at a fixed interval into the forecast (e.g., 0 hr, 6 hr, 12 hr, etc). The comparison series is then the observed variable at the time the forecast is valid. If there are, for example, two forecasts per day, then there will be two 6-hr projection values, separated by 12 hours in time.

Note that all model and observational data units are to conform to the International Standard, although occasionally English units (e.g., feet, knots) may appear for reference.

2.4. Harmonic Analysis and Other Data Analysis Techniques

The following are techniques that are useful in analyzing the 6-min time series and the series created from forecasts.

Harmonic Analysis - Tidal harmonic constants (amplitudes and phases or epochs) for the 37 NOS constituents are analyzed for a tide-only model simulation and compared with the accepted values available from CO-OPS. The 37 constituents used by NOS are the M_2, S_2, N_2, K_1, M_4, O_1, M_6, MK_3, S_4, MN_4, ν_2, S_6, μ_2, 2N, OO, λ_2, S_1, M_1, J_1, Mm, Ssa, Sa, Msf, Mf, ρ_1, Q_1, T_2, R_2, 2Q, P_1, 2SM, M_3, L_2, $2MK_3$, K_2, M_8, and MS_4. A description of the constituents can be found in Schureman (1958). A least-squares program (Zervas, 1999) should be used to analyze the modeled 180- to 365-day long time series of both water levels and currents (1 hr intervals are suitable). Techniques are also available to analyzing 15- and 29-day series (Zervas, 1999).

Extracting Extrema - The maximum or minimum values (i.e., the extrema) can be extracted by searching for largest or smallest values within a given time period in a series. For tidal variations, the time period is 25 hr. For lakes, the time interval can vary. Filtering or Singular Value Decomposition (see below) can be used to create a smooth time series. See below for extracting extrema from forecasts.

Gap-filling - The extraction of extrema cannot be accomplished in a time series with gaps. If a gap is not more than 3 hrs, synthetic values can be created by fitting a sine curve or a cubic curve through the data using either a least squares or singular value decomposition approach.

Filtering - Filtering of values in a time series is necessary to select accurately the extrema (i.e., maximum or minimum) values and times. A Fourier filter is preferred since it computes amplitudes of the components of the signal at various frequencies and reduces the amplitudes at selected frequencies. Simple smoothing is to be avoided since it reduces extrema amplitudes.

Singular Value Decomposition (SVD) - SVD fits a low order polynomial through a large number of data values, and is used by NOS to extract the time and amplitude of high and low water levels. It may be used instead of filtering.

Finding Principal Current Direction (PCD) - The PCD is required for computing the harmonic constants from a tidal current series, and is computed as follows. For an eastward current, u, and northward current, v, the PCD [Preisendorfer (1988), Eqn. 2.9] is

$$PCD = \frac{1}{2}\arctan\left(\frac{2\sum_{i=1}^{N}(u_i - \bar{u})(v_i - \bar{v})}{\sum_{i=1}^{N}(u_i - \bar{u})^2 - \sum_{i=1}^{N}(v_i - \bar{v})^2}\right) + m\frac{\pi}{2} \quad (1)$$

where m is either 0 or 1, whichever gives a PCD that maximizes the variance, s^2 (Preisendorfer,1988; Eqn. 2.6), defined as

$$s^2 = \cos^2(PCD)\sum_{n=1}^{N}(u_i - \bar{u})^2 + 2\sin(PCD)\cos(PCD)\sum_{n=1}^{N}(u_i - \bar{u})(v_i - \bar{v}) + \sin^2(PCD)\sum_{n=1}^{N}(v_i - \bar{v})^2 \quad (2)$$

PCD is counterclockwise from east and may represent either the flood or ebb direction. For example, for the current shown in Figure 1, the PCD is -19° (counterclockwise from east), and for the current shown in Figure 2, the PCD is 26° (counterclockwise from east).

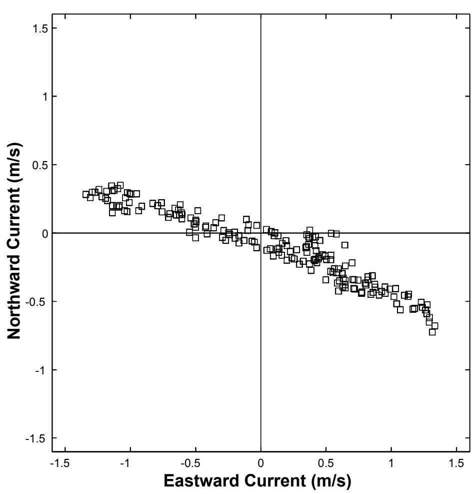

Figure 1. Unidirectional currents at Bolivar Roads, Texas, at 6-min intervals for April 12, 2003. R = 0.01.

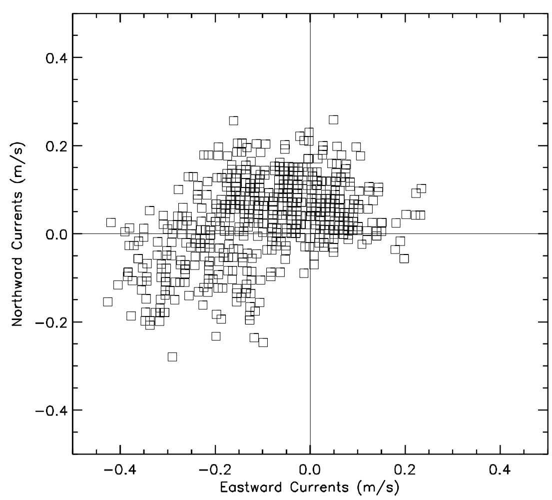

Figure 2. Rotary currents at Queens Gate, California, at 6-min intervals for April 23, 2003. R = 0.35.

Currents may be analyzed along only one direction (the PCD) unless there is significant cross directional flow. Significant cross directional flow occurs when *R*, the ratio of the standard deviations (where the standard deviation is *s* in Eqn. 2), is greater than 0.25.

$$R = \frac{s^2(PCD + \frac{\pi}{2})}{s^2(PCD)} \tag{3}$$

For example, for the current shown in Figure 1, the ratio is 0.01 and for Figure 2 the ratio is 0.35.

Construct a Time Series from Nowcasts - A single time series can be constructed from a set of nowcasts by simply appending them. If there are multiple nowcasts during one day, the skill analysis can be made on either an average or each separate series.

Construct a Time Series from Forecasts - A set of time series for Group 3 data (e.g., Hnn) can be constructed from forecasts as follows. Suppose there are a fixed number of forecasts per day, with forecast fields saved at a fixed time interval. Consider a forecast variable, F_i^n, representing for example water level, valid at the selected i^{th} projection time of the n^{th} forecast in the day. A series can be made using all the *Fs* at the i^{th} projection from each forecast for all days, namely $F_i^1, F_i^2 \ldots$, having the valid times $T_i^1, T_i^2, \ldots$ These values can be compared to the observations at the same times. Other series can be made for different projection times.

Extracting Extrema from Forecasts – Extracting the extrema (e.g., for AHW) from a limited-duration forecast can be accomplished by neglecting values at all times before the first crossing (a crossing occurs when the sign of the variable changes), and the values at times later than the last crossing in the forecast. This eliminates the possibility of obtaining an erroneous value for the first and last extrema. It will, however, reduce the number of extrema found in any forecast; for a 24-hour forecast, the MDPO and MDNO based on L = 25 hours (see Sec. 2.7) will always be met. Therefore, MDPO and MDNO for a forecast need not meet specific critera.

2.5. Model Run Scenarios

There are five scenarios (Table 2) under which the model is run to produce the data for skill assessment, and they are discussed in the order they would occur during model development.

The scenarios begin with the (1) Astronomical Tide Only simulations because in most coastal regions tidal variations are generally dominant, they may account for a significant part of the error, and because there are extensive data available for validation (but see Section 6 on evaluation in non-tidal regions). Modeled time series can be harmonically analyzed to produce constituent amplitudes and phases for comparison with accepted values. These values provide information on the model's behavior in frequency space and can also illuminate the role of friction and non-linear processes. The (2) Hindcast is a long simulation using the best available gap-filled data for observed boundary water levels, winds, and river flows. The (3) Test Forecast is made in a hindcast mode, but using the best available gap-filled data for forecast boundary water levels, winds, and river flows. The (4) Semioperational Nowcast and (5) Semioperational

Forecast are made in an operational environment (i.e., running daily with real-time input) and so they will occasionally encounter missing observations and forecasts; the system must be able to handle these conditions without significant loss of accuracy.

Table 2. Model Run Scenarios. Scenarios needed to produce skill assessment variables. Model forcing includes ocean boundary water levels and water density variations, wind stresses at the air-water interface, and river flows.

Scenario	Explanation
1. Astronomical Tide Only	In this scenario, the model is forced with only harmonically-predicted astronomical tides for the ocean boundary water levels. A run typically covers several months.
2. Hindcast	In this scenario, model forcing is based on historical, gap-filled observational data. A run typically covers several months.
3. Test Forecast	(Optional) In this scenario, the model forcing is based on archived, gap-filled forecasts. A run typically covers a few days to a few months.
4. Semioperational Nowcast	In this scenario, the model forcing is based on recent observed values, even though some data could be missing. A run typically covers a few hours to days.
5. Semioperational Forecast	In this scenario, the model forcing is based on recent forecast values from other models, even though some data could be missing. Initial conditions are generated by observed data or the output from a nowcast. A run typically covers 1 or 2 days.

The length of time each scenario is to be run is, ideally, 365 days in order to capture all expected seasonal conditions. However, some scenarios can be run concurrently to reduce the time required for implementation. When significant data are missing or other circumstances arise, the Technical Review Team may reduce the 365-day requirement. Normally, data at 6 min intervals is required.

2.6. Skill Assessment Statistics

Each SDIT is responsible for generating the set of statistical values that will be used for model evaluation. Although no single set of statistics can quantify model performance perfectly, we have chosen several, easily-calculated quantities that provide relevant information on the important categories of model behavior. A summary of relevant terms is shown in Table 3.

For a global assessment of errors, both the series mean (SM) and the frequency with which errors lie within specified limits (herein termed the central frequency, CF) are used. The SM will indicate how well the model reproduces the observed mean and the CF indicates how often the error is within acceptable limits. The Root Mean Square Error (RMSE) and Standard Deviation (SD) are to be calculated, but have limited use since we do not expect errors to be normally distributed and CF is easier to explain to users lacking a technical background. The CF concept has been used previously in NOS for data quality assurance standards (Williams et al., 1989).

The frequency of times of poor performance is determined by analyzing the outliers, which are values that exceed specified limits. The Positive Outlier Frequency (POF) measures how often the nowcast/forecast is higher than the observed. The maximum duration of positive outliers (MDPO) indicates whether there are long periods when the model overpredicts. The Negative Outlier Frequency (NOF) measures how often the nowcast/forecast is lower than the observed. The maximum duration of negative outliers (MDNO) indicates whether there are long periods when the model underpredicts. The MDPO and MDNO will be computed with data without gaps. For water levels, the 'worst case', from a model-based nowcast/forecast viewpoint, is when actual water level turns out to be low but the model erroneously predicted much higher water levels and the user would have been better off using the astronomical tide water level prediction. This is called the Worst Case Outlier Frequency (WOF).

Table 3. Skill Assessment Statistics . The variables and statistics used in the skill assessment are explained below.

Variable	Explanation
Error	The error is defined as the predicted value, p, minus the reference (observed or astronomical tide value, r : $e_i = p_i - r_i$.
SM	Series Mean. The mean value of a series y. Calculated as $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$.
RMSE	Root Mean Square Error. Calculated as $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_i^2}$.
SD	Standard Deviation. Calculated as $SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (e_i - \bar{e})^2}$
CF(X)	Central Frequency. Fraction (percentage) of errors that lie within the limits $\pm$X.
POF(X)	Positive Outlier Frequency. Fraction (percentage) of errors that are greater than X.
NOF(X)	Negative Outlier Frequency. Fraction (percentage) of errors that are less than -X.
MDPO(X)	Maximum Duration of Positive Outliers. A positive outlier event is two or more consecutive occurrences of an error greater than X. MDPO is the length of time (based on the number of consecutive occurrences) of the longest event.
MDNO(X)	Maximum Duration of Negative Outliers. A negative outlier event is two or more consecutive occurrences of an error less than -X. MDNO is the length of time (based on the number of consecutive occurrences) of the longest event.
WOF(X)	Worst Case Outlier Frequency. Fraction (percentage) of errors that, given an error of magnitude exceeding X, either (1) the simulated value of water level is greater than the astronomical tide and the observed value is less than the astronomical tide, or (2) the simulated value of water level is less than the astronomical tide and the observed value is greater than the astronomical tide.

These statistics fall within the two categories described by Dingman and Bedford (1986), namely traditional and non-parametric statistics. For example, the RMSE of the error between two time series is a traditional statistic, while the RMSE of a set of times or amplitudes of high water is non-parametric. Dingman and Bedford (1986) conclude that non-parametric statistics are better able to assess the ability of a model to simulate extreme events.

2.7. Target Frequencies and Durations

Most of the statistics described above have an associated target frequency of occurrence. For example,

$$S(X) \leq P \tag{4}$$

where S is the statistic, X is the acceptable error magnitude (defined by the user), and P is the target frequency (percentage). The targets for the distribution of errors have the general form:

$$CF(X_1) \geq 90\%, \quad POF(X_2) \leq 1\%, \quad NOF(X_2) \leq 1\%. \tag{5}$$

For water levels

$$WOF(X_2) \leq 0.5\%. \tag{6}$$

To insure that the positive and negative outliers track relatively large errors, we take $X_2 = 2X_1$.

The set of conditions (5) means that 90% of the errors will be within $\pm X$, only 1% of the errors will exceed X, and 1% of the errors will be less than X. The limit of 1% of the time is equivalent to about 87 hours (about 3½ days) per year. Note that for a normal (Gaussian) distribution, the requirement that $CF(X) \leq 90\%$ implies that SD $=0.608X$ and that $POF(2X)=0.05\%$. However, errors cannot be expected to be Gaussian (Figure 3).

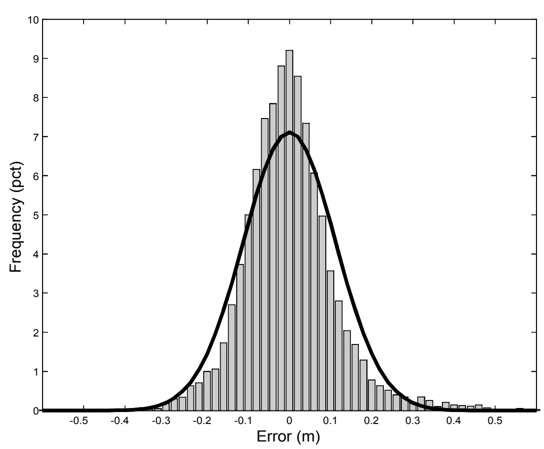

Figure 3. Water level errors at Baltimore for 1996. The solid line is a Gaussian distribution.

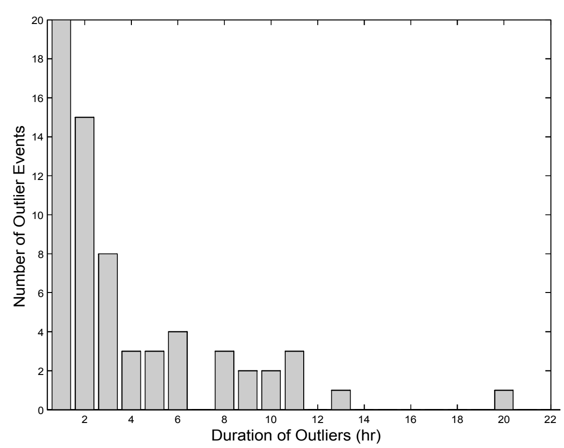

Figure 4. Number and duration of outlier events. The number of events of duration less that 1 hr has been truncated at 20.

Other statistics are expressed as limits on the duration of errors, such as

$$S(X) \leq L \tag{7}$$

where L is the time limit or maximum allowable duration. For example, the target time duration (defined as the length of time bracketing consecutive occurrences of an outlier) that applies to data of Group 1, 2, or 3 is:

$$MDPO(X_2) \leq L \quad and \quad MDNO(X_2) \leq L, \tag{8}$$

where L is the target time limit in hours. Figure 4 shows a typical distribution of outliers.

2.8. The Standard Suite of Statistics

The above two sets of target frequencies and durations, plus SM, RMSE, SD, and for water levels WOF, are required for the assessment of nearly all variables, and are collectively called the Standard Suite of statistics (Table 4).

Table 4. Standard Suite of statistics and Standard Criteria for Skill Assessment. The table lists the required skill parameters in terms of acceptable error (X), target frequencies (expressed as a percentage), and limiting durations (L). Definitions of skill parameters are as described in Table 3. [1] WOF applies only to water levels.

Variable	SM	RMSE	SD	NOF(2X)	CF(X)	POF(2X)	MDPO(2X)	MDNO(2X)	WOF(2X)[1]
Criterion	none	none	none	≤ 1%	≥ 90%	≤ 1%	≤ L	≤ L	≤ ½%

2.9. Comparison of Forecast Method

For this comparison, the Standard Suite of statistics for the appropriate variables must be evaluated for the model-based forecast and for at least one other forecast method. The requirement is that the model-based forecast should be better than the forecast based on other methods. If the model forecast is not an improvement over the others, the model should not be implemented.

For water levels and currents, these methods are (1) the astronomical tidal prediction alone, and (2) the astronomical tidal prediction to which a persisted non-tidal component has been added. The astronomical tide forecast is a prediction at each station using accepted harmonic constants for that station. A tide plus persistence water level forecast is constructed by adding an offset value, based on an observed offset at that station during some time period before the forecast is made, to the tide prediction at each station. For currents, the offset may be a mean current. In equation form, the requirements for the comparison of forecast method for tidal water levels and currents are

$$
\begin{aligned}
CF_{astronomical} &\leq CF_{persistence} \leq CF_{model} \\
POF_{astronomical} &\geq POF_{persistence} \geq POF_{model} \\
NOF_{astronomical} &\geq NOF_{persistence} \geq NOF_{model} \\
MDNO_{astronomical} &\geq MDNO_{persistence} \geq MDNO_{model} \\
MDPO_{astronomical} &\geq MDPO_{persistence} \geq MDPO_{model}
\end{aligned} \tag{9}
$$

and, for water levels only,

$$WOF_{astronomical} \geq WOF_{persistence} \geq WOF_{model}. \tag{10}$$

For salinity, temperature, and non-tidal water levels, a comparison forecast can possibly be made by using another method. For these variables, the criteria for forecasts is

$$
\begin{aligned}
CF_{other} &\leq CF_{model} \\
POF_{other} &\geq POF_{model} \\
NOF_{other} &\geq NOF_{model} \\
MDNO_{other} &\geq MDNO_{model} \\
MDPO_{other} &\geq MDPO_{model}
\end{aligned} \tag{11}
$$

2.10. Acceptance Criteria

As discussed above, the approach here is to measure the performance of the model in (1) simulating astronomical tidal variability (for tidal water levels and currents), (2) simulating total variability in both the model development stage and the operational environment, and (3) giving a more accurate forecast than another method can (e.g., the tide tables and/or persistence).

For a nowcast or forecast at a particular station to be approved for release to the public, the statistics related to model performance at that station must (a) meet or exceed *all* target frequencies or durations, or (b) meet or exceed *most* of the target frequencies or durations and be granted a *waiver* by the Technical Review Team. A waiver may be needed to allow for the wide variety of coastal areas and their dynamics, for changes in the priorities of users, and because a forecast is not likely to be as accurate as a nowcast. However, the basis for any waiver will be judicially considered. Legitimate reasons may include the fact that a time series of required length is unavailable or that a numerical criterion is missed by only a small amount. The Technical Review Team may approve for dissemination a limited forecast (i.e., only a few forecast projections) or full (24 hour) forecasts for a limited number of locations. The Technical Review Team may alter the accuracy requirements based on present day modeling capabilities.

The numerical values appearing in the target frequencies and durations were selected on the basis of the estimated utility of the specific nowcasts/forecasts to users; it is expected that as nowcasts/forecasts become more widely disseminated, some values will change. Statistics with no specific acceptance requirement are still necessary for scientific model evaluation. To assist the decision-making process, all statistics generated from the data will be presented in a precision at least one place beyond that of the criterion or comparison value. The Technical Review Team will have final say on targets, criteria, and model acceptance.

3. EVALUATION OF PREDICTED WATER LEVELS IN TIDAL REGIONS

The following is a discussion of the statistics and acceptance criteria that are necessary for the evaluation of predicted water levels in regions with significant tidal variations. They are to be generated, analyzed, and documented by the SDIT with a precision sufficient for an accurate comparison.

The three requirements for evaluating water levels (at locations where data are available) are:

- comparison of tidal harmonic constants,
- computation of the Standard Suite of statistics, and
- comparison of forecast methods

3.1. Comparison of Tidal Harmonic Constants

For this comparison, tidal harmonic constants (37 amplitudes and phases) are analyzed from a tide-only model simulation and compared with NOS accepted values. Table 5 is a template of the variables required. Harmonic constants are available from CO-OPS. A least-squares program (Zervas, 1999) should be used to analyze the modeled 365-day time series (1 hr intervals are suitable). This comparison is for model checking only, and there are no target values. Typical results are shown in Appendix A (Table A.1).

Table 5. Template for comparison of tidal constituent amplitudes and epochs for water levels and currents. Boxes filled with gray require values although there are no targets. Constituents 3 to 36 have been omitted here for convenience but are required.

N	Name	Oberved Amplitude	Observed Epoch	Modeled Amplitude	Modeled Epoch	Modeled minus Observed Amplitude	Modeled minus Observed Epoch
1	M_2						
2	S_2						
3...	...						
37	MS_4						

3.2. Computation of the Standard Suite of Statistics

The Standard Suite of statistics (as shown in Table 4) for water level variables [h, ahw, alw, thw, and tlw (Table 1)] are to be evaluated for the five scenarios in Table 2 (although results for the Test Forecast are optional). For h, hnn, ahw, and alw*,* the acceptable error *X* is 15 cm (0.5 ft) and is based on estimates of pilot's needs for under keel clearance. The limiting duration for these time series variables, *L*, is 24 hrs (1 day), and is based on the length of a typical nowcast or forecast cycle. For thw and tlw, *X* is ½ hr and *L* is 25 hrs (or the approximate time of two M_2 cycles), to assist in selecting port arrival and departure times. The template of the required values is shown in Table 6 and typical results are shown in Appendix A (Table A.3).

Although the acceptance requirements given here for the forecast are identical to those for the nowcast, it is recognized that in practice a forecast will generally be less accurate due to the uncertainties in the boundary forcing. Potential relaxation of the acceptance criteria must rely on further research.

Table 6. Template for water levels: the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used. For variable names, see Table 1.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L	WOF(2X) < 0.5%
1 Astro-Tide Only	H												
	h												
	H-h	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻	✻	✻	
	ALW-alw	″	″					✻	✻	✻	✻	✻	
	THW-thw	½ hr	25 hr					✻	✻	✻	✻	✻	
	TLW-tlw	″	″					✻	✻	✻	✻	✻	
2.Hind-cast	H												
	h												
	H-h	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻	✻	✻	
	ALW-alw	″	″					✻	✻	✻	✻	✻	
	THW-thw	½ hr	25 hr					✻	✻	✻	✻	✻	
	TLW-tlw	″	″					✻	✻	✻	✻	✻	
3. Test Fcst	H00 - h00	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻			
	ALW-alw	″	″					✻	✻	✻			
	THW-thw	½ hr	25 hr					✻	✻	✻			
	TLW-tlw	″	″					✻	✻	✻			
4. Semi-oper Ncst	H												
	h												
	H-h	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻	✻	✻	
	ALW-alw	″	″					✻	✻	✻	✻	✻	
	THW-thw	½ hr	25 hr					✻	✻	✻	✻	✻	
	TLW-tlw	″	″					✻	✻	✻	✻	✻	
5. Semi-oper Fcst	H00 - h00	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻			
	ALW-alw	″	″					✻	✻	✻			
	THW-thw	½ hr	25 hr					✻	✻	✻			
	TLW-tlw	″	″					✻	✻	✻			

The time series representing the nowcast will be constructed by appending individual nowcasts at a given time. For the forecast, there will be two sets. The first consists of forecasts all for the same projection time. The second consists of multiple time series used for extracting high and low waters, one or more per day depending on how often forecasts are made.

3.3. Comparison of Forecast Method

For this comparison, the Standard Suite of statistics for water level variables (h, hnn, ahw, alw, thw, tlw) must be evaluated for the three comparison forecasts: semioperational model-based forecast, astronomical tide, and tide plus persistence. A template is shown in Table 7 to help in comparing forecast skill. Note that the information for the semioperational forecast is identical to that in Table 6. A sample of the results is shown in Appendix A (Table A.3). The requirements to be met are given in Eqns. 9 and 10.

Table 7. Template for the Forecast Method Comparison for tidal water levels; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used. Note that the results for Scenario 5 (Semioperational Forecast) are identical to those in the template for the Standard Suite (Table 6).

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L	WOF(2X) < 0.5%
5. Semi-oper Fcst	H00 - h00	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻			
	ALW-alw	″	″					✻	✻	✻			
	THW-thw	½ hr	25 hr					✻	✻	✻			
	TLW-tlw	″	″					✻	✻	✻			
Astro-Tide Fcst	H-h	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻			
	ALW-alw	″	″					✻	✻	✻			
	THW-thw	½ hr	25 hr					✻	✻	✻			
	TLW-tlw	″	″					✻	✻	✻			
Astro-Tide & Persist	H00 - h00	15 cm	24 hr					✻	✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻			
	ALW-alw	″	″					✻	✻	✻			
	THW-thw	½ hr	25 hr					✻	✻	✻			
	TLW-tlw	″	″					✻	✻	✻			

4. EVALUATION OF PREDICTED CURRENTS IN TIDAL REGIONS

The following is a discussion of the statistics and acceptance criteria that are necessary for the evaluation of predicted water currents in regions where there are significant tidal variations. They are to be generated, analyzed, and documented by the SDIT with a precision sufficient for an accurate comparison. Current data will be at either NOS prediction depth (15 ft below MLLW) or one-half the MLLW depth, whichever is smaller.

The three sets of requirements for currents at each location where data are available are:

- comparison of tidal harmonic constants,
- computation of the Standard Suite of statistics, and
- comparison of forecast methods

The first requirement applies to currents along, and possibly normal to, the PCD, and the other requirements apply to both the current speed and the current direction, regardless of the PCD.

4.1. Comparison of Tidal Harmonic Constants

Here, tidal harmonic constants (37 amplitudes and phases) are analyzed from a tide-only model simulation and compared with values obtained directly from observations or from CO-OPS' historical harmonic constant data. First, the PCD of the modeled output (Eqns. 1, 2) and the value of the ratio R (Eqn. 3) must be computed. Then the comparison is made between the harmonic constants from the simulated, along-PCD currents and the harmonic constants obtained from the data along its (possibly different) PCD. It is expected that the PCD for model and observations will be within about 30 degrees. If the SD of the cross-PCD currents is large as compared to the SD of the along-PCD currents (i.e., R from Eqn. 3 greater than 0.25), a similar comparison must be made for them. The template for comparison of tidal constituents for current appears in Table 5 and a sample is shown in the Appendix (Table A.2). Note that the PCD and the value of R are required for both the modeled and observed time series.

4.2. Computation of the Standard Suite of Statistics

The Standard Suite (see Table 4) requires current speed [u, unn (see Table 1)], current direction (d, dnn), slack water times (tsf, tef, tse, tee), amplitude of maximum flood and ebb currents (afc, aec), times for maximum flood and ebb currents (tfc, tec), and direction of maximum flood and ebb currents (dfc, dec), all evaluated for the five scenarios in Table 2 (although results for the Test Forecast are optional). For current speed and maximum flood and ebb speeds, X is 26 cm/sec (0.5 kt); for time of maximum flood or ebb, X = 30 min; for slack water times, X = 15 minutes; and for current direction, X = 22.5 degrees, provided the current speed is not less than 26 cm/s (0.5 kt). These values of acceptable error are based on estimates of pilot's needs for maneuvering in ports and dredged channels. The limiting duration for these variables, L, is 24 hrs for the time series values, and 25 hrs for slack, flood, and ebb variables. For this requirement, the data should be arranged as shown in the template for current speed in Table 8 and for current direction in Table 9. Samples are shown in the Appendix (Tables A.4 and A.5)

Table 8. Template for tidal current speeds; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
1. Astro-Tide Only	U											
	u											
	U-u	26 cm/s	24 hr					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻	✻	✻
	AEC-aec	″	″					✻	✻	✻	✻	✻
	TFC-tfc	½ hr	25 hr					✻	✻	✻	✻	✻
	TEC-tec	″	″					✻	✻	✻	✻	✻
	TSF-tsf	¼ hr	″					✻	✻	✻	✻	✻
	TEF-tef	″	″					✻	✻	✻	✻	✻
	TSE-tse	″	″					✻	✻	✻	✻	✻
	TEE-tee	″	″					✻	✻	✻	✻	✻
2. Hind-cast	U											
	u											
	U-u	26 cm/s	24 hr					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻	✻	✻
	AEC-aec	″	″					✻	✻	✻	✻	✻
	TFC-tfc	½ hr	25 hr					✻	✻	✻	✻	✻
	TEC-tec	″	″					✻	✻	✻	✻	✻
	TSF-tsf	¼ hr	″					✻	✻	✻	✻	✻
	TEF-tef	″	″					✻	✻	✻	✻	✻
	TSE-tse	″	″					✻	✻	✻	✻	✻
	TEE-tee	″	″					✻	✻	✻	✻	✻
3. Test Fcst	U00 - u00	26 cm/s	24 hr					✻	✻	✻	✻	✻
	U06 - u06	″	″					✻	✻	✻	✻	✻
	U12 - u12	″	″					✻	✻	✻	✻	✻
	U18 - u18	″	″					✻	✻	✻	✻	✻
	U24 - u24	″	″					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻		
	AEC-aec	″	″					✻	✻	✻		
	TFC-tfc	½ hr	25 hr					✻	✻	✻		
	TEC-tec	″	″					✻	✻	✻		
	TSF-tsf	¼ hr	″					✻	✻	✻		
	TEF-tef	″	″					✻	✻	✻		
	TSE-tse	″	″					✻	✻	✻		
	TEE-tee	″	″					✻	✻	✻		
4. Semi-oper Ncst	U											
	u											
	U-u	26 cm/s	24 hr					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻	✻	✻
	AEC-aec	″	″					✻	✻	✻	✻	✻
	TFC-tfc	½ hr	25 hr					✻	✻	✻	✻	✻
	TEC-tec	″	″					✻	✻	✻	✻	✻
	TSF-tsf	¼ hr	″					✻	✻	✻	✻	✻
	TEF-tef	″	″					✻	✻	✻	✻	✻
	TSE-tse	″	″					✻	✻	✻	✻	✻
	TEE-tee	″	″					✻	✻	✻	✻	✻

(table continued on next page)

Table 8. Template for tidal current speeds (Continued).

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	≥		MDNO(2X) ≤L	MDPO(2X) ≤L
5. Semi-oper Fcst	U00 - u00	26 cm/s	24 hr					✻	✻	✻	✻	✻
	U06 - u06	″	″					✻	✻	✻	✻	✻
	U12 - u12	″	″					✻	✻	✻	✻	✻
	U18 - u18	″	″					✻	✻	✻	✻	✻
	U24 - u24	″	″					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻		
	AEC-aec	″	″					✻	✻	✻		
	TFC-tfc	½ hr	25 hr					✻	✻	✻		
	TEC-tec	″	″					✻	✻	✻		
	TSF-tsf	¼ hr	″					✻	✻	✻		
	TEF-tef	″	″					✻	✻	✻		
	TSE-tse	″	″					✻	✻	✻		
	TEE-tee	″	″					✻	✻	✻		

Table 9. Template for tidal current directions; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used. See Table 8 for currents speeds.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
1. Astro-Tide Only	D											
	d											
	D-d	22.5 dg	24 hr					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻	✻	✻
	DEC-dec	″	″					✻	✻	✻	✻	✻
2. Hind-cast	D											
	d											
	D-d	22.5 dg	24 hr					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻	✻	✻
	DEC-dec	″	″					✻	✻	✻	✻	✻
3. Test Fcst	D00 - d 00		24 hr					✻	✻	✻	✻	✻
	D06 - d 06	″	″					✻	✻	✻	✻	✻
	D12 - d 12	″	″					✻	✻	✻	✻	✻
	D18 - d 18	″	″					✻	✻	✻	✻	✻
	D24 - d 24	″	″					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻		
	DEC-dec	″	″					✻	✻	✻		
4. Semi-oper Ncst	D											
	d											
	D-d	22.5 dg	24 hr					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻	✻	✻
	DEC-dec	″	″					✻	✻	✻	✻	✻
5. Semi-oper Fcst	D00 - d 00		24 hr					✻	✻	✻	✻	✻
	D06 - d 06	″	″					✻	✻	✻	✻	✻
	D12 - d 12	″	″					✻	✻	✻	✻	✻
	D18 - d 18	″	″					✻	✻	✻	✻	✻
	D24 - d 24	″	″					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻		
	DEC-dec	″	″					✻	✻	✻		

4.3. Comparison of Forecast Method

For this comparison, the Standard Suite of statistics for current variables (u, d) must be evaluated for the three comparison forecasts: semioperational model-based forecast, astronomical tide, and tide plus persistence. The model-based forecasts should be better than the other forecasts. A template is shown in Table 10 for current speed and Table 11 for current direction to help in comparing forecast skill. Note that the information for the semioperational forecast is identical to that in Tables 8 and 9.

Table 10. Template for the Forecast Method Comparison for tidal current speeds; required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
5. Semi-oper Fcst	U00 – u00	26 cm/s	24 hr					✻	✻	✻	✻	✻
	U06 – u06	″	″					✻	✻	✻	✻	✻
	U12 – u12	″	″					✻	✻	✻	✻	✻
	U18 – u18	″	″					✻	✻	✻	✻	✻
	U24 – u24	″	″					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻		
	AEC-aec	″	″					✻	✻	✻		
	TFC-tfc	½ hr	25 hr					✻	✻	✻		
	TEC-tec	″	″					✻	✻	✻		
	TSF-tsf	¼ hr	25 hr					✻	✻	✻		
	TEF-tef	″	″					✻	✻	✻		
	TSE-tse	″	″					✻	✻	✻		
	TSF-tsf	″	″					✻	✻	✻		
Astro-Tide Fcst	U-u	26 cm/s	24 hr					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻		
	AEC-aec	″	″					✻	✻	✻		
	TFC-tfc	½ hr	25 hr					✻	✻	✻		
	TEC-tec	″	″					✻	✻	✻		
	TSF-tsf	¼ hr	25 hr					✻	✻	✻		
	TEF-tef	″	″					✻	✻	✻		
	TSE-tse	″	″					✻	✻	✻		
	TSF-tsf	″	″					✻	✻	✻		
Astro-Tide & Persist	U00 - u00	26 cm/s	24 hr					✻	✻	✻	✻	✻
	U06 - u06	″	″					✻	✻	✻	✻	✻
	U12 - u12	″	″					✻	✻	✻	✻	✻
	U18 - u18	″	″					✻	✻	✻	✻	✻
	U24 - u24	″	″					✻	✻	✻	✻	✻
	AFC-afc	″	″					✻	✻	✻		
	AEC-aec	″	″					✻	✻	✻		
	TFC-tfc	½ hr	25 hr					✻	✻	✻		
	TEC-tec	″	″					✻	✻	✻		
	TSF-tsf	¼ hr	25 hr					✻	✻	✻		
	TEF-tef	″	″					✻	✻	✻		
	TSE-tse	″	″					✻	✻	✻		
	TSF-tsf	″	″					✻	✻	✻		

Table 11. Template for the Forecast Method Comparison for tidal current directions; required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
5. Semi-oper Fcst	D00 - d 00	″	24 hr					✻	✻	✻	✻	✻
	D06 - d 06	″	″					✻	✻	✻	✻	✻
	D12 - d 12	″	″					✻	✻	✻	✻	✻
	D18 - d 18	″	″					✻	✻	✻	✻	✻
	D24 - d 24	″	″					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻		
	DEC-dec	″	″					✻	✻	✻		
Astro Tide Fcst	D-d	″	24 hr					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻		
	DEC-dec	″	″					✻	✻	✻		
Astro-Tide & Persist	D00 - d 00	″	24 hr					✻	✻	✻	✻	✻
	D06 - d 06	″	″					✻	✻	✻	✻	✻
	D12 - d 12	″	″					✻	✻	✻	✻	✻
	D18 - d 18	″	″					✻	✻	✻	✻	✻
	D24 - d 24	″	″					✻	✻	✻	✻	✻
	DFC- dfc	″	25 hr					✻	✻	✻		
	DEC-dec	″	″					✻	✻	✻		

For this analysis, the operational data streams with gaps must be preserved to recreate the operational environment. A sample of the results is shown in Appendix A (Tables A.4 and A.5). The requirements for current speed and direction and time of slack water are the same as Eqn. 9; in other words, the model-based forecast should give better results than the persistence or the tide-only forecasts.

5. EVALUATION OF PREDICTED WATER DENSITY, SALINITY, AND TEMPERATURE IN TIDAL REGIONS

The following is a discussion of the statistics and acceptance criteria that are necessary for the evaluation of predicted water density, salinity, and temperature. They are to be generated, analyzed, and documented by the SDIT with a precision sufficient for an accurate comparison.

5.1. Density and Draft

Water density, which depends upon salinity and temperature, is important since it determines a vessel's buoyancy. Based on input from the pilots, the desired accuracy of a forecast of a ship's draft is to the nearest 7.5 cm (3 in). If a vessel with a nominal draft D sits in water of depth-averaged density, ρ, the buoyant force per unit area (where area is perpendicular to the vertical) will be $\rho g D$, where g is the gravitational acceleration. Therefore, an error in the value of density, ρ', will lead to an error in draft, D', such that

$$(\rho + \rho')gD = \rho g(D - D') \quad (12)$$

Thus if the water is less dense than estimated, the buoyancy will be less, and the vessel's draft will be greater than estimated. Given an acceptable error in density, D', Eqn. 12 leads to a condition on the density error

$$|\rho'| \leq \rho D'/D \quad (13)$$

Equation 13 was evaluated using a multi-parameter density formulation (Mellor, 1991), assuming that salinity and temperature are depth-averaged values.

For an acceptable error in draft of 7.5 cm (3 inches) and a vessel draft of 15.25 m (50 ft) (approximately the largest existing today), the acceptable error X is 3.5 for salinity and 7.7°C for temperature. Note that for a draft error of 5.0 cm (2 inches), the corresponding values for X are 2.0 for salinity and 5.4°C for temperature.

The acceptable errors for salinity and temperature were determined as follows. For an arbitrary pair of values for salinity and temperature (S, T) and errors (S', T'), the density error is

$$\rho' = \rho(S + S', T + T') - \rho(S, T). \quad (14)$$

Now a region exists (Figure 5) that is bounded by an upper condition [density = $\rho(S, T) - \rho'$] and lower condition [density = $\rho(S, T) + \rho'$], within which the density error satisfies (13). By testing various values of salinity and temperature errors, the values S' and T' for this particular S and T can be found that both (a) satisfy (13) and (b) have the maximum area within the dashed rectangle (Figure 5) as defined by the product of S' and T'. For a set of test salinities (0, 10, 20, and 30) and test temperatures (0, 6, 12, 16, and 24°C), sets of S' and T' were found. The minimum values of S' and T' within each respective set then became the final criteria.

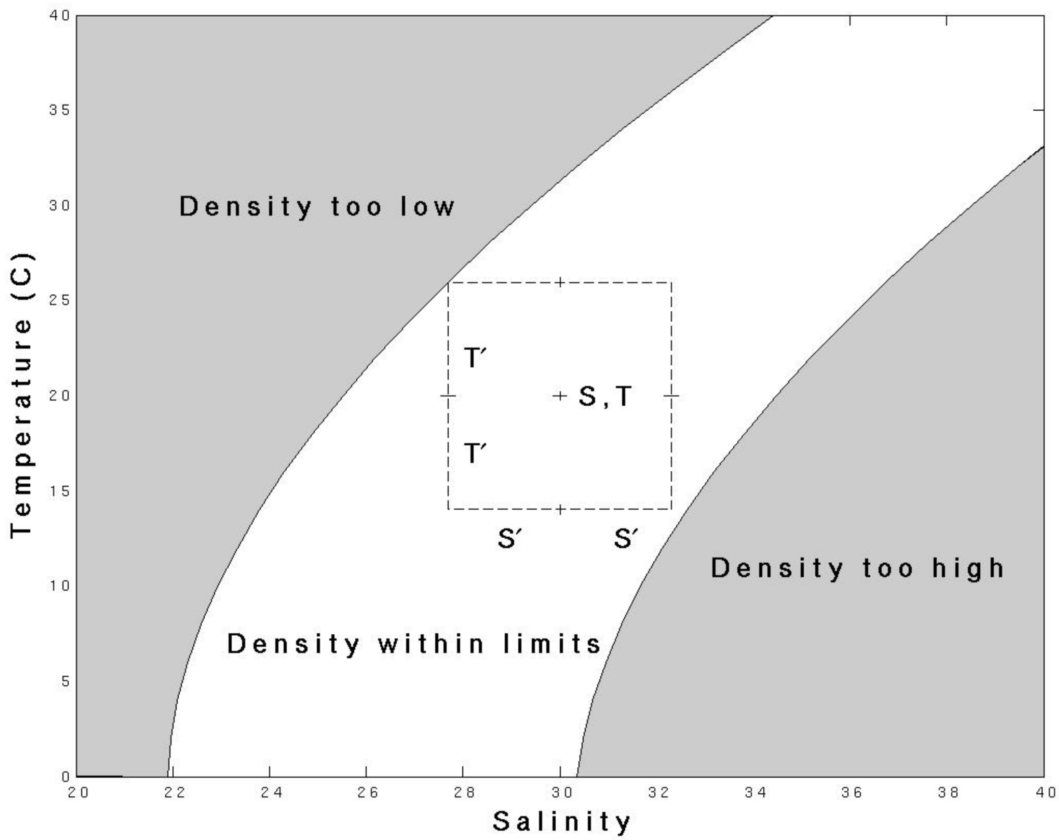

Figure 5. Water density characteristics as a function of salinity and temperature. For a given salinity, *S*, and temperature, *T*, there are maximum errors in salinity, *S'*, and temperature, *T'*, for which the error in density remains within the accepted range (denoted by the dashed rectangle). The region where density is within limits is bounded from above by density = $\rho(S, T) - \rho'$ and below by density = $\rho(S, T) + \rho'$.

5.2. Computation of the Standard Suite of Statistics

The requirements for salinity and temperature at each location and depth where data are available are the Standard Suite (Table 4) for salinity [*S* (Table 1)] and temperature (*T*) evaluated for the four scenarios in Table 2 (not the Astronomical Tide Only scenario). No harmonic constant data are available, so the comparison of harmonic constants is not required. The acceptable error *X* is as described above and the duration, *L*, is 24 hrs. The requirements are summarized in Tables 12 and 13, and samples are shown in the Appendix (Tables A.6, A.7).

Table 12. Template for salinity; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
2. Hind-cast	S											
	s											
	S-s	3..5	24 hr					✻	✻	✻	✻	✻
3. Test Fcst	S00 - s00	3.5	24 hr					✻	✻	✻	✻	✻
	S06 - s06	〃	〃					✻	✻	✻	✻	✻
	S12 - s12	〃	〃					✻	✻	✻	✻	✻
	S18 - s18	〃	〃					✻	✻	✻	✻	✻
	S24 - s24	〃	〃					✻	✻	✻	✻	✻
4. Semi-oper Ncst	S											
	s											
	S-s	3.5	24 hr					✻	✻	✻	✻	✻
5. Semi-oper Fcst	S00 - s00	3.5	24 hr					✻	✻	✻	✻	✻
	S06 - s06	〃	〃					✻	✻	✻	✻	✻
	S12 - s12	〃	〃					✻	✻	✻	✻	✻
	S18 - s18	〃	〃					✻	✻	✻	✻	✻
	S24 - s24	〃	〃					✻	✻	✻	✻	✻

Table 13. Template for water temperature; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
2 Hind-cast	T											
	t											
	T-t	7.7C	24 hr					✻	✻	✻	✻	✻
3. Test Fcst	T00 - t00	7.7C	24 hr					✻	✻	✻	✻	✻
	T06 - t06	〃	〃					✻	✻	✻	✻	✻
	T12 - t12	〃	〃					✻	✻	✻	✻	✻
	T18 - t18	〃	〃					✻	✻	✻	✻	✻
	T24 - t24	〃	〃					✻	✻	✻	✻	✻
4. Semi-oper Ncst	T											
	t											
	T-t	7.7C	24 hr					✻	✻	✻	✻	✻
5. Semi-oper Fcst	T00 - t00	7.7C	24 hr					✻	✻	✻	✻	✻
	T06 - t06	〃	〃					✻	✻	✻	✻	✻
	T12 - t12	〃	〃					✻	✻	✻	✻	✻
	T18 - t18	〃	〃					✻	✻	✻	✻	✻
	T24 - t24	〃	〃					✻	✻	✻	✻	✻

5.3. Comparison of Forecast Method

Comparison of forecast method can only be made if another method exists. One possibility is to take the mean of the observed values as the forecast value. For certain coastal areas, other methods, such as using the annual climatological variation, may exist. The comparisons of forecast methods are shown in Tables 14 and 15. Samples are shown in the Appendix (Tables A.6, A.7).

Table 14. Template for the Forecast Method Comparison for salinity; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
5. Semi-oper Fcst	S00 - s00	3.5	24 hr					✻	✻	✻	✻	✻
	S06 - s06	″	″					✻	✻	✻	✻	✻
	S12 - s12	″	″					✻	✻	✻	✻	✻
	S18 - s18	″	″					✻	✻	✻	✻	✻
	S24 - s24	″	″					✻	✻	✻	✻	✻
Fcst By Other Method	S00 - s00	″	″					✻	✻	✻	✻	✻
	S06 - s06	″	″					✻	✻	✻	✻	✻
	S12 - s12	″	″					✻	✻	✻	✻	✻
	S18 - s18	″	″					✻	✻	✻	✻	✻
	S24 - s24	″	″					✻	✻	✻	✻	✻

Table 15. Template for the Forecast Method Comparison for temperature; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error, L=limiting duration, and Imax is the number of data values used.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤1%	CF(X) ≥ 90%	POF(2X) ≤1%	MDNO(2X) ≤L	MDPO(2X) ≤L
5. Semi-oper Fcst	T00 - t00	7.7C	24 hr					✻	✻	✻	✻	✻
	T06 - t06	″	″					✻	✻	✻	✻	✻
	T12 - t12	″	″					✻	✻	✻	✻	✻
	T18 - t18	″	″					✻	✻	✻	✻	✻
	T24 - t24	″	″					✻	✻	✻	✻	✻
Fcst by Other Method	T00 - t00	″	″					✻	✻	✻	✻	✻
	T06 - t06	″	″					✻	✻	✻	✻	✻
	T12 - t12	″	″					✻	✻	✻	✻	✻
	T18 - t18	″	″					✻	✻	✻	✻	✻
	T24 - t24	″	″					✻	✻	✻	✻	✻

6. EVALUATION OF PREDICTED WATER LEVELS IN NON-TIDAL REGIONS

In areas without significant tidal variations (e.g., Great Lakes) or model simulations where the tides are excluded, the criteria of the preceding sections must be modified since the tidal statistics cannot be calculated. For example, there can be no comparison of harmonic constants for water levels or currents.

6.1. Computation of the Standard Suite of Statistics

In the Standard Suite of statistics, there can still be a comparison of time series differences for water levels using the SM, RMSE, SD, NOF, POF, MDPO, and MDNO statistics. These statistics can be computed for four scenarios (neglecting the Astronomical Tide Only). Although there will not be a comparison of the times and amplitudes of tidally-forced high and low waters, there should be some analysis of the simulation of large amplitude events. For example, given a time series of water level, a running average can be computed, and from it a departure. Then, using an acceptable error departure, an analysis of events can be constructed (c.f., Dingman and Bedford, 1986). For example, consider the time series of water levels measured at Toledo, Ohio, on Lake Erie for January 1 to 30, 2003 (Figure 6). By defining a running mean value (defined by averaging over a given time interval), a departure from the running mean can be computed. By defining a suitable threshold value (e.g., 0.5 m), significantly large events can be defined. Then, when the departure is significant, a comparison can then be made of model-based predictions to observed values (both timing and amplitude).

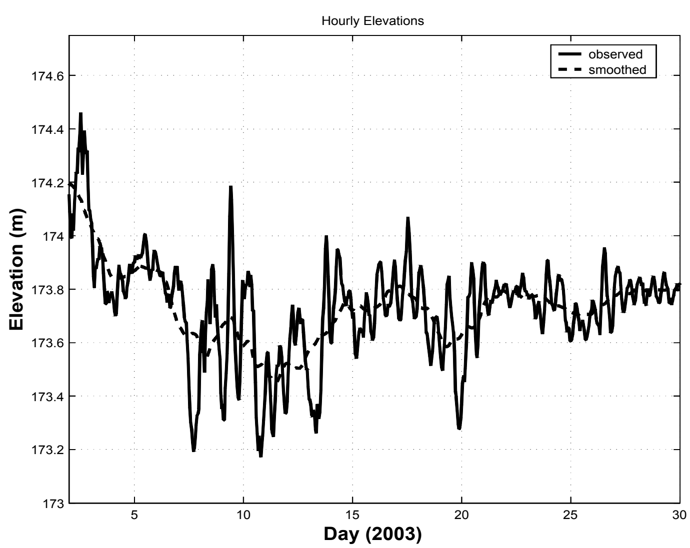

Figure 6. Hourly observed (solid line) and smoothed (dashed line) water levels at Toledo, Ohio, in January 2003. Smoothing is the simple average over ±36 hrs.

A similar approach, but geared to extremely high or low total water levels, has been described by Richardson and Schmalz (2002a). They searched for events, i.e., times when water levels exceeded a given threshold, and compared simulated and observed levels for both amplitude and phase of the peaks. In addition, they devised a scheme to categorize the simulation as a success, failure, or false alarm. The same authors also analyzed large current events in a similar way (Richardson and Schmalz, 2002b).

Since the same navigational considerations apply, the acceptable error in extrema amplitude is again 15 cm and the acceptable error in extrema time is 0.5 hr. However, the time period, *L*, for MDPO and MDNO must be defined. For a lake, the natural period of the lake would be a candidate. However, for predictions of non-tidal water levels in some coastal areas, *L* may not be able to be defined; in this case, MDPO and MDNO need not be calculated for extrema. The template for non-tidal water levels appears in Table 16.

Table 16. Template for non-tidal water levels; the required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error and Imax is the number of data values used. Note that P is must be defined for the individual area.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤ 1%	CF(X) ≥ 90%	POF(2X)≤ 1%	MDNO(2X) ≤L	MDPO(2X)≤L
2.Hind-cast	H											
	h											
	H-h	15 cm	P					✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻	✻	✻
	ALW-alw	″	″					✻	✻	✻	✻	✻
	THW-thw	½ hr	″					✻	✻	✻	✻	✻
	TLW-tlw	″	″					✻	✻	✻	✻	✻
3. Test Fcst	H00 - h00	15 cm	″					✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻		
	ALW-alw	″	″					✻	✻	✻		
	THW-thw	½ hr	″					✻	✻	✻		
	TLW-tlw	″	″					✻	✻	✻		
4. Semi-oper Ncst	H											
	h											
	H-h	15 cm	P					✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻	✻	✻
	ALW-alw	″	″					✻	✻	✻	✻	✻
	THW-thw	½ hr	″					✻	✻	✻	✻	✻
	TLW-tlw	″	″					✻	✻	✻	✻	✻
5. Semi-oper Fcst	H00 - h00	15 cm	″					✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻		
	ALW-alw	″	″					✻	✻	✻		
	THW-thw	½ hr	″					✻	✻	✻		
	TLW-tlw	″	″					✻	✻	✻		

6.2. Comparison of Forecast Method

It may be possible to compare the forecast of a model with a prediction made another way (e.g., normal mode analysis). If this can be done, the comparison of methods can be completed. The template for a forecast comparison of non-tidal water levels appears in Table 17.

Table 17. Template for the Forecast Method Comparison for non-tidal water levels; required statistics for a single station. Shaded boxes show required values and ✻ means that a target condition applies. X=acceptable error and Imax is the number of data values used. Note that P is must be defined for the individual area.

Scenario	Variable	X	L	Imax	SM	RMSE	SD	NOF(2X) ≤ 1%	CF(X) ≥ 90%	POF(2X)≤ 1%	MDNO(2X) ≤L	MDPO(2X)≤L
5. Semi-oper. Fcst	H00 - h00	15 cm	P					✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻		
	ALW-alw	″	″					✻	✻	✻		
	THW-thw	½ hr	″					✻	✻	✻		
	TLW-tlw	″	″					✻	✻	✻		
Other Fcst	H00 - h00	15 cm	P					✻	✻	✻	✻	✻
	H06 - h06	″	″					✻	✻	✻	✻	✻
	H12 - h12	″	″					✻	✻	✻	✻	✻
	H18 - h18	″	″					✻	✻	✻	✻	✻
	H24 - h24	″	″					✻	✻	✻	✻	✻
	AHW-ahw	″	″					✻	✻	✻		
	ALW-alw	″	″					✻	✻	✻		
	THW-thw	½ hr	″					✻	✻	✻		
	TLW-tlw	″	″					✻	✻	✻		

7. EVALUATION OF PRODUCTS AT SITES LACKING OBSERVATIONAL DATA

User needs may require issuing nowcast/forecast products for locations within an estuary where either real-time and/or historical observations are not available. Such products include both time series at individual locations and entire two-dimensional fields. The skill assessment of nowcast/forecast products at these locations must be handled differently than at locations where validation data are available. The procedures are as follows.

For a time series nowcast or forecast for a location where historical time series data exist, the model will be run to simulate the historical period (provided either wind data are available or wind effects are unimportant), or the historical data will be harmonically analyzed and compared to model-generated constituents. With constituent data only, model output will be harmonically analyzed for direct comparison.

For the case of locations where no historical data exist, the SDIT will analyze the model-generated data (time series or a field) and make a professional judgment as to its realism and the extent to which it captures actual features. The team will accomplish this by comparing the product to observational data from nearby stations and by assessing whether there are oceanographic reasons that it may be unrepresentative (e.g., location in an embayment separated from the main bay by a narrow, flow-restricting channel). This is especially important in assessing a current pattern which contains (or lacks) eddies and other features either known or hypothesized to exist.

The SDIT will present their results from the above analyses to the Technical Review Team. If the product is accepted, it can be disseminated for a trial period, but only with cautionary information. During initial dissemination, the nowcast field will be accompanied by a cautionary note (see below note 1) and will show the maximum of the errors at the locations with data for the time of the nowcast. A time loop will show the maximum error at the gauges for the period of the loop. A forecast field will be accompanied by a message showing/describing the forecast error bars. The SDIT will arrange for users to provide feedback on the accuracy and utility of the product. After a length of time sufficient for users to respond, the product will be reassessed by the SDIT for accuracy in light of comments. If revisions in the product or methodology are required, they will be implemented and the product will be issued for another trial period. If no revisions are necessary, the product will be disseminated with a *cautionary note* (see Note 2 below).

1. CAUTION: This product is presently being evaluated and should not be used for navigation or other official purposes. Probable error of water levels less than xx cm. (Probable error in current speed is less than xx cm/s and direction is less than xx degrees).

2. CAUTION: This product, although based on the best available information, should be used with caution for navigation or other official purposes. Probable error of water levels less than xx cm. (Probable error in current speed is less than xx cm/s and direction is less than xx degrees).

8. FUTURE DIRECTIONS

Although the major part of the skill assessment problem has been addressed in this document, there are other topics that remain to be researched.

Skill Assessment Software - Once the model scenarios have been run, the skill assessment scores could, in principle, be computed automatically by a software package. Using data files containing observed, nowcast, and forecast variables, these data could be processed and the results displayed in tables such as those in Appendix A. Possibly the numerical values could be color-coded to identify where the criteria were not being met. The processing routines could include harmonic analysis, gap filling, filtering (or singular value decomposition), and the other methods discussed in Section 2.4. Of particular value would be ways of concatenating forecasts and in extracting water level and current extrema.

Skill Assessment for Non-navigational Applications - Non-navigation uses of forecast system products include oil spill advection, algal bloom transport, ecosystem simulation, and coastal inundation. Each has special requirements for accuracy that need to be expressed before the skill measures can be selected.

Skill Assessment of Water Density Stratification - Although a description of the assessment of water density appears in this report, assessment of vertical density stratification does not. Since water density is nearly uniform, it is likely that salinity and temperature stratification would be assessed instead. Stratification data tends to consist of either (a) vertical profiles at widely-spaced intervals in space and time, or (b) continuous measurements at fixed vertical levels, usually near the surface and the bottom. Often, stratification is used as an indicator of vertical mixing. Therefore, useful magnitudes of mixing need to be described and related to stratification and other variables such as currents.

Predicted Forecast Error - Forecast products could be accompanied by an indication of predicted forecast error or uncertainty for each forecast. The purpose of displaying the predicted error is to give the user a measure of the likelihood that the forecast will be reliable. Predicted error can be depicted graphically in two ways: as error bars or as an ensemble of forecasts. Error bars will show, at each forecast hour, the upper and lower limits that bound a fixed percentage (e.g., 90%) of the test forecast results for that hour. Event-dependent error bars show the limits for a specific set of situations. For example, error bars may depend on whether the forecast winds over the region, when averaged over 12 hours, have a direction from either the north or the south. Other categories for developing event-dependent error bars could be based on a different averaging period, on averaged wind speed, or on the averaged non-tidal water level. Implementation of error bars will ultimately depend on the requirements of the user.

The ensemble approach requires several forecast runs, each with a different wind, ocean water level, or river flow forecast. Each combination will produce a unique water level forecast. The upper and lower bounding forecasts would be the limits, with the mean at each hour forming the forecast. Note that each model and input combination would independently have to meet all skill assessment requirements.

Operational Assessment of Error - Normally, the operational system differs somewhat from the semi-operational version due to changes in the code and/or the inputs. Therefore, the operational system ordinarily requires either additional periodic assessments or a continually running assessment. A continually running assessment of model skill includes daily calculation and update of skill parameters, and will alert system users and developers to potential problems. Continual assessment may require time series of variables to be archived so that the skill parameters for a daily, weekly, monthly, and/or other long-term assessment is possible.

Summary Statistics - For each location where data are available, the required number of statistics is relatively large (c.f., Appendix A). Therefore, a statistic or small number of statistics that summarizes the total would be useful for comparison purposes.

Event Analysis – It may be desirable to assess how the nowcast/forecast system performs during large events such as storm surges or high wind conditions.

9. SUMMARY OF REQUIREMENTS

The following summarizes the requirements for the skill assessment of water levels, currents, salinity, and water temperatures for each location in the model domain where observational data are available (see Table 18). Recall that the relevant variables appear in Table 1, the model run scenarios are given in Table 2, definitions of skill assessment statistics are given in Table 3, and the Standard Suite of statistics is given in Table 4.

As discussed in Sec. 2.6, the length of time each scenario is to be run is, ideally, 365 days in order to capture all expected seasonal conditions. However, some scenarios can be run concurrently to reduce the time required for implementation. When significant data are missing or other circumstances arise, the Technical Review Team may reduce the 365-day requirement. Normally, data at 6 min intervals is required.

Table 18. Summary of the requirements for water levels, currents, salinity, and temperatures. Notes: (1) This is required only if the ratio of the standard deviations of the across and along-PCD speeds is greater than 0.25. (2) For current speeds not less than 0.5 kts. (3) This is required providing another forecast method exists.

Variable	Requirement	Template Table No.	Example Table No.
Water Levels in Tidal Regions	Harmonic Constant Comparison	5	A.1
	Standard Suite	6	A.3
	Comparison of Forecast Method	7	A.3
Currents in Tidal Regions	Along-PCD Harmonic Constant Comparison	5	A.2
	Across-PCD Harmonic Constant Comparison[1]	5	A.2
	Standard Suite for Current Speed	8	A.4
	Comparison of Forecast Method for Current Speed	10	A.4
	Standard Suite for Current Direction[2]	9	A.5
	Comparison of Forecast Method for Current Direction[2]	11	A.5
Salinity	Standard Suite	12	A.6
	Comparison of Forecast Method[3]	13	A.6
Temperature	Standard Suite	14	A.7
	Comparison of Forecast Method[3]	15	A.7
Water Levels in Non-Tidal Regions	Standard Suite	16	A.8
	Comparison of Forecast Method[3]	17	A.8

9.1. Harmonic Constant Comparison

For regions where there are significant tidal variations, the model is run in the Astronomical Tide Only scenario. From the model output for water levels and currents at each location in the model domain where data area available, a time series of hourly values covering at least 180 days (preferably 365 days) is constructed. See Section 2.4 for an explanation of gap-filling and filtering techniques. These time series are harmonically analyzed and compared to harmonic constants obtained from either a database of historical values (i.e., from CO-OPS) or by harmonic analysis of an observational series.

For modeled currents, the Principal Current Direction (PCD) is computed (Eqns. 1 and 2) for both the model and the observations, and the harmonic analysis is applied for currents along these directions. The value of R (Eqn. 3) is also computed. If the value of R is greater than 0.25, then the harmonic analysis is required for the currents in the direction normal to the PCD. If historical harmonic constants are available, these may be used for comparison as long as their principal direction close to the modeled PCD (e.g., is within 20 degrees of). If observational currents are available, they are analyzed for their principal direction and the harmonic analysis is applied for currents along this direction.

9.2. Computation of the Standard Suite of Statistics

For all regions, the model is run in the Hindcast, Test Forecast (this is optional), Semioperational Nowcast, and Semioperational Forecast scenarios. For regions where there are significant tidal variations, the model is also run in the Astronomical Tide Only scenario. From the model output for water levels and currents at each location in the model domain where data area available (except for the Semioperational Forecast), a time series of 6-minute values covering at least 180 days (preferably 365 days) is constructed. For the Test Nowcast and the Semioperational Nowcast, all nowcasts are appended to produce a single series. For the Semioperational Forecast, the time series used is each individual forecast at 6-min intervals; another data set consists of all forecast values at a given projection time.

Using these modeled and observed time series, the statistics in the Standard Suite are computed. Current directions are computed only for speeds not less than 0.5 kt.

9.3. Comparison of Forecast Method

For tidal water levels and currents, the values in the Standard Suite for the Semioperational Forecast are compared to the values in the Standard Suite for the astronomical tide (based on harmonic constants at a location), and for a tide plus persistence forecast. A tide plus persistence water level forecast is constructed by adding an offset value, based on an observed offset at that station during some time period before the forecast is made, to the tide prediction at each station. For currents, the offset may be a mean current.

For salinity, temperature, and non-tidal water levels, some other method of forecasting should be found. For example, for salinity and temperature, the climatological variation could be used. For lakes, a normal-mode prediction could be used.

9.4. Acceptance Criteria

For a nowcast or forecast at a particular station to be approved for release to the public, the statistics related to model performance at that station must (a) meet or exceed *all* target frequencies or durations, or (b) meet or exceed *most* of the target frequencies or durations and be granted a *waiver* by the Technical Review Team.

ACKNOWLEDGMENTS

This report is a product of a cooperative effort between personnel from CSDL and CO-OPS. CSDL's Bruce Parker supplied additional guidance for the model evaluation process. CSDL's Lloyd Huff and Maureen Kenny helped identify relevant variables and quantify accuracy requirements.

REFERENCES

Dingman, J. S., and K. W. Bedford, 1986. Skill tests and parametric statistics for model evaluation. **Journal of Hydraulic Engineering**, 112, 124 - 141.

Gill, S., W. Stoney, and T. Bethem, 1997. System Development Plan. CORMS: Continuous Operational Real-Time Monitoring System. **NOAA Technical Report** NOS OES 14. U.S Department of Commerce, NOAA. 41 pp.

Mellor, G. L., 1991: An equation of state for numerical models of oceans and estuaries. **Journal of Atmospheric and Oceanic Technology**, v. 8 (4), 609 - 611.

NOS, 1999. NOS Procedures for Developing and Implementing Operational Nowcast and Forecast Systems for PORTS. **NOAA Technical Report** NOS CO-OPS 20, 33 pp.

Preisendorfer, R. W., 1988. Principal Component Analysis in Meteorology and Oceanography. **Developments in Atmospheric Science**, v. 17. Elsevier, Amsterdam. pp 425.

Richardson, P. H., and R. A. Schmalz Jr., 2002a. Water Level Event Analysis: Program Documentation. **CSDL Informal Technical Note** 1, 90 pp.

Richardson, P. H., and R. A. Schmalz Jr., 2002b. Principal Component Direction Current Event Analysis: Program Documentation. **CSDL Informal Technical Note** 2, 100 pp.

Schureman, P., 1958. Manual of Harmonic Analysis and Prediction of Tides. U.S. Department of Commerce, Coast and Geodetic Survey, **Special Publication** No. 98 [revised 1940 edition, reprinted 1988]. 317 pp.

Vincent, M., K. W. Hess, and J. G. W. Kelley, 2003. NOS Procedures for Developing and Implementing Operational Nowcast and Forecast Hydrodynamic Model Systems, **NOAA Technical Report** NOS CO-OPS 39, 17 pp.

Williams, R. G., T. D. Bethem and H. R. Frey, 1989: Tampa Bay Current Prediction Quality Assurance Miniproject. **NOAA Technical Memorandum** NOS OMA 50, National Ocean Service, Office of Oceanography and Marine Assessment, Rockville, Maryland, 11 pp+figs.

Zervas, C., 1999. Tidal Current Analysis Procedures and Associated Computer Programs. **NOAA Technical Report** NOS CO-OPS 21, 101 pp.

APPENDIX A. SAMPLES OF REQUIRED DATA SHEETS

The following tables show samples of all required skill assessment data sheets. Numerical values in the tables are for demonstration purposes only.

Table A.1. Comparison of tidal constituent amplitudes and epochs for TIDAL WATER LEVELS. The amplitudes for water levels are in meters and the epochs are in degrees. Tidal epoch may be local or Greenwich, but must be consistent.

		Observed		Modeled		Difference	
N	Constituent	Amplitude	Epoch	Amplitude	Epoch	Amplitude	Epoch
1	M2	0.580	210.6	0.607	213.4	0.027	2.8
2	S2	0.137	218.4	0.141	220.7	0.004	2.3
3	N2	0.123	184.9	0.129	187.0	0.006	2.1
4	K1	0.368	226.5	0.376	226.8	0.008	0.3
5	M4	0.023	142.0	0.029	135.9	0.006	-6.1
6	O1	0.230	210.1	0.231	209.7	0.001	-0.4
7	M6	0.000	335.6	0.000	335.6	0.000	0.0
8	MK3	0.019	129.0	0.026	124.0	0.007	-5.0
9	S4	0.000	120.0	0.000	120.0	0.000	0.0
10	MN4	0.009	119.3	0.011	108.7	0.002	-10.6
11	NU2	0.026	188.3	0.027	191.6	0.001	3.3
12	S6	0.000	0.0	0.000	0.0	0.000	0.0
13	MU2	0.007	100.9	0.007	88.2	0.000	-12.7
14	2N	0.014	158.4	0.014	154.2	0.000	-4.2
15	OO	0.011	261.1	0.013	262.7	0.002	1.6
16	LAM2	0.006	217.5	0.008	215.0	0.002	-2.5
17	S1	0.007	284.8	0.007	276.5	0.000	-8.3
18	M1	0.011	244.5	0.013	248.6	0.002	4.1
19	J1	0.019	243.8	0.019	248.6	0.000	4.8
20	MM	0.000	4.3	0.000	4.3	0.000	0.0
21	SSA	0.039	286.9	0.039	287.0	0.000	0.1
22	SA	0.038	221.4	0.038	221.4	0.000	0.0
23	MSF	0.000	8.1	0.000	8.1	0.000	0.0
24	MF	0.000	8.7	0.000	8.8	0.000	0.1
25	RHO1	0.009	202.3	0.010	198.5	0.001	-3.8
26	Q1	0.040	203.2	0.040	202.8	0.000	-0.4
27	T2	0.009	196.6	0.008	201.9	-0.001	5.3
28	R2	0.001	218.7	0.001	221.0	0.000	2.3
29	2Q	0.004	208.0	0.005	186.3	0.001	-21.7
30	P1	0.116	223.3	0.119	223.9	0.003	0.6
31	2SM	0.000	248.1	0.000	248.1	0.000	0.0
32	M3	0.005	26.8	0.009	29.8	0.004	3.0
33	L2	0.016	225.8	0.019	225.5	0.003	-0.3
34	2MK3	0.014	96.7	0.019	91.5	0.005	-5.2
35	K2	0.040	209.9	0.041	211.6	0.001	1.7
36	M8	0.000	207.4	0.000	207.5	0.000	0.1
37	MS4	0.010	154.3	0.012	147.2	0.002	-7.1

Table A.2. Comparison of tidal constituent amplitudes and epochs for TIDAL CURRENTS. In the abbreviated table, the amplitudes for currents are in m/s and the epochs are in degrees. Note that currents require PCD and values of R. Tidal epoch may be local or Greenwich, but must be consistent.

		Observed Current (R=0.85)		Modeled Current (R=0.92)		Difference	
N	Constituent	Amplitude	Epoch	Amplitude	Epoch	Amplitude	Epoch
CURRENT ALONG PCD:		DIRECTION = 272		DIRECTION = 282			
1	M2	0.580	210.6	0.607	213.4	0.027	2.8
2	S2	0.137	218.4	0.141	220.7	0.004	2.3
3	N2	0.123	184.9	0.129	187.0	0.006	2.1
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
33	L2	0.016	225.8	0.019	225.5	0.003	-0.3
34	2MK3	0.014	96.7	0.019	91.5	0.005	-5.2
35	K2	0.040	209.9	0.041	211.6	0.001	1.7
36	M8	0.000	207.4	0.000	207.5	0.000	0.1
37	MS4	0.010	154.3	0.012	147.2	0.002	-7.1
CURRENT ACROSS PCD:		DIRECTION = 92		DIRECTION = 102			
1	M2	0.580	210.6	0.607	213.4	0.027	2.8
2	S2	0.137	218.4	0.141	220.7	0.004	2.3
3	N2	0.123	184.9	0.129	187.0	0.006	2.1
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
33	L2	0.016	225.8	0.019	225.5	0.003	-0.3
34	2MK3	0.014	96.7	0.019	91.5	0.005	-5.2
35	K2	0.040	209.9	0.041	211.6	0.001	1.7
36	M8	0.000	207.4	0.000	207.5	0.000	0.1
37	MS4	0.010	154.3	0.012	147.2	0.002	-7.1

Table A.3. Example of skill assessment scores for TIDAL WATER LEVELS.

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO	WOF
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L	<.5%
SCENARIO 1: ASTRONOMICAL TIDE ONLY												
H			4801	-0.002								
h			4801	-0.001								
H-h	15cm	24h	4801	-0.001	0.110	0.110	0.0	80.5	0.0	0.0	0.0	11.35
AHW-ahw	15cm	24h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0	
ALW-hwl	15cm	24h	40	0.048	0.084	0.070	0.0	92.5	0.0	0.0	0.0	
THW-thw	.5h	25h	40	-0.145	0.278	0.241	0.0	85.0	0.0	0.0	0.0	
TLW-twl	.5h	25h	40	-0.255	0.383	0.290	0.0	77.5	0.0	0.0	0.0	
SCENARIO 2: HINDCAST												
H			4801	-0.006								
h			4801	-0.087								
H-h	15cm	24h	4801	0.081	0.082	0.014	0.0	100.0	0.0	0.0	0.0	0.00
AHW-ahw	15cm	24h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0	
ALW-alw	15cm	24h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0	
THW-thw	.5h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0	
SCENARIO 3: TEST FORECAST												
H00-h00	15cm	24h	76	0.075	0.082	0.034	0.0	98.7	0.0	0.0	0.0	0.00
H06-h06	15cm	24h	76	0.071	0.092	0.059	1.3	98.7	0.0	0.0	0.0	1.32
H12-h12	15cm	24h	75	0.069	0.092	0.061	1.3	98.7	0.0	0.0	0.0	1.33
H18-h18	15cm	24h	74	0.070	0.090	0.057	1.4	97.3	0.0	0.0	0.0	1.35
H24-h24	15cm	24h	73	0.067	0.090	0.059	1.4	97.3	0.0	0.0	0.0	1.37
AHW-ahw	15cm	24h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0	
HLW-alw	15cm	24h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0	
THW-thw	.5h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0	
SCENARIO 4: SEMI-OPERATIONAL NOWCAST												
H			4801	-0.006								
h			4801	-0.087								
H-h	15cm	24h	4801	0.081	0.082	0.014	0.0	100.0	0.0	0.0	0.0	0.00
AHW-ahw	15cm	24h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0	
ALW-alw	15cm	24h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0	
THW-thw	.5h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0	
SCENARIO 5: SEMI-OPERATIONAL FORECAST												
H00-h00	15cm	24h	76	0.075	0.082	0.034	0.0	98.7	0.0	0.0	0.0	0.00
H06-h06	15cm	24h	76	0.071	0.092	0.059	1.3	98.7	0.0	0.0	0.0	1.32
H12-h12	15cm	24h	75	0.069	0.092	0.061	1.3	98.7	0.0	0.0	0.0	1.33
H18-h18	15cm	24h	74	0.070	0.090	0.057	1.4	97.3	0.0	0.0	0.0	1.35
H24-h24	15cm	24h	73	0.067	0.090	0.059	1.4	97.3	0.0	0.0	0.0	1.37
AHW-ahw	15cm	24h	41	0.092	0.160	0.132	0.0	70.7	9.8	0.0	0.1	
ALW-alw	15cm	24h	41	0.069	0.174	0.162	0.0	70.7	7.3	0.0	0.1	
THW-thw	.5h	25h	41	0.846	0.917	0.357	0.0	9.8	31.7	0.0	0.3	
TLW-tlw	.5h	25h	41	0.868	0.906	0.261	0.0	4.9	22.0	0.0	0.1	
COMPARISON: ASTRONOMICAL TIDE ONLY FORECAST												
H-h	15cm	24h	4801	0.086	0.140	0.110	0.0	60.8	0.0	0.0	0.0	0.00
AHW-ahw	15cm	24h	42	0.078	0.137	0.114	0.0	61.9	0.0	0.0	0.0	
ALW-alw	15cm	24h	42	0.098	0.150	0.115	0.0	57.1	2.4	0.0	0.0	
THW-thw	.5h	25h	42	0.019	0.295	0.298	2.4	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	-0.033	0.248	0.249	2.4	95.2	0.0	0.0	0.0	
COMPARISON: PERSISTENCE FORECAST												
H00-h00	15cm	24h	76	0.087	0.099	0.049	0.0	96.1	1.3	0.0	0.0	1.32
H06-h06	15cm	24h	76	0.089	0.107	0.060	0.0	90.8	0.0	0.0	0.0	0.00
H12-h12	15cm	24h	75	0.098	0.114	0.058	0.0	92.0	0.0	0.0	0.0	0.00
H18-h18	15cm	24h	74	0.086	0.105	0.060	0.0	94.6	1.4	0.0	0.0	1.35
H24-h24	15cm	24h	73	0.080	0.106	0.070	1.4	94.5	1.4	0.0	0.0	2.74
AHW-ahw	15cm	24h	41	0.100	0.141	0.100	0.0	65.9	4.9	0.0	0.1	
ALW-alw	15cm	24h	42	0.090	0.168	0.143	0.0	66.7	4.8	0.0	0.1	
THW-thw	.5h	25h	41	0.866	0.931	0.345	0.0	7.3	26.8	0.0	0.2	
TLW-tlw	.5h	25h	42	0.850	0.896	0.288	0.0	4.8	19.0	0.0	0.2	

Table A.4. Example of skill assessment scores for TIDAL CURRENT SPEEDS.

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L
SCENARIO 1: ASTRONOMICAL TIDE ONLY											
U			4801	-0.132							
u			4801	-0.091							
U-u	26cm/s	24h	4801	-0.001	0.110	0.110	0.0	80.5	0.0	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	40	-0.046	0.128	0.127	0.0	75.0	0.0	0.0	0.0
TEF-tef	.25h	25h	40	0.048	0.084	0.070	0.0	92.5	0.0	0.0	0.0
TSE-tse	.25h	25h	40	-0.145	0.278	0.241	0.0	85.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	40	-0.255	0.383	0.290	0.0	77.5	0.0	0.0	0.0
SCENARIO 2: HINDCAST											
U			4801	-0.006							
u			4810	-0.087							
U-u	26cm/s	24h	4810	0.081	0.082	0.014	0.0	100.0	0.0	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0
TEF-tef	.25h	25h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0
TSE-tse	.25h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0
TSF-tsf	.25h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0
SCENARIO 3: TEST FORECAST											
U00-u00	26cm/s	24h	76	0.075	0.082	0.034	0.0	98.7	0.0	0.0	0.0
U06-u06	26cm/s	24h	76	0.071	0.092	0.059	1.3	98.7	0.0	0.0	0.0
U12-u12	26cm/s	24h	75	0.069	0.092	0.061	1.3	98.7	0.0	0.0	0.0
U18-u18	26cm/s	24h	74	0.070	0.090	0.057	1.4	97.3	0.0	0.0	0.0
U24-u24	26cm/s	24h	73	0.067	0.090	0.059	1.4	97.3	0.0	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0
TEF-tef	.25h	25h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0
TSE-tse	.25h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0
TLW-twl	.25h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0
SCENARIO 4: SEMI-OPERATIONAL NOWCAST											
U			4810	-0.006							
u			4810	-0.087							
U-u	26cm/s	24h	4810	0.081	0.082	0.014	0.0	100.0	0.0	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0
TEF-tef	.25h	25h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0
TSE-tse	.25h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0
TSF-tsf	.25h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0

(Continued on next page)

Table A.4 (Continued).

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L
SCENARIO 5: SEMI-OPERATIONAL FORECAST											
U00-u00	26cm/s	24h	76	0.075	0.082	0.034	0.0	98.7	0.0	0.0	0.0
U06-u06	26cm/s	24h	76	0.071	0.092	0.059	1.3	98.7	0.0	0.0	0.0
U12-u12	26cm/s	24h	75	0.069	0.092	0.061	1.3	98.7	0.0	0.0	0.0
U18-u18	26cm/s	24h	74	0.070	0.090	0.057	1.4	97.3	0.0	0.0	0.0
U24-u24	26cm/s	24h	73	0.067	0.090	0.059	1.4	97.3	0.0	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	41	0.092	0.160	0.132	0.0	70.7	9.8	0.0	0.1
TEF-tef	.25h	25h	41	0.069	0.174	0.162	0.0	70.7	7.3	0.0	0.1
TSE-tse	.25h	25h	41	0.846	0.917	0.357	0.0	9.8	31.7	0.0	0.3
TSF-tsf	.25h	25h	41	0.868	0.906	0.261	0.0	4.9	22.0	0.0	0.1
COMPARISON: ASTRONOMICAL TIDAL CURRENT ONLY FORECAST											
U-u	26cm/s	24h	4801	0.086	0.140	0.110	0.0	60.8	0.0	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	42	0.078	0.137	0.114	0.0	61.9	0.0	0.0	0.0
TEF-tef	.25h	25h	42	0.098	0.150	0.115	0.0	57.1	2.4	0.0	0.0
TSE-tse	.25h	25h	42	0.019	0.295	0.298	2.4	95.2	2.4	0.0	0.0
TSF-tsf	.25h	25h	42	-0.033	0.248	0.249	2.4	95.2	0.0	0.0	0.0
COMPARISON: PERSISTENCE FORECAST											
U00-u00	26cm/s	24h	76	0.087	0.099	0.049	0.0	96.1	1.3	0.0	0.0
U06-u06	26cm/s	24h	76	0.089	0.107	0.060	0.0	90.8	0.0	0.0	0.0
U12-u12	26cm/s	24h	75	0.098	0.114	0.058	0.0	92.0	0.0	0.0	0.0
U18-u18	26cm/s	24h	74	0.086	0.105	0.060	0.0	94.6	1.4	0.0	0.0
U24-u24	26cm/s	24h	73	0.080	0.106	0.070	1.4	94.5	1.4	0.0	0.0
AFC-afc	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
AEC-aec	26cm/s	25h	40	0.101	0.057	0.044	0.2	82.2	1.3	1.0	0.0
TFC-tfc	.25h	25h	40	-0.046	0.128	0.121	0.0	75.0	0.0	0.0	0.0
TEC-tec	.25h	25h	40	-0.046	0.128	0.133	0.0	75.0	0.0	0.0	0.0
TSF-tsf	.25h	25h	41	0.100	0.141	0.100	0.0	65.9	4.9	0.0	0.1
TEF-tef	.25h	25h	42	0.090	0.168	0.143	0.0	66.7	4.8	0.0	0.1
TSE-tse	.25h	25h	41	0.866	0.931	0.345	0.0	7.3	26.8	0.0	0.2
TSF-tsf	.25h	25h	42	0.850	0.896	0.288	0.0	4.8	19.0	0.0	0.2

Table A.5. Example of skill assessment scores for TIDAL CURRENT DIRECTIONS.

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L
SCENARIO 1: ASTRONOMICAL TIDE ONLY											
D			4801	183.1							
d			4801	174.2							
D-d	22.5deg	24h	4801	-14.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
SCENARIO 2: HINDCAST											
D			4801	183.1							
d			4801	174.2							
D-d	22.5deg	24h	4801	-14.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
SCENARIO 3: TEST FORECAST											
D00-d00	22.5deg	24h	76	-14.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
D06-d06	22.5deg	24h	76	-14.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
D12-d12	22.5deg	24h	75	-14.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
D18-d18	22.5deg	24h	74	-14.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
D24-d24	22.5deg	24h	73	-14.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4
SCENARIO 4: SEMI-OPERATIONAL NOWCAST											
D			4801	183.1							
d			4801	174.2							
D-d	22.5deg	24h	4801	-14.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
SCENARIO 5: SEMI-OPERATIONAL FORECAST											
D00-d00	22.5deg	24h	76	-14.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
D06-d06	22.5deg	24h	76	-14.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
D12-d12	22.5deg	24h	75	-14.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
D18-d18	22.5deg	24h	74	-14.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
D24-d24	22.5deg	24h	73	-14.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4
COMPARISON: ASTRONOMICAL TIDAL CURRENT ONLY FORECAST											
D-d	22.5deg	24h	4801	-13.3	0.11	0.11	0.0	80.5	0.0	0.0	0.0
COMPARISON: PERSISTENCE FORECAST											
D00-d00	22.5deg	24h	76	-14.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
D06-d06	22.5deg	24h	76	-14.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
D12-d12	22.5deg	24h	75	-14.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
D18-d18	22.5deg	24h	74	-14.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
D24-d24	22.5deg	24h	73	-14.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4

Table A.6. Example of skill assessment scores for SALINITY.

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L
SCENARIO 2: HINDCAST											
S			4801	24.3							
s			4801	27.1							
S-s	3.5 ppt	24h	4801	3.1	4.1	3.5	3.1	91.3	3.0	22.0	24.4
SCENARIO 3: TEST FORECAST											
S00-s00	3.5 ppt	24h	76	4.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
S06-s06	3.5 ppt	24h	76	5.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
S12-s12	3.5 ppt	24h	75	6.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
S18-s18	3.5 ppt	24h	74	8.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
S24-s24	3.5 ppt	24h	73	10.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4
SCENARIO 4: SEMI-OPERATIONAL NOWCAST											
S			4801	24.3							
s			4801	27.1							
S-s	3.5 ppt	24h	4801	3.1	4.1	3.5	3.1	91.3	3.0	22.0	24.4
SCENARIO 5: SEMI-OPERATIONAL FORECAST											
S00-s00	3.5 ppt	24h	76	4.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
S06-s06	3.5 ppt	24h	76	5.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
S12-s12	3.5 ppt	24h	75	6.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
S18-s18	3.5 ppt	24h	74	8.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
S24-s24	3.5 ppt	24h	73	10.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4
COMPARISON: OTHER FORECAST METHOD											
S00-s00	3.5 ppt	24h	76	4.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
S06-s06	3.5 ppt	24h	76	5.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
S12-s12	3.5 ppt	24h	75	6.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
S18-s18	3.5 ppt	24h	74	8.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
S24-s24	3.5 ppt	24h	73	10.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4

Table A.7. Example of skill assessment scores for WATER TEMPERATURE.

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L
SCENARIO 2: HINDCAST											
T			4801	24.3							
t			4801	27.1							
T-t	7.7 degC	24h	4801	3.1	4.1	3.5	3.1	91.3	3.0	22.0	24.4
SCENARIO 3: TEST FORECAST											
T00-t00	7.7 degC	24h	76	4.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
T06-t06	7.7 degC	24h	76	5.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
T12-t12	7.7 degC	24h	75	6.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
T18-t18	7.7 degC	24h	74	8.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
T24-t24	7.7 degC	24h	73	10.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4
SCENARIO 4: SEMI-OPERATIONAL NOWCAST											
T			4801	24.3							
t			4801	27.1							
T-t	7.7 degC	24h	4801	3.1	4.1	3.5	3.1	91.3	3.0	22.0	24.4
SCENARIO 5: SEMI-OPERATIONAL FORECAST											
T00-t00	7.7 degC	24h	76	4.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
T06-t06	7.7 degC	24h	76	5.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
T12-t12	7.7 degC	24h	75	6.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
T18-t18	7.7 degC	24h	74	8.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
T24-t24	7.7 degC	24h	73	10.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4
COMPARISON: OTHER FORECAST METHOD											
T00-t00	7.7 degC	24h	76	4.3	10.1	8.5	4.1	92.3	2.0	12.0	14.4
T06-t06	7.7 degC	24h	76	5.3	12.1	10.5	4.1	92.3	2.0	12.0	14.4
T12-t12	7.7 degC	24h	75	6.3	14.1	12.5	4.1	92.3	2.0	12.0	14.4
T18-t18	7.7 degC	24h	74	8.3	18.1	13.5	4.1	92.3	2.0	12.0	14.4
T24-t24	7.7 degC	24h	73	10.3	22.1	15.5	4.1	92.3	2.0	12.0	14.4

Table A.8. Example of skill assessment scores for NON-TIDAL WATER LEVELS.

Variable	X	L	Imax	SM	RMSE	SD	NOF	CF	POF	MDNO	MDPO	WOF
Criterion	-	-	-	-	-	-	<1%	>90%	<1%	<L	<L	<.5%
SCENARIO 2: HINDCAST												
H			4801	-0.006								
h			4801	-0.087								
H-h	15cm	24h	4801	0.081	0.082	0.014	0.0	100.0	0.0	0.0	0.0	0.00
AHW-ahw	15cm	24h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0	
ALW-alw	15cm	24h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0	
THW-thw	.5h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0	
SCENARIO 3: TEST FORECAST												
H00-h00	15cm	24h	76	0.075	0.082	0.034	0.0	98.7	0.0	0.0	0.0	0.00
H06-h06	15cm	24h	76	0.071	0.092	0.059	1.3	98.7	0.0	0.0	0.0	1.32
H12-h12	15cm	24h	75	0.069	0.092	0.061	1.3	98.7	0.0	0.0	0.0	1.33
H18-h18	15cm	24h	74	0.070	0.090	0.057	1.4	97.3	0.0	0.0	0.0	1.35
H24-h24	15cm	24h	73	0.067	0.090	0.059	1.4	97.3	0.0	0.0	0.0	1.37
AHW-ahw	15cm	24h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0	
ALW-alw	15cm	24h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0	
THW-thw	.5h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0	
SCENARIO 4: SEMI-OPERATIONAL NOWCAST												
H			4801	-0.006								
h			4801	-0.087								
H-h	15cm	24h	4801	0.081	0.082	0.014	0.0	100.0	0.0	0.0	0.0	0.00
AHW-ahw	15cm	24h	42	0.084	0.097	0.049	0.0	97.6	2.4	0.0	0.0	
ALW-alw	15cm	24h	42	0.093	0.116	0.070	0.0	95.2	4.8	0.0	0.0	
THW-thw	.5h	25h	42	0.052	0.234	0.231	0.0	95.2	2.4	0.0	0.0	
TLW-tlw	.5h	25h	42	0.057	0.243	0.239	0.0	92.9	2.4	0.0	0.0	
SCENARIO 5: SEMI-OPERATIONAL FORECAST												
H00-h00	15cm	24h	76	0.075	0.082	0.034	0.0	98.7	0.0	0.0	0.0	0.00
H06-h06	15cm	24h	76	0.071	0.092	0.059	1.3	98.7	0.0	0.0	0.0	1.32
H12-h12	15cm	24h	75	0.069	0.092	0.061	1.3	98.7	0.0	0.0	0.0	1.33
H18-h18	15cm	24h	74	0.070	0.090	0.057	1.4	97.3	0.0	0.0	0.0	1.35
H24-h24	15cm	24h	73	0.067	0.090	0.059	1.4	97.3	0.0	0.0	0.0	1.37
AHW-ahw	15cm	24h	41	0.092	0.160	0.132	0.0	70.7	9.8	0.0	0.1	
ALW-alw	15cm	24h	41	0.069	0.174	0.162	0.0	70.7	7.3	0.0	0.1	
THW-thw	.5h	25h	41	0.846	0.917	0.357	0.0	9.8	31.7	0.0	0.3	
TLW-tlw	.5h	25h	41	0.868	0.906	0.261	0.0	4.9	22.0	0.0	0.1	
COMPARISON: OTHER FORECAST METHOD												
H00-h00	15cm	24h	76	0.087	0.099	0.049	0.0	96.1	1.3	0.0	0.0	1.32
H06-h06	15cm	24h	76	0.089	0.107	0.060	0.0	90.8	0.0	0.0	0.0	0.00
H12-h12	15cm	24h	75	0.098	0.114	0.058	0.0	92.0	0.0	0.0	0.0	0.00
H18-h18	15cm	24h	74	0.086	0.105	0.060	0.0	94.6	1.4	0.0	0.0	1.35
H24-h24	15cm	24h	73	0.080	0.106	0.070	1.4	94.5	1.4	0.0	0.0	2.74
AHW-ahw	15cm	24h	41	0.100	0.141	0.100	0.0	65.9	4.9	0.0	0.1	
ALW-alw	15cm	24h	42	0.090	0.168	0.143	0.0	66.7	4.8	0.0	0.1	
THW-thw	.5h	25h	41	0.866	0.931	0.345	0.0	7.3	26.8	0.0	0.2	
TLW-tlw	.5h	25h	42	0.850	0.896	0.288	0.0	4.8	19.0	0.0	0.2	