

Sea Surface Height Predictions from the Global Navy Coastal Ocean Model during 1998–2001*

CHARLIE N. BARRON, A. BIROL KARA, HARLEY E. HURLBURT, C. ROWLEY, AND LUCY F. SMEDSTAD

Naval Research Laboratory, Oceanography Division, Stennis Space Center, Mississippi

(Manuscript received 9 September 2003, in final form 19 June 2004)

ABSTRACT

A $\frac{1}{8}^\circ$ global version of the Navy Coastal Ocean Model (NCOM), operational at the Naval Oceanographic Office (NAVOCEANO), is used for prediction of sea surface height (SSH) on daily and monthly time scales during 1998–2001. Model simulations that use 3-hourly wind and thermal forcing obtained from the Navy Operational Global Atmospheric Prediction System (NOGAPS) are performed with/without data assimilation to examine indirect/direct effects of atmospheric forcing in predicting SSH. Model–data evaluations are performed using the extensive database of daily averaged SSH values from tide gauges in the Atlantic, Pacific, and Indian Oceans obtained from the Joint Archive for Sea Level (JASL) center during 1998–2001. Model–data comparisons are based on observations from 282 tide gauge locations. An inverse barometer correction was applied to SSH time series from tide gauges for model–data comparisons, and a sensitivity study is undertaken to assess the impact of the inverse barometer correction on the SSH validation. A set of statistical metrics that includes conditional bias (B_{cond}), root-mean-square (rms) difference, correlation coefficient (R), and nondimensional skill score (SS) is used to evaluate the model performance. It is shown that global NCOM has skill in representing SSH even in a free-running simulation, with general improvement when SSH from satellite altimetry and sea surface temperature (SST) from satellite IR are assimilated via synthetic temperature and salinity profiles derived from climatological correlations. When the model was run from 1998 to 2001 with NOGAPS forcing, daily model SSH comparisons from 612 yearlong daily tide gauge time series gave a median rms difference of 5.98 cm (5.77 cm), an R value of 0.72 (0.76), and an SS value of 0.45 (0.51) for the $\frac{1}{8}^\circ$ free-running (assimilative) NCOM. Similarly, error statistics based on the 30-day running averages of SSH time series for 591 yearlong daily tide gauge time series over the time frame 1998–2001 give a median rms difference of 3.63 cm (3.36 cm), an R value of 0.83 (0.85), and an SS value of 0.60 (0.64) for the $\frac{1}{8}^\circ$ free-running (assimilated) NCOM. Model–data comparisons show that skill in 30-day running average SSH time series is as much as 30% higher than skill for daily SSH. Finally, SSH predictions from the free-running and assimilative $\frac{1}{8}^\circ$ NCOM simulations are validated against sea level data from the tide gauges in two different ways: 1) using original detided sea level time series from tide gauges and 2) using the detided data with an inverse barometer correction derived using daily mean sea level pressure extracted from NOGAPS at each location. Based on comparisons with 612 yearlong daily tide gauge time series during 1998–2001, the inverse barometer correction lowered the median rms difference by about 1 cm (15%–20%). Results presented in this paper reveal that NCOM is able to predict SSH with reasonable accuracies, as evidenced by model simulations performed during 1998–2001. In an extension of the validation over broader ocean regions, the authors find good agreement in amplitude and distribution of SSH variability between NCOM and other operational model products.

1. Introduction

The usefulness of any particular ocean general circulation model (OGCM) is usually tested by examining how well its output conforms to a given set of known (i.e., observed or measured) conditions. In particular, free-running or assimilative versions of OGCMs that

are developed toward operational use must go through a rigorous evaluation. This evaluation will determine whether or not the accuracy of the model predictions is acceptable. In general, free-running simulations determine model performance in response to the atmospheric forcing. Therefore, it is necessary to have free-running simulations that are as realistic as possible before applying any kind of ocean data assimilation. This is especially true when a global ocean model is developed toward a real-time forecasting system using real-time atmospheric forcing.

Global ocean prediction is usually driven by the need for ocean models to balance a requirement for high vertical resolution required near the surface to resolve the physics of the surface boundary layers, with require-

* Naval Research Laboratory Contribution Number NRL/JA/7320/03/0112.

Corresponding author address: Charlie N. Barron, Naval Research Laboratory, Code 7323, Bldg. 1009, Stennis Space Center, MS 39529-5004.
E-mail: barron@nrlssc.navy.mil

ments for horizontal resolution to resolve eddy dynamics. Prediction of sea surface height (SSH) is important in global ocean prediction because it is a better indication of subsurface frontal location than is sea surface temperature (SST). Analysis of the SSH field allows one to better track front and eddy movements, equatorial and coastal trapped waves, storm surges, and geostrophically balanced surface currents. In addition, SSH is a good integral measure of subsurface temperature and salinity variations.

A $\frac{1}{8}^\circ$ global version of the recently developed Navy Coastal Ocean Model (NCOM) has been introduced, along with general features of the model and the assimilation scheme (Barron et al. 2005, manuscript submitted to *Ocean Modelling*). Model–data intercomparisons with unassimilated observations were performed to measure NCOM effectiveness in predicting upper-ocean quantities such as SST, sea surface salinity (SSS), mixed layer depth (MLD), and subsurface temperature and salinity (Kara et al. 2004). In this paper, daily SSH fields from NCOM are compared with unassimilated tide gauge data using two interannual simulations over 1998–2001 with prescribed 3-hourly atmospheric forcing. One simulation is free-running; the other includes assimilation of SSH from satellite altimetry and SST from satellite IR via synthetic temperature and salinity profiles derived using climatological correlations. These results help demonstrate the capability and assess the skill of global NCOM, which will be a real-time operational global ocean prediction system in 2004.

The major difficulty in evaluating global OGCMs has been a lack of global oceanic datasets of sufficient quality and duration to characterize the error statistics. It is obvious that validating results from a global ocean model at a single location is not very informative in measuring the model's actual skill. Such a validation reflects few details of the model's successes or deficiencies over the global ocean. In the particular case of a global ocean model, it is vital to validate model results in both coastal regions and in open ocean regions. It is also necessary to establish a quantitative framework to examine model errors with respect to observations. Within the quantitative framework, a set of validation criteria along with statistical error analysis is needed for quality assessment of an OGCM.

High-resolution, quality-controlled, daily averaged SSH time series are available from tide gauges in the Atlantic, Pacific, and Indian Oceans (e.g., Caldwell and Merrifield 2000). With the availability of these observations, model–data comparisons can be made using a variety of statistical metrics. In this paper, these data are applied in a detailed set of statistical metrics to evaluate NCOM performance in predicting SSH variability over the global ocean. In particular, daily averaged SSH from 612 yearlong daily time series from 282 tide gauges are used during 1998–2001 to determine effectiveness of NCOM in predicting SSH. For the analysis, we use a nondimensional metric (skill score) in

addition to using a classical rms difference. The former allows one to quantify model discrepancies that are not taken into account by the correlation coefficient. In addition, statistical model–data comparisons are presented at selected tide gauge locations individually, so that any other OGCM prediction of SSH could be compared with the $\frac{1}{8}^\circ$ NCOM results.

This paper is organized as follows. Section 2 gives a brief overview of NCOM parameterizations with a specific focus on SSH. Section 3 presents the atmospheric forcing along with the SSH assimilation scheme used in the model. Section 4 introduces sea level data obtained from many tide gauges over the global ocean. Section 5 provides model–data comparisons in predicting SSH when using either the free-running or data-assimilative version of NCOM. Section 5 also examines sensitivity of NCOM SSH validation to application of an inverse barometer correction to original tide gauge values and presents comparisons of spatial SSH variability among various models. Finally, section 6 gives the conclusions of this study.

2. Ocean model description

The OGCM used in this paper is the global version of NCOM with a variable horizontal resolution of $\sim\frac{1}{8}^\circ$ in latitude at midlatitudes. The model has a free surface and is based on the primitive equations with the hydrostatic, Boussinesq, and incompressible approximations. The physics and numerics of NCOM are based largely on the Princeton Ocean Model (POM) as described in Blumberg and Mellor (1987), with some aspects from the Sigma/Z-level Model (SZM) (Martin et al. 1998).

In this application, NCOM includes 19 terrain-following σ levels and 21 fixed z levels (a total of 40 material levels). The NCOM vertical grid uses σ coordinates from the surface down to a user-specified depth (z_s) and z levels below. Within the σ portion of the grid, each level is a fixed fraction of the total thickness occupied by the σ levels, as in POM. In the z level portion of the grid, bottom depth is rounded to the nearest z level. In general, the use of combined σ and z level coordinates in a single ocean model provides some flexibility in grid configuration. For example, σ coordinates in the upper levels improve resolution of changes in sea surface elevation and enhance representation of vertical shear and bottom depth in shallow regions. Underlying z levels below z_s avoid the problems associated with σ levels in regions of steep topography and limit the combined thickness of all σ levels. Since each σ level is a fixed fraction of the σ total, the underlying z levels enable NCOM to maintain consistent high resolution near the surface in deep-water regions, one important advantage of σ/z over pure σ coordinates. The vertical coordinate is logarithmically stretched to focus resolution near the surface with a maximum rest surface level thickness of 1 m.

NCOM surface boundary conditions are the surface

stress for the momentum equations, the surface heat flux for the temperature equation, and the effective surface salt flux for the salinity equation. The bottom boundary conditions are the bottom drag for the momentum equations, which is parameterized by a quadratic drag law, and zero flux for the temperature and salinity equations. The model equations are solved on an Arakawa C grid and the horizontal grid is orthogonal curvilinear, as in POM (Blumberg and Herring 1987).

The basic NCOM equations can be found in Barron et al. (2005). Here, we focus on SSH predictions from the free-running and data-assimilative simulations from NCOM. Thus, we only give a brief summary of the representation of SSH, which is the main focus of this paper. NCOM SSH is the free-surface elevation (ζ) above the undisturbed reference at $z = 0$, or equivalently the deviation of the total water column thickness from its rest thickness H , where H is the bottom depth (positive downward). Thus, a SSH of 0 would imply that the ocean is at its rest thickness at that point. Its time evolution is parameterized as

$$\frac{\partial \zeta}{\partial t} = -\frac{\partial[(\zeta + H)\bar{u}]}{\partial x} - \frac{\partial[(\zeta + H)\bar{v}]}{\partial y} + (\zeta + H)\bar{Q}, \quad (1)$$

where \bar{Q} is the depth-averaged mass flux source, and \bar{u} and \bar{v} are the depth-averaged horizontal velocities. The coordinate transformation for the sigma-coordinate part of the grid depends on ζ , given by

$$\sigma = \frac{z - \zeta}{\zeta + \min(H, z_s)} \quad (2)$$

(Martin 2000). The free-surface evolution is calculated implicitly. Hence, the surface pressure gradients and the divergence terms in the surface-elevation equation have a component at the new time level being calculated. These pressure gradient and divergence terms are incorporated according to user-selectable weightings at the new ($n + 1$), current (n), and previous ($n - 1$) time levels. The weighting currently being employed is an even split between $n + 1$ and $n - 1$.

River inflow in NCOM is a positive mass flux that, in a closed domain like global NCOM, would lead to a global net increase in SSH in the absence of a balancing evaporative flux. Since this net precipitation and evaporation are not presently included, a pseudo-evaporative mass flux is uniformly applied over the ocean surface to balance the river inflow each time step and maintain at zero the global area-averaged SSH.

3. Model simulations

This section describes global NCOM simulations along with the atmospheric forcing and the data assimilation scheme.

a. Global NCOM setup

The NCOM simulations in this paper were performed on a curvilinear grid extending from 80°S to the full arctic cap, a global configuration designed to maintain a grid-cell horizontal aspect ratio near 1. Horizontal resolution varies from 19.5 km near the equator to 8 km or finer in the Arctic Ocean, with midlatitude resolution of $\sim 1/8^\circ$ latitude (14 km) at 45° latitude. Horizontal resolution has been sacrificed to allow increased vertical resolution. The model has a total of 40 vertical material layers: 19 σ levels from the surface to $z_s = 137$ m depth and 21 z levels from 137 m to the maximum depth at 5500 m. The coordinates are stretched to focus resolution near the surface so that NCOM maintains a maximum 1-m upper-level rest thickness in this hybrid σ/z vertical configuration. Model depth and coastline are based on a 2' Digital Bathymetric Data Base (DBDB2) bathymetry produced at the Naval Research Laboratory (NRL), Stennis Space Center. The model includes all continental shelves for depths >5 m over the global ocean.

The model forcing is calculated from the following time-varying atmospheric fields: wind stress, air temperature, air mixing ratio, and net solar radiation. Most of these are taken from the Navy Operational Global Atmospheric Prediction System (NOGAPS), whose main features are discussed in Rosmond et al. (2002). The sensible and latent heat fluxes are strongly dependent on SST and are calculated every time step using the model SST in bulk formulations that include effects of air-sea stability through the exchange coefficients (Kara et al. 2002). The annual climatological SST cycle is built into the model to a limited extent. Including air temperature in the formulations for latent and sensible heat flux along with model SST in the bulk formulation automatically provides a physically realistic tendency toward the correct SST, as discussed in another OGCM study by Kara et al. (2003). Although radiation fluxes also depend on SST to some extent, these fluxes are obtained directly from NOGAPS in order to use the atmospheric cloud mask.

Before performing any interannual NCOM simulations, the model was first spun up to statistical energy equilibrium for about 6 years with monthly mean wind stress and surface heat fluxes obtained from NOGAPS. The model simulation was then extended interannually using 6-hourly atmospheric wind and thermal forcing from NOGAPS. The first year under realistic forcing, 1997, included assimilation to produce a suitable initial state for starting validation in 1998. There are two $1/8^\circ$ global NCOM model simulations used in this paper: 1) free-running (i.e., atmospheric forcing only) and 2) with ocean data assimilation. Both simulations started from the same initial state and were extended using the interannual NOGAPS wind/thermal forcing from 1998 to 2001, as explained above. These model simulations are used to examine the performance of NCOM in pre-

dicting SSH over the global ocean on daily time scales. Because all forcing is interannual, the SSH predicted by NCOM can be compared with interannual sea level data obtained from many tide gauges around the world, as explained in detail in section 4.

b. NCOM ocean data assimilation

In data-assimilative mode, NCOM assimilates SSH and SST via synthetic 3D temperature and salinity fields provided by the Modular Ocean Data Analysis System (MODAS). Thus, SSH is assimilated only indirectly, according to statistics from the historical hydrographic database as compiled within MODAS, which is explained in detail by Fox et al. (2002a,b). The basic 3D MODAS temperature fields are derived according to bimonthly climatological correlations in the historical hydrographic database, between steric SSH anomalies, SST anomalies, and subsurface temperature anomalies, where these anomalies are relative to the MODAS climatology. SST is taken from the operational model-independent MODAS 2D global analysis of multichannel SST (MCSST). These analyses cover the global ocean up to 80°N at uniform 1/8° resolution (Fox et al. 2002b). Steric SSH is extracted from the 1/16° global Naval Research Laboratory Layered Ocean Model (NLOM), an operational eddy-resolving global ocean nowcast/forecast system (Smedstad et al. 2003). Both are run operationally for the U.S. Navy at the Naval Oceanographic Office (NAVOCEANO). NLOM assimilates altimeter SSH track data using an optimal interpolation (OI) analysis (Smedstad et al. 2003) with mesoscale covariance calculated from TOPEX/Poseidon and *European Remote Sensing Satellite-2 (ERS-2)* altimeter data (Jacobs et al. 2002). NLOM also assimilates the MODAS 2D SST analyses. A detailed description of NLOM and its prediction capabilities is given in Wallcraft et al. (2003). Kara et al. (2003, 2005, manuscript submitted to *Ocean Modelling*) discuss NLOM performance in predicting SST in a free-running model. Operational NLOM, while excluding the Arctic and water shallower than 200 m, has higher horizontal resolution than global NCOM. Thus, it is better suited for direct assimilation of altimeter data and forecasting the location of fronts and eddies.

NLOM SSH is paired with MODAS 2D SST to derive the MODAS 3D temperature and salinity fields. In regions where the NLOM domain does not overlap the NCOM domain, SST-only synthetics are used with blending to joint SSH–SST synthetics along the edge of the NLOM domain. In the Arctic, the fields are blended toward the Generalized Digital Environmental Model (GDEM)-3 Arctic climatology (M. Carnes 2003, personal communication). MODAS salinity is derived from the MODAS temperature field according to the local climatological temperature and salinity (T – S) relations.

NCOM 3D relaxation in the assimilative simulations is parameterized by

$$T_{\text{ASSIM}} = (1 - b)T_{\text{PRE}} + bT_{\text{MODAS}}, \quad (3)$$

where T_{MODAS} is the assimilated field interpolated to the NCOM time step, T_{PRE} is the NCOM field before assimilation, T_{ASSIM} is the NCOM field after assimilation, and b is the spatially variable weighting function. Relaxation is spatially applied to each model time step. In general, b is a variable in three dimensions and in time. For the assimilative case in the paper,

$$b = \frac{\Delta t}{t_{\text{scale}}} [1 - e^{(-\text{depth}/d_{\text{scale}})}], \quad (4)$$

where Δt is the model time step, t_{scale} is the relaxation time scale, d_{scale} is the relaxation depth scale, and the depth is the depth of the model point. Here, t_{scale} is 48 h, Δt is 6 min, and d_{scale} is 200 m. These values are based on prior experiments and are expected to be revised in more advanced data assimilation.

Surface assimilation of SST and SSS is via adjustment of the surface temperature and salinity fluxes. For a temperature flux R ($^{\circ}\text{C m s}^{-1}$),

$$R_{\text{adjusted}} = R + a(\text{SST}_{\text{ASSIM}} - \text{SST}_{\text{NCOM}}), \quad (5)$$

where a is relaxation rate (in m day^{-1}). In the present application, $a_{\text{SST}} = 2^{\circ}\text{C m s}^{-1}$ and $a_{\text{SSS}} = 0.5 \text{ m day}^{-1}$, reflecting lower confidence in MODAS SSS with parameters based on sensitivity studies.

4. Sea level data

Since global NCOM has been developed to predict ocean quantities in both open ocean and coastal regions, it is necessary to examine model performance in as many places as possible by including both coastal and open ocean locations. In this paper, NCOM sea level intercomparisons are performed using sea level measurements from a total of 282 tide gauges located in different regions of the global ocean (Fig. 1).

The sea level data from those tide gauges are maintained by the Joint Archive for Sea Level (JASL) Center, which is a cooperative effort between the U.S. National Oceanographic Data Center (NODC) and the University of Hawaii Sea Level (UHSL) Center (Caldwell 1992; Caldwell and Merrifield 2000). JASL acquires, quality controls, manages, and distributes sea level data over the global ocean, including hourly data from regional and national sea level networks. At JASL the data are inspected and obvious errors, such as data spikes and time shifts, are corrected. Gaps less than 25 h are interpolated. This quality-controlled JASL dataset is presently the largest global collection of sea level measurements. For NCOM–data comparisons we use daily averaged detided sea level data from JASL over the period from 1998 to 2001. The tide gauges are located both at islands in the open ocean and in coastal regions where the depth of the water is often very shallow, and the continental shelf varies from wide to minimal. The variety and coverage of the tide gauge locations make

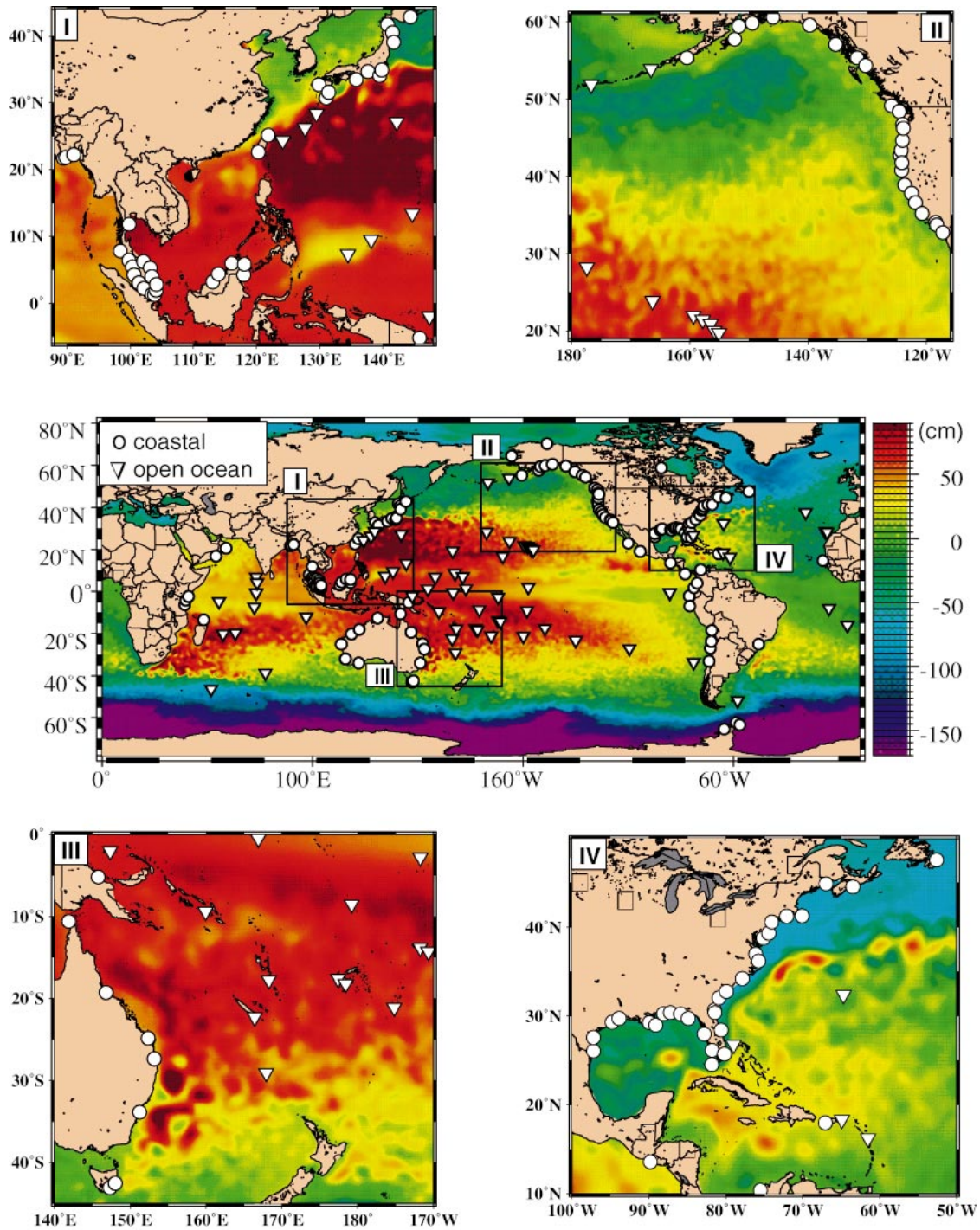


FIG. 1. Tide gauge locations superimposed on the annual mean of climatological sea surface height from the data-assimilative $\frac{1}{8}^\circ$ NCOM. All the tide gauges are maintained by the University of Hawaii JASL center and are located in both coastal and open ocean island regions of the Atlantic, Pacific, and Indian Oceans. The tide gauges are separated into two categories located at (a) coastal regions and (b) islands in open ocean.

these data very useful in validating daily NCOM SSH over the global ocean.

In order to compare the sea level time series from JASL with those from NCOM, daily SSH values from

each tide gauge were first adjusted for the inverse barometer effect (e.g., Gill 1982). The inverse barometer correction provides a simple approximation of the sea level response to forcing from air pressure changes. The

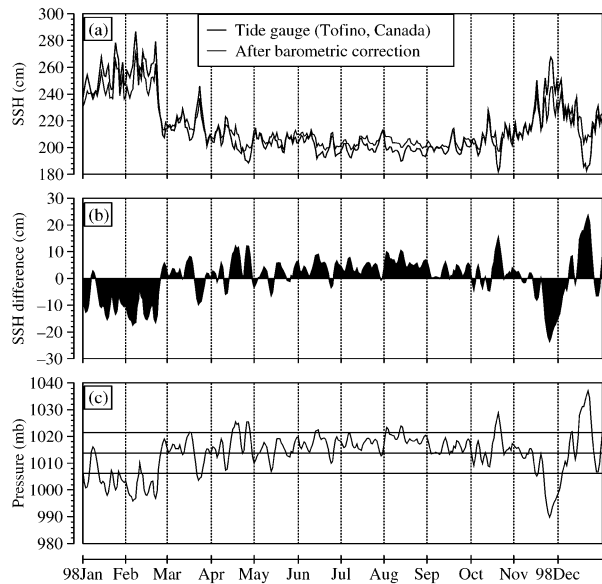


FIG. 2. A sample application of the inverse barometer correction to a tide gauge time series at Tofino, Canada (49°N, 126°W), in 1998. (a) Daily SSH time series directly obtained from JASL tide gauge data and the one obtained after applying the barometer correction, (b) daily difference between original tide gauge values and corrected ones, and (c) daily mean sea level atmospheric pressure values obtained from NOGAPS and mean pressure \pm standard deviations shown with horizontal lines. Note that an inverse barometer correction value of 0.9955 cm mb⁻¹ is applied to SSH from tide gauges. The reader is referred to the text for further details of the inverse barometer correction.

correction requires knowledge of daily mean atmospheric pressure sea level at all 282 tide gauge locations. For this purpose a daily averaged mean sea level pressure time series was derived from an archive of NOGAPS mean sea level pressure fields at each tide gauge location. The time series of pressure every 6 hours were filtered using a 13-point boxcar filter, a simple weighted average with uniform weight. Thus, daily pressure values are at each day a 72-h mean centered on that day.

Taking gravitational acceleration $g = 9.8 \text{ m s}^{-2}$ and water density $\rho = 1025 \text{ kg m}^{-3}$, we have $\rho \times g = 10\,045 \text{ kg m}^{-2} \text{ s}^{-2}$. Given that $1 \text{ Pa m}^{-1} = 100.45 \text{ m m}^{-1}$, we then obtain $dh/dp = 0.009\,955 \text{ m mb}^{-1}$. An inverse barometer adjustment of $0.009\,955 \text{ m mb}^{-1}$ was then added (subtracted) to (from) the sea level observation for pressure above (below) a reference value of 1013.3 mb. In summary, the pressure correction is a height increase where the NOGAPS pressure is high (the sea level would be higher in the absence of high pressure) and a height decrease where the NOGAPS pressure is low. The barometer correction was applied to SSH time series at each individual tide gauge location.

As an example, the inverse barometer correction is shown for Tofino, located on the Canadian coast at (49°N, 126°W) in Fig. 2. At this particular location there is no data void in 1998, allowing one to see changes in

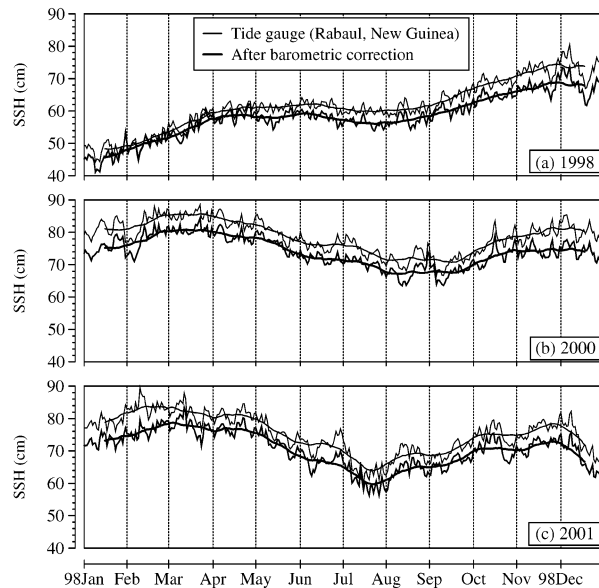


FIG. 3. Daily mean SSH values from a tide gauge at 2°S, 147°E in 1998, 2000, and 2001. Daily SSH time series are shown before/after the barometric correction was applied in each year. Also shown are 30-day running averages of SSH values.

SSH before/after the correction for a full 1-yr period. The most obvious effect of the barometric correction is a change in the standard deviation of SSH. While the standard deviation of SSH is 22 cm for the original tide gauge data, it becomes 16 cm after applying the inverse barometric correction in 1998 (Fig. 2a). The effects of barometric correction are greatest in the Northern Hemisphere winter (January–March) and the Northern Hemisphere fall (October–December) with SSH differences as large as 20 cm (Fig. 2b). Mean NOGAPS sea level pressure at Tofino is 1013.8 mb with a standard deviation of 7.6 mb in 1998 (Fig. 2c). For June–October 1998 SSH after the inverse barometer correction is higher than the one before the barometer correction was applied from June to October because the NOGAPS mean sea level pressure during this time period is greater than the mean pressure (1013.3 mb) used in calculating the correction.

Another example of the inverse barometer correction is shown at Rabaul, New Guinea (2°S, 147°E), in 1998, 2000, and 2001 to illustrate how SSH and the inverse barometer correction change from 1998 to 2001 (Fig. 3). It is noted that for this site the inverse barometer correction usually results in SSH values lower than the original for all years. Standard deviations of the original daily averaged SSH values are 7.7, 5.0, and 6.2 cm in 1998, 2000, and 2001, respectively. However, they are 6.6, 4.6, and 5.7 cm after applying the barometric correction. Similarly, standard deviations of the 30-day running averages of SSH values are 6.8, 4.6, and 5.5 cm, while they are 5.6, 4.2, and 5.2 cm after applying the correction in 1998, 2000, and 2001, respectively. Daily

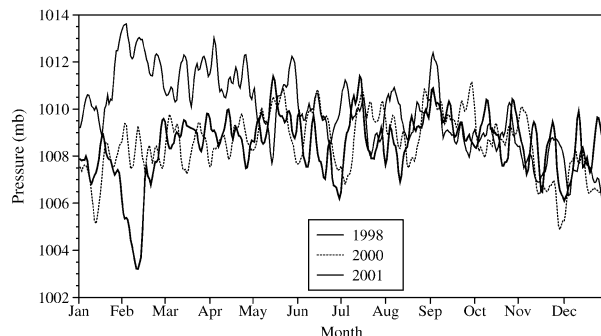


FIG. 4. Daily mean sea level atmospheric pressure values at 2°S, 147°E used for the inverse barometer correction in 1998, 2000, and 2001. Atmospheric pressure values were obtained from NOGAPS. Standard deviations of the atmospheric pressure values are very small, with values of 1.6, 1.3, and 1.3 mb in 1998, 2000, and 2001, respectively.

mean sea level pressure also shows large interannual variability at this location (Fig. 4). This is especially evident from January to May. The annual mean atmospheric sea level pressure values are 1009.9, 1008.6, and 1008.5 mb in 1998, 2000, and 2001.

5. NCOM sea level validation

A yearlong daily SSH time series from each tide gauge location is used in 1998, 1999, 2000, and 2001, separately because our purpose is to validate interannual NCOM simulations. Some tide gauges had large data voids in some years. To avoid extreme cases, we only used series that have at least 100 daily SSH measurements in a given year. Data from the 282 tide stations (see Fig. 1) used in this paper yielded a total of 612 yearlong SSH time series during 1998–2001.

There should be no expectation that the mean would be the same between NCOM SSH and any tide gauge or between any two tide gauges because SSHs are with respect to an essentially arbitrary reference level. In the case of tide gauges it is a measurement, and in the case of NCOM the domainwide mean SSH is zero, that is, free-surface elevation with respect to the ocean basin at rest. Because SSH values from JASL and NCOM have different reference values, the annual mean is removed from both time series before performing model–data comparisons. In the case of locations that have data voids in a given year, the NCOM and observation means were calculated for the days when daily SSH values were available, and the mean was then removed during that time period. Model–data comparisons are performed for each individual year from 1998 to 2001, separately, because our purpose is to measure the performance of NCOM on daily and monthly time scales. No smoothing was applied on sea level data to eliminate some of high frequency effects due to wind and thermal atmospheric forcing used in the model.

a. Statistical metrics

Several statistical metrics are used for comparing the SSH time series predicted from the model (NCOM) with those from the JASL tide gauge data. Let X_i ($i = 1, 2, \dots, n$) be the set of n SSH values from tide gauges (i.e., reference) values, and let Y_i ($i = 1, 2, \dots, n$) be the set of corresponding NCOM estimates. Also, let \bar{X} , \bar{Y} and σ_X , σ_Y be the mean and standard deviations of the reference (estimate) values, respectively. Given this information, the statistical metrics (e.g., Murphy 1988) between the buoy and the model time series can be expressed as follows:

$$R = \frac{1}{n} \sum_{i=1}^n \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \sigma_Y}, \quad (6)$$

$$\text{RMS} = \left[\frac{1}{n} \sum_{i=1}^n (Y_i - X_i)^2 \right]^{1/2}, \quad (7)$$

$$\text{SS} = R^2 - \underbrace{\left[R - \left(\frac{\sigma_Y}{\sigma_X} \right) \right]^2}_{B_{\text{cond}}} - \underbrace{\left[\frac{\bar{Y} - \bar{X}}{\sigma_X} \right]^2}_{B_{\text{uncond}}}, \quad (8)$$

where R is the correlation coefficient, RMS is the rms difference, and SS is the skill score. In our interannual analysis, n is a number between 100 and 366, depending on the availability of daily buoy SSH data for a given year during 1998–2001. Note that n can be as large as 366 only in 2000 (leap year).

It is important to note that we need to examine more than the shape of the SSH seasonal cycle using R ; thus a nondimensional SS is used. Here, SS measures the accuracy of NCOM SSH relative to SSH from tide gauges. The SS is 1.0 for perfect NCOM predictions and can be negative if the NCOM prediction has normalized amplitudes larger than the correlation or large biases in the mean (Murphy and Epstein 1989). An alternative definition for SS, based on rms difference, is $\text{SS} = 1 - \text{RMS}^2/\sigma_X^2$, as given in Murphy and Daan (1985). In our context, SS is calculated based on standard deviations because differences in the mean values have been removed from the data. Thus, the rms difference is used as the basic measure of NCOM accuracy.

As seen from (8), the SS includes conditional (B_{cond}) and unconditional biases (B_{uncond}), as explained in Murphy (1992). The B_{cond} is the bias in standard deviation of the NCOM SSH, while the B_{uncond} reflects the mismatch between the mean NCOM and tide gauge SSH. In our analysis, it should be emphasized that B_{uncond} in (8) is zero because SSH time series from both tide gauges and NCOM are de-meaned. Therefore, SS accounts only for B_{cond} , making σ_X an important quantity in calculations. As noted in section 4, the σ_X is properly calculated after applying the inverse barometer correction to the time series. Based on (6), R^2 is equal to SS only when B_{cond} is zero. The value of R^2 can be considered a measure of “potential” skill, that is, the skill that we

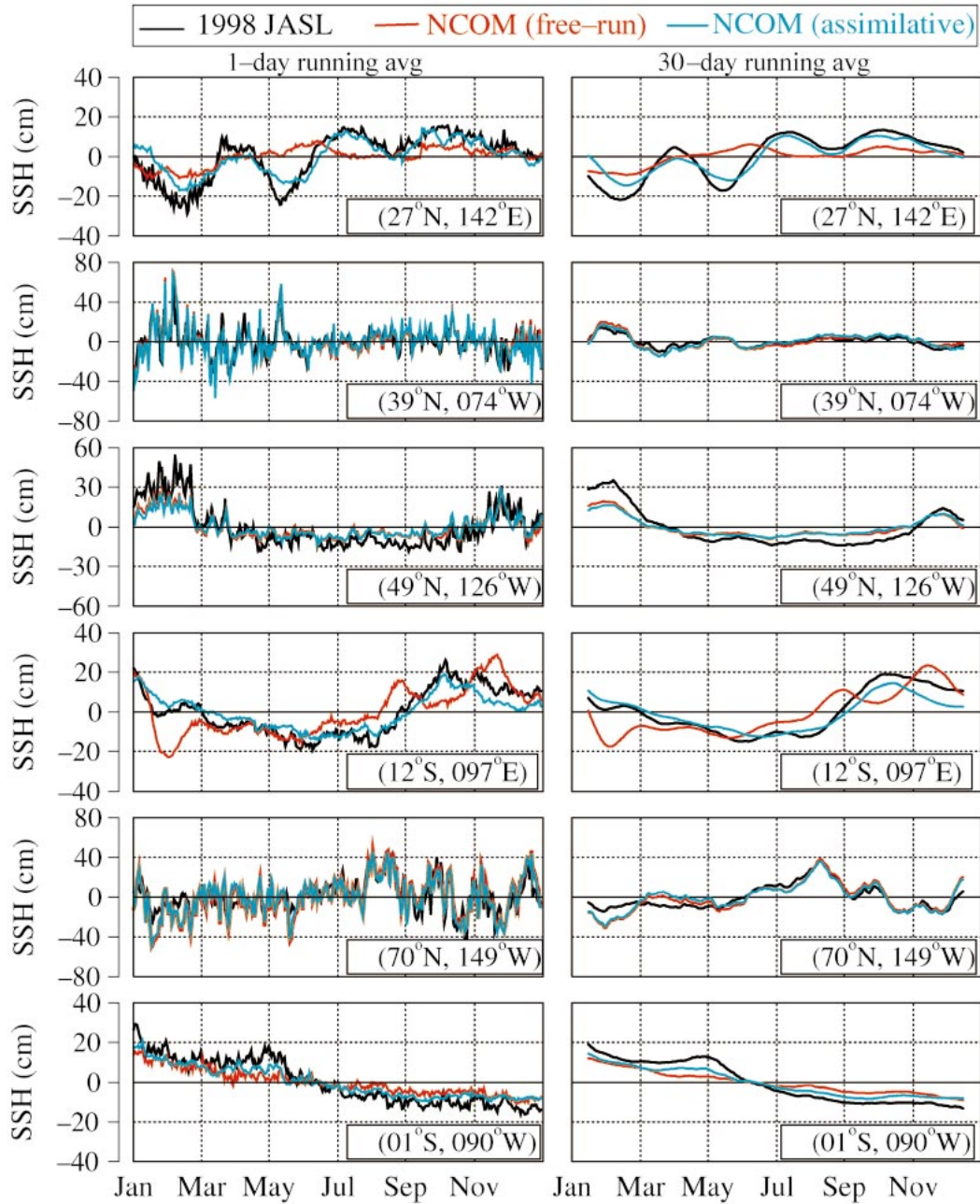


FIG. 5. Daily SSH time series comparisons between the JASL and $\frac{1}{8}^\circ$ NCOM values at various tide gauge locations in 1998. Note that the inverse barometer correction is already applied to SSH at each tide gauge. Model results are shown from free-running and data-assimilative NCOM simulations, separately. The right-hand panels show results with a 30-day running average applied to daily SSH values from tide gauges and NCOM simulations.

can obtain by eliminating bias in standard deviation from the NCOM SSH.

b. Model–data comparisons

In this section, we discuss examples illustrating the model assessment procedure used throughout the paper. Analyses were performed using varying temporal break-

downs of SSH values from $\frac{1}{8}^\circ$ NCOM and from JASL tide gauge locations. These comparisons were based on extractions ranging from daily to multiple years. Model results were analyzed for free-running and data-assimilative simulations. Daily SSH values from the $\frac{1}{8}^\circ$ NCOM were extracted at the tide gauge locations for the model–data comparisons. Figure 5 shows de-meaned daily SSH values from $\frac{1}{8}^\circ$ NCOM (both free-running

TABLE 1. Sample tide gauge locations from JASL. The list includes station name, country, ocean basin, and exact latitude and longitude. JASL receives hourly data from regional and national sea level networks, and the list also provides the contributors for each tide gauge location.

Location	Country	Basin	Lat, lon	Contributor
Chichijima	Japan	Pacific	(27°06'N, 142°11'E)	Japan Meteorological Agency
Atlantic City	United States	Atlantic	(39°21'N, 74°25'W)	U.S. National Ocean Service
Tofino	Canada	Pacific	(49°09'N, 125°55'W)	Canadian Marine Environmental Data Service
Cocos	Australia	Indian	(12°07'S, 96°53'E)	Australian National Tidal Facility
Prudhoe Bay	United States	Pacific	(70°24'N, 148°32'W)	National Ocean Service
Santa Cruz	Ecuador	Pacific	(0°45'S, 90°19'W)	University of Hawaii Sea Level Center

and data-assimilative) and JASL at various locations located near the United States, Canada, Japan, Australia, and Ecuador (Table 1), clearly presenting model success even for the free-running simulations.

Following the daily SSH model–data comparisons, a running-averaged smoothing is applied to reduce high frequency fluctuations in the SSH time series. By using the smoothing technique, the purpose is to filter high frequency fluctuations not resolved by the altimeter data and clarify the longer trends in the time series. As examples, the running average is applied to 1- and 30-day SSH time series obtained from the tide gauges and the

$\frac{1}{8}^\circ$ free-running and data-assimilative NCOM simulations for the locations shown in Fig. 5. Statistical metrics (i.e., RMS, R , and SS) are calculated after removing the mean from the time series (Table 2). Here $\frac{1}{8}^\circ$ NCOM statistics clearly improve after applying the 30-day running average to the time series. An example of the high daily SSH variability is observed at 70°N, 149°W. At this location, the RMS is 11.06 cm and the R value is 0.82 in 1998 when comparing daily SSH time series from JASL and free-running $\frac{1}{8}^\circ$ NCOM. However, the RMS decreases to 4.73 from 11.06 cm ($\sim 70\%$) and the R value increases to 0.90 from 0.82 ($\sim 30\%$) after ap-

TABLE 2. Statistical verification of SSH between the JASL tide gauge data and free-running and data-assimilative $\frac{1}{8}^\circ$ NCOM simulations at a few tide gauge locations in 1998. The reader is referred to the text for a detailed description of each statistical metric. Note that n is the number of days over which the statistics were calculated. An inverse barometer correction was applied to SSH time series from the tide gauges. Standard deviations of SSH time series from the tide gauges and NCOM are denoted as σ_x and σ_y , respectively. It must be emphasized that using only one statistical metric might not be sufficient to assess differences in performance of $\frac{1}{8}^\circ$ NCOM between various tide gauge locations. For example, the rms difference value of 9.95 cm near Atlantic City located at 39°N, 74°W is larger than that of 9.38 cm near Cocos located at 12°S, 97°E when analyzing the free-running $\frac{1}{8}^\circ$ NCOM simulation. This might give the indication that SSH prediction from the model at Atlantic City is worse than at Cocos. However, an examination of the nondimensional SS values reveals a SS value of 0.43 at Atlantic City that is larger than the value of 0.34 at Cocos (11.54 cm), contributing to the relatively large RMS difference at Atlantic City. Thus, the nondimensional SS analysis serves to clarify the distinction between the two cases.

Tide gauge	$\frac{1}{8}^\circ$ NCOM simulation	RMS (cm)	σ_x (cm)	σ_y (cm)	R	SS	n (days)
JASL vs $\frac{1}{8}^\circ$ NCOM: 1-day avg							
27°N, 142°E	Free-run	9.36	11.57	4.70	0.63	0.34	365
	Assimilative	5.65	11.57	8.30	0.89	0.76	365
39°N, 74°W	Free-run	9.95	13.22	15.85	0.78	0.43	365
	Assimilative	10.04	13.22	16.05	0.78	0.42	365
49°N, 126°W	Free-run	8.61	16.00	9.17	0.91	0.71	365
	Assimilative	9.40	16.00	8.29	0.89	0.65	365
12°S, 97°E	Free-run	9.38	11.54	12.01	0.68	0.34	365
	Assimilative	4.99	11.54	8.62	0.92	0.81	365
70°N, 149°W	Free-run	11.06	15.95	19.29	0.82	0.52	365
	Assimilative	10.92	15.95	18.77	0.81	0.53	365
1°S, 90°W	Free-run	5.99	11.17	6.22	0.92	0.71	365
	Assimilative	4.37	11.17	7.68	0.96	0.85	365
JASL vs $\frac{1}{8}^\circ$ NCOM: 30-day avg							
27°N, 142°E	Free-run	8.63	10.88	4.30	0.67	0.37	335
	Assimilative	4.24	10.88	7.82	0.95	0.85	335
39°N, 74°W	Free-run	3.05	5.09	7.02	0.92	0.64	335
	Assimilative	2.76	5.09	6.73	0.93	0.71	335
49°N, 126°W	Free-run	6.59	13.65	7.49	0.97	0.77	335
	Assimilative	7.27	13.65	6.85	0.96	0.72	335
12°S, 97°E	Free-run	8.35	10.98	11.00	0.71	0.42	335
	Assimilative	4.19	10.98	8.09	0.95	0.85	335
70°N, 149°W	Free-run	4.73	9.46	10.72	0.90	0.75	335
	Assimilative	4.82	9.46	9.88	0.88	0.74	335
1°S, 90°W	Free-run	5.26	10.01	5.33	0.95	0.72	335
	Assimilative	3.36	10.01	6.86	0.99	0.89	335

plying the 30-day running average to the time series. This improvement after applying the smoothing is especially evident from the nondimensional SS value. The 30-day running average applied to the time series increases the SS value almost $\sim 50\%$ in comparison to the daily statistics (from 0.52 to 0.75). Thus, NCOM has greater skill in representing the lower frequency components of the SSH.

Analyses mentioned above indicate significant improvement in the model results when using the data-assimilative simulation from $\frac{1}{8}^\circ$ NCOM as opposed to using the free-running simulation from $\frac{1}{8}^\circ$ NCOM (see Table 2). Examples of the importance of SSH prediction are Chichijima, Japan (27°N , 142°E), and Cocos, Australia (12°S , 97°E). After the assimilation, SS increases from 0.34 to 0.76 and RMS decreases from 9.36 to 5.65 cm at Chichijima when comparing daily SSH time series in 1998. Similarly, SS increases from 0.34 to 0.81 and RMS decreases from 9.38 to 4.99 cm at Cocos. An increase in R values is also evident at both tide gauge locations.

Spatial variation of the statistical results is also examined over the global ocean in each year. Figure 6 shows R values calculated between SSH values from the JASL tide gauge data and results from the data-assimilative simulation from $\frac{1}{8}^\circ$ NCOM at all locations in 1998, 1999, 2000, and 2001. A notable improvement in R values is evident after applying a 30-day running average to the daily SSH time series. This is seen at tide gauge stations located near the Japan Sea for all analyzed years. Similarly, a substantial increase in R values is seen at tide gauge locations near the U.S. and Canada coasts. The median rms differences over the years from 1998 to 2001 are ~ 6 cm for 1-day statistics and ~ 3 cm for 30-day statistics. Both free-running and assimilated $\frac{1}{8}^\circ$ NCOM simulations reproduce the correct SSH phase with median $R > 0.70$ in all years. The model success is evident from large and positive SS values that are close to 1, indicating that NCOM is able to simulate SSH with an acceptable accuracy at nearly all locations.

Most of the tide gauges over the open ocean are located between 100°E and 160°W over the global ocean. For these stations, correlations tend to be larger near the equator and smaller near 20°N or 20°S , perhaps reflecting the larger and thus better-resolved equatorial scales. Ocean data assimilation in the model improves the median SS value, evidenced in 2000 when using either 1- or 30-day statistics. The nondimensional SS value is especially useful when evaluating relative model performance at two locations, where one has little seasonal cycle and the other has a large seasonal cycle. Because the standard deviations of SSH from the tide gauges are different, an rms difference might yield misleading results. The model success in predicting SSH can also be seen in nondimensional SS values (Fig. 7) calculated using (8) in section 5a. The stations are plot-

ted in order of decreasing SS, similar to the correlation plots. Any positive SS indicates model skill in predicting SSH, with perfect skill over the year indicated by a SS of one. Positive skill is indicated over most of the global ocean, with negative skill primarily for scattered stations in the central basins or along western boundaries. Skill is particularly high in eastern boundary regions, which may reflect relatively higher dependence on direct local wind forcing.

NCOM performance in predicting daily SSH is now summarized by providing median values of each statistical metric for each year over all stations, rather than providing individual statistical values at each tide gauge location. The same verification procedure as in Table 2 is applied to SSH time series at all tide gauge locations available from 1998 to 2001 to examine overall NCOM performance on daily and monthly time scales. Median RMS, R , and SS values are given Table 3. These were calculated using 1- and 30-day running averages of SSH time series from tide gauges and $\frac{1}{8}^\circ$ NCOM. In the median statistics analysis, there are a total of 189, 181, 151, and 91 tide gauges based on the daily averaged SSH time series and 187, 177, 137, and 90 tide gauges based on the 30-day running-averaged time series in 1998, 1999, 2000, and 2001, respectively.

Finally, overall NCOM performance in predicting SSH is assessed for the 4-yr period during 1998–2001 by combining tide gauge statistics from all years. This yields a total of 612 (591) yearlong SSH time series over 1998–2001 based on the 1-day (30-day running average) statistics. It should be noted that we only use tide gauge locations that have at least 100 daylong SSH time series in a given year. For each set of statistics we also calculate class intervals for RMS R , and SS to examine the distribution of statistical metrics using histograms. As seen from Fig. 8, there are 134 (146) yearlong tide gauge time series that have rms differences between 2 and 3 cm from the free-running (data assimilative) $\frac{1}{8}^\circ$ NCOM simulations when using the 30-day running averages of time series. Most of the R values are > 0.7 , and assimilation generally improves R values. A notable improvement from the data assimilation in predicting SSH is seen again from the SS class intervals. There are 118 out of 591, or $\sim 20\%$, yearlong tide gauge time series where the free-running $\frac{1}{8}^\circ$ NCOM simulations did not yield skillful SSH prediction ($\text{SS} < 0$). With data assimilation this number is reduced to 71 out of 591 cases, or $\sim 12\%$, of unskillful predictions ($\text{SS} < 0$) during 1998–2001.

The statistics in the 4-yr study are also substantially improved by using a 30-day versus a 1-day running average. In part, this is due to the temporal resolution of the altimeter data as evidenced by the larger percentage-wise improvement in rms difference from data assimilation with the 30-day running-averaged time series (Fig. 9). The median RMS is 5.98 cm (3.63 cm) when using daily (30-day running averaged) SSH time series in the free-running model–data comparisons.

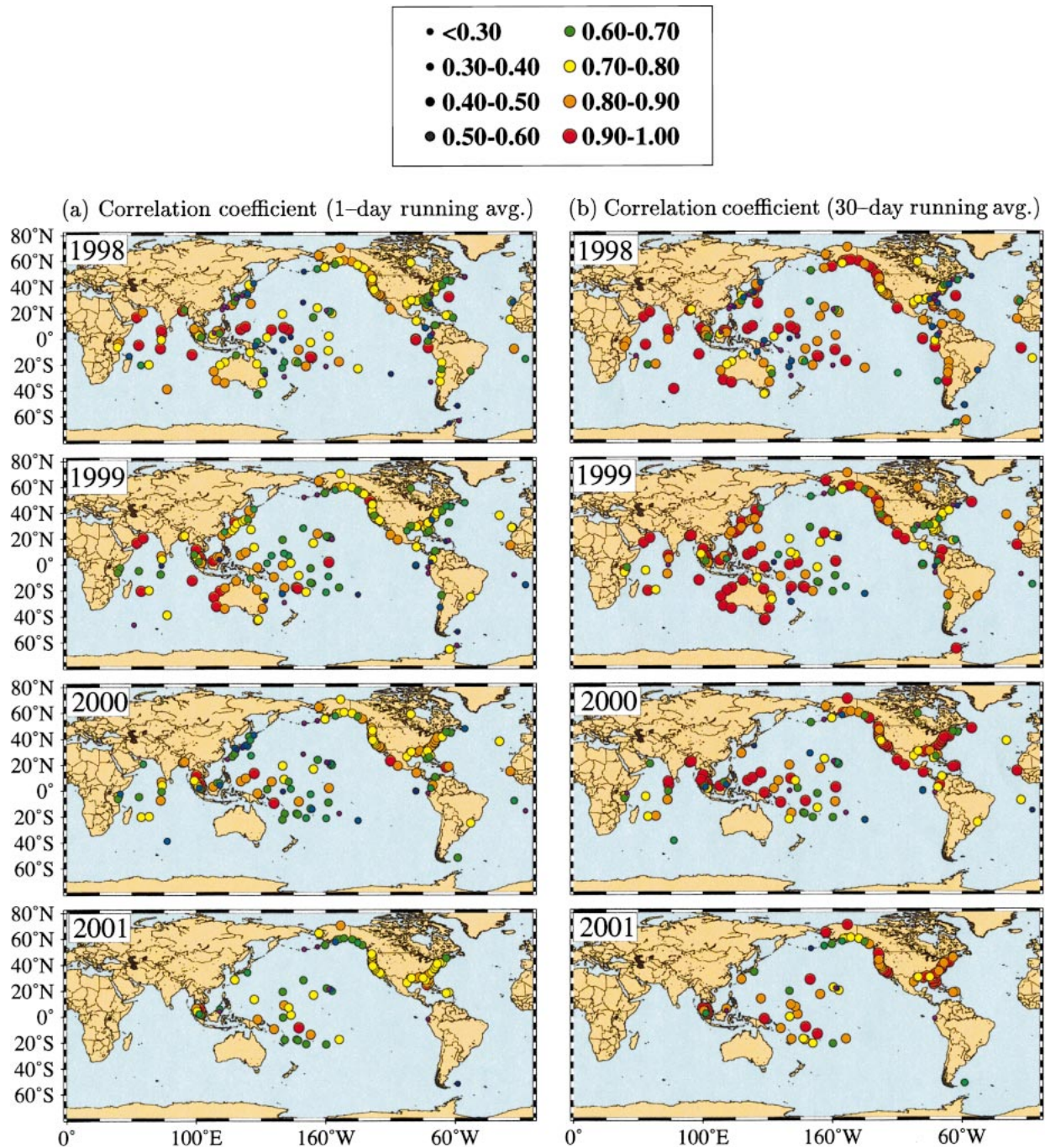


FIG. 6. Correlation coefficients calculated using SSH time series from the JASL tide gauge data and the data-assimilative $\frac{1}{2}^\circ$ NCOM simulation at all the tide gauge locations over the global ocean from 1998 to 2001. Correlation values are calculated based on (a) daily averaged SSH time series and (b) 30-day running averages of SSH time series. Stations are plotted in order of decreasing correlation, so that a station with lower correlation will never be obscured behind a station with higher correlation.

Similarly, the median RMS is 5.77 cm (3.36 cm) when using daily (30-day running averaged) SSH time series in the data-assimilative model–data comparisons. There are approximately 52% (66%) cases when $R > 0.7$ calculated from daily (30-day running averaged)

SSH time series for the free-running model simulations. In the case of the data-assimilative model simulations, $\sim 61\%$ (75%) cases have $R > 0.7$ when using daily (30-day running averaged) SSH time series in calculations.

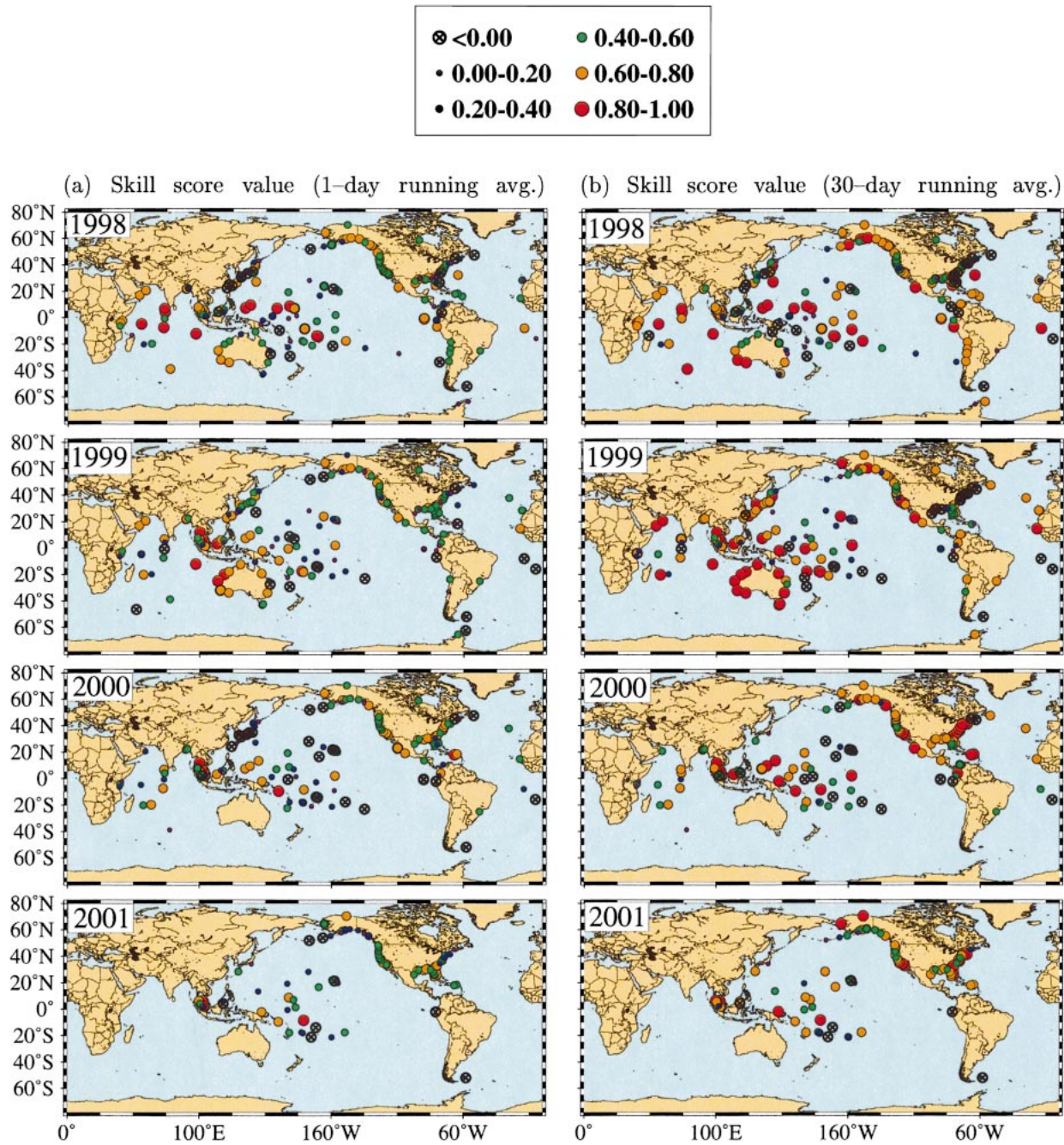


FIG. 7. As in Fig. 6 but for the nondimensional skill score. The nondimensional SS value is especially useful when evaluating relative model performance at two locations, where one has little seasonal cycle and the other has a large seasonal cycle. Because the standard deviations of SSH from the tide gauges could be quite different, using only rms difference might yield misleading results for the model–data comparisons.

c. Sensitivity of model evaluations to the inverse barometer correction

Until the recent atmospheric reanalyses by the European Centre for Medium-Range Weather Forecasts (ECMWF) project (Gibson et al. 1997) and National Centers for Environmental Prediction (NCEP; Kalnay

et al. 1996), accurate multidecadal daily time series of atmospheric sea level pressure that were consistent through time were not available unless there was a reliable weather data recording station nearby. That made correcting tide gauge sea level data for atmospheric pressure loading a challenge. A particular question becomes how would the model–data comparisons change

TABLE 3. Median statistics based on 1- and 30-day running averages of SSH time series from NCOM and demeaned tide gauge sea level data in 1998, 1999, 2000, and 2001, separately. Median statistics are also given for the 4-yr period during 1998–2001. Both data series have had the mean over coincident samples removed. Tide gauge time series have been corrected for the static inverse barometer effect. Results are shown for free-running and data-assimilative $\frac{1}{8}^\circ$ NCOM simulations. Standard deviations of SSH time series from the tide gauge and NCOM are denoted as σ_x and σ_y , respectively.

Year	$\frac{1}{8}^\circ$ NCOM simulation	RMS (cm)	σ_x (cm)	σ_y (cm)	R	SS	No. of buoys
JASL vs $\frac{1}{8}^\circ$ NCOM: 1-day avg							
1998	Free-run	6.33	8.92	6.71	0.73	0.47	189
	Assimilative	6.25	8.92	6.79	0.76	0.51	189
1999	Free-run	6.01	7.97	6.27	0.68	0.39	181
	Assimilative	5.38	7.97	6.33	0.75	0.50	181
2000	Free-run	5.39	7.19	5.81	0.70	0.35	151
	Assimilative	5.21	7.19	5.69	0.74	0.51	151
2001	Free-run	6.64	9.42	6.58	0.72	0.50	91
	Assimilative	6.32	9.42	6.38	0.77	0.53	91
1998–2001	Free-run	5.98	7.98	6.28	0.72	0.45	612
	Assimilative	5.77	7.98	6.22	0.76	0.51	612
JASL vs $\frac{1}{8}^\circ$ NCOM: 30-day avg							
1998	Free-run	3.91	5.96	5.37	0.82	0.61	187
	Assimilative	4.19	5.96	4.93	0.85	0.61	187
1999	Free-run	4.03	5.34	4.56	0.80	0.56	177
	Assimilative	3.12	5.34	4.61	0.85	0.62	177
2000	Free-run	3.15	4.96	5.06	0.85	0.53	137
	Assimilative	2.82	4.96	4.48	0.88	0.65	137
2001	Free-run	3.78	5.69	4.89	0.88	0.65	90
	Assimilative	3.51	5.69	4.73	0.88	0.68	90
1998–2001	Free-run	3.63	5.77	4.67	0.83	0.60	591
	Assimilative	3.36	5.77	4.67	0.85	0.64	591

if one used sea level data from tide gauges without applying barometer correction. In this subsection, we further investigate the impact of the inverse barometer correction on the evaluation of NCOM sea level results.

As seen from the SSH time series in 1998 (Fig. 10), there is no significant difference between the uncorrected and corrected daily sea level values from the Key West, tide gauge (25°N , 82°W). However, the standard deviation of the sea level data is 7.85 cm (7.01 cm) before (after) applying the inverse barometer correction. This is reflected in SS values. For the free-running $\frac{1}{8}^\circ$ NCOM simulations, we obtain a SS value of 0.49 (0.56) before (after) applying the inverse barometer correction. For the data-assimilative $\frac{1}{8}^\circ$ NCOM simulations, a notable improvement is evident from the SS value of 0.49 before to 0.67 after applying the inverse barometer correction. The same is true for the shape of the SSH scatter (Fig. 10). In comparison to an R value of 0.70 calculated using uncorrected sea level data and the data-assimilative $\frac{1}{8}^\circ$ NCOM simulation, we obtain relatively large R value of 0.82 after the inverse barometer correction was applied to sea level data from the tide gauge. The least squares lines in the scatterplots are intended to show slope values. The slope values are 0.61 and 0.51 (0.74 and 0.67) for the free-running and data-assimilative $\frac{1}{8}^\circ$ NCOM simulations before (after) applying the inverse barometer correction, respectively.

Figure 11 shows median RMS, R , and SS values by year to summarize effects of applying barometer correction during 1998–2001. The model–data comparisons were performed using the 30-day running averages

of SSH time series from NCOM and tide gauges. There is usually a modest improvement in median statistics when SSH values from both free-running and assimilative NCOM simulations are compared to sea level data with the inverse barometer correction. The overall drop in the median rms difference is not usually larger than 1 cm.

Based on daily averaged SSH time series (Table 4), a remarkable improvement in median SS and R values and decrease in median RMS values are evident in some years when model–data comparisons are performed after the inverse barometer correction was applied to sea level time series from the tide gauges. For example, when the original sea level time series from tide gauges are used the median SS value is 0.39 between data-assimilative $\frac{1}{8}^\circ$ NCOM and the tide gauges, while the median SS value becomes 0.51 after applying an inverse barometer correction to sea level data in 2000. In general, R and SS increase after applying an inverse barometer correction to sea level data when considering all 612 tide gauges during 1998–2001. This is true for both the free-running and data-assimilative model simulations.

d. Comparisons of spatial SSH variability among various model products

In this section, the spatial SSH variability predicted from NCOM is compared with that interpolated directly from altimeter observations or produced by alternate OGCM outputs to further assess NCOM performance.

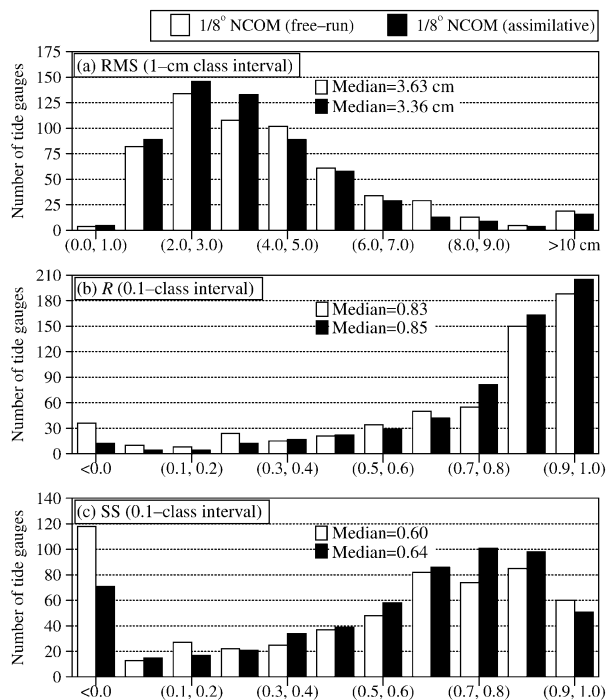


FIG. 8. Histograms of rms difference (RMS), correlation coefficient (R), and nondimensional skill score (SS) used for verifying the free-running and data-assimilative $1/8^\circ$ NCOM simulations during 1998–2001: (a) RMS values with 1-cm class interval, (b) R values with 0.01 class interval, and (c) SS values with 0.1 class interval. All results are based on 30-day running averages of SSH time series. The total number of yearlong tide gauge time series used in the analysis is 591 for the 4-yr time period. Note that an RMS class interval of, e.g., (2.0, 3.0) on the x axis represents the RMS values ≥ 2.0 cm but < 3.0 cm, similarly for R and SS class intervals.

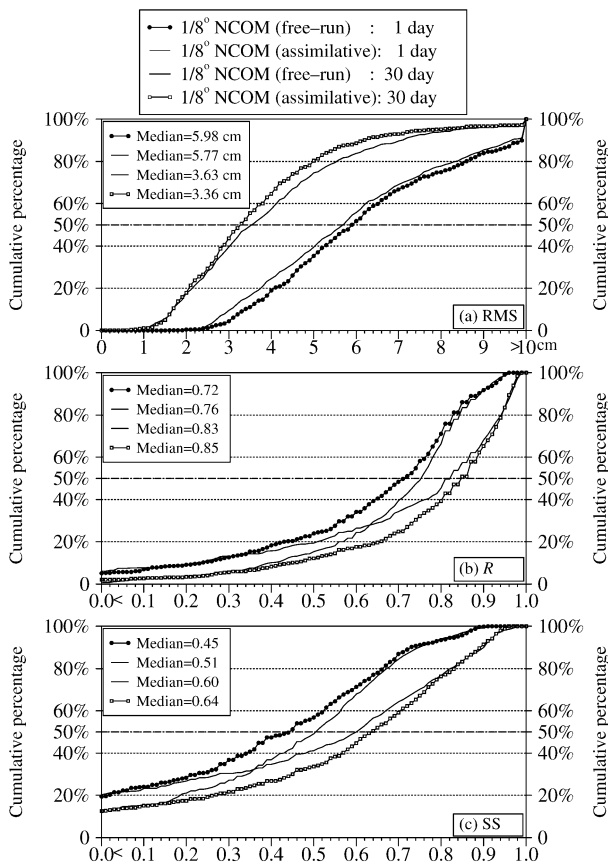


FIG. 9. Cumulative percentage of RMS, R , and SS values calculated using year-long SSH time series from all the tide gauges and $1/8^\circ$ NCOM over the global ocean during 1998–2001. The results are shown for both free-running and data-assimilative $1/8^\circ$ NCOM simulations when applying 1- and 30-day running averages on the time series. Note that an inverse barometer correction has been applied to the tide gauge data before the statistical analysis.

Two regions, the equatorial ocean and the Kuroshio region, are selected for such model intercomparisons. Both regions have been the subject of studies examining SSH variability (e.g., Mitchell et al. 1996; Holland and Mitchum 2003; Qiu 2003). SSH variability in both regions presents challenges for ocean model studies and reflects the influences of local and basin-scale processes.

Since global NCOM in this study does not assimilate SSH directly, there is no strong constraint on the overall level and distribution of SSH variability. As it assimilates temperature and salinity profiles from the MODAS dynamic climatology, profiles themselves derived using SST and SSH, does global NCOM produce realistic standard deviations of SSH? To test this aspect, standard deviations calculated over a common 3-yr interval (1998–2000) are compared. This 3-yr interval is selected to cover most of the tide gauge analysis, excluding 2001, to use an NLOM reanalysis study that ended in 2000. The data are taken from $1/8^\circ$ MODAS 2D SSH, $1/16^\circ$ NLOM SSH, $1/16^\circ$ NLOM baroclinic SSH, and $1/8^\circ$ NCOM SSH. The MODAS 2D SSH is the operational interpolation of the altimeter observations to a spherical grid. NLOM directly assimilates the track-by-track al-

timeter data, and the NLOM SSH is the anomaly of the model from its long-term mean. NLOM baroclinic SSH (SSHB), is equal to NLOM SSH minus the height equivalent to the NLOM pressure anomaly in its deepest layer. Thus, SSHB is representative of a steric height anomaly, the height anomaly that is due to specific volume changes in the water column rather than mass changes, which would cause a bottom pressure signal. It is this SSHB proxy for steric height anomaly that is actually used for the height terms in the MODAS dynamic climatology calculations for assimilation into global NCOM. NCOM SSH is its free surface elevation relative to a global mean of zero.

Figure 12 shows the standard deviations of the various SSH fields in four regions. All present similar amplitudes and patterns. In the equatorial Pacific (Fig. 12a), the 6.6-cm area-averaged SSH standard deviation for NCOM falls between the 6.5 cm for MODAS and 6.8 cm for NLOM, with the baroclinic component of NLOM showing the highest variability at 7.0-cm stan-

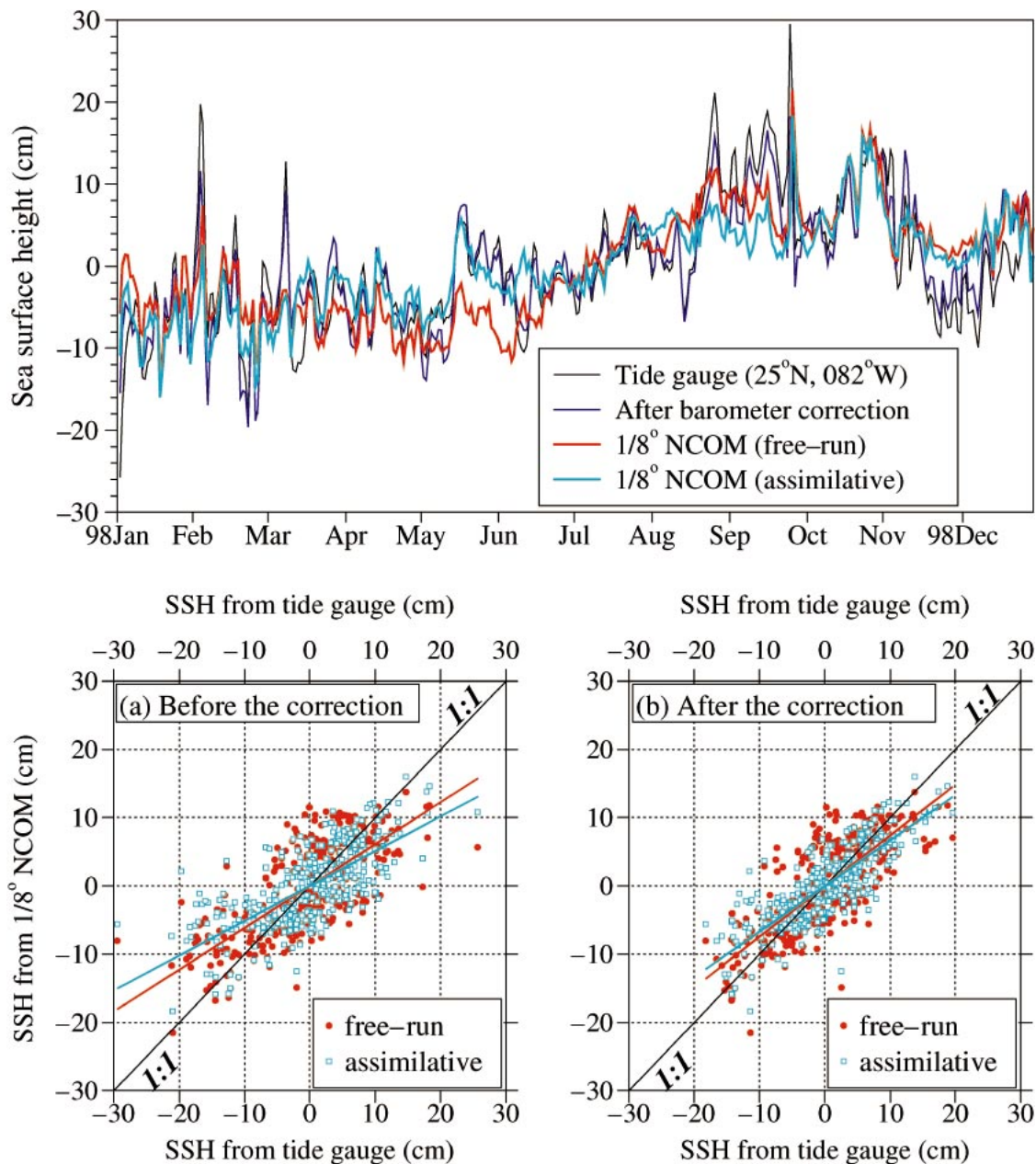


FIG. 10. (top) De-meaned daily SSH time series at Key West, FL, (25°N , 82°W) before/after the inverse barometer correction was applied to the original tide gauge time series in 1998. Also shown are daily SSH values from the free-running and assimilated $1/8^{\circ}$ NCOM simulations. Scatterplots of SSH between tide gauge and the free-running and assimilated $1/8^{\circ}$ NCOM simulations are also shown (a) before the barometer correction was applied to tide gauge time series and (b) after the barometer correction was applied to tide gauge time series.

dard deviation. Some differences arise from shallow regions, which tend to have high SSH variability and are only adequately represented by global NCOM. Altimeter data on the shelf is presently not used by MODAS, and these areas are excluded entirely by the deep water NLOM.

In more open ocean regions, NCOM has somewhat higher variability in the South China Sea and western

boundary currents and lower variability in the less energetic equatorial and eastern Pacific. In a narrower focus to the Japan/East Sea and Kuroshio Extension (Fig. 12b), MODAS (13.0 cm), NLOM (13.1 cm), and NCOM (13.2 cm) estimate similar mean levels of SSH standard deviation, with a slightly smaller baroclinic component in NLOM SSHB (11.9 cm). NCOM variability is a little low in the meander/straight path region south of Japan

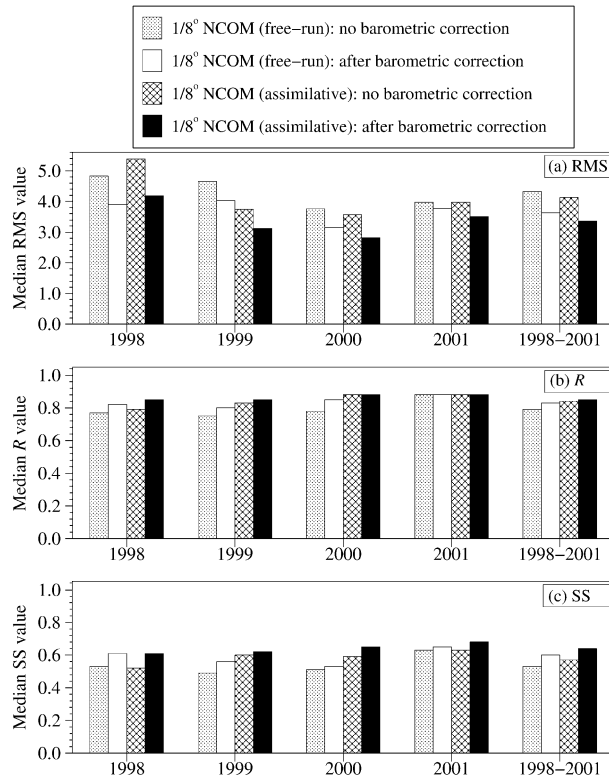


FIG. 11. Median RMS, R , and SS values calculated using 30-day running averages of SSH time series from the JASL tide gauges and data-assimilative $1/8^\circ$ NCOM simulation at all tide gauge locations over the global ocean by year. Also shown is the median value for the 4-yr time period (1998–2001), which is based on 591 yearlong tide gauge time series. Comparisons are shown before and after the inverse barometer correction has been applied to the tide gauge data.

but has a good extension of relatively high variability out to 170°E . Overall the comparisons of SSH variability show good agreement in amplitude and distribution from initial observations to the NCOM model product.

6. Conclusions

In this paper, sea surface height (SSH) values as estimated from free-running and data-assimilative $1/8^\circ$ global Navy Coastal Ocean Model (NCOM) simulations have been compared to detided sea level data measured by a set of in situ tide gauges over the global ocean from 1998 to 2001. There is no SSH assimilated directly, and no tide gauge data enters any part of the assimilation process. SSH from the Naval Research Laboratory Layered Ocean model (NLOM) is indirectly assimilated via synthetic temperature and salinity profiles based on climatological correlations. NLOM assimilates tracks of SSH anomaly data from satellite altimetry.

One potential impediment to comparisons with historical observations is the probability that the historical data will be incomplete, particularly if multiple data types are needed to derive the compared quantity. For this study, we paired observations of detided sea level and atmospheric pressure, the former as archived by the Joint Archive for Sea Level (JASL) center, and the latter interpolated to the tide gauge location from the Navy Operational Global Atmospheric Prediction System (NOGAPS). How would the absence of atmospheric pressure data impact the conclusions of such a study since there are likely historical SSH time series without corresponding records of atmospheric pressure? By evaluation with and without the inverse barometer ad-

TABLE 4. Median statistics based on daily SSH time series from $1/8^\circ$ NCOM and de-meaned tide gauge sea level data in 1998, 1999, 2000, and 2001, separately. Median statistics are shown before and after the inverse barometer correction has been applied to SSH time series from tide gauges. Standard deviations of SSH time series from the tide gauge and NCOM are denoted as σ_x and σ_y , respectively.

Year	Barometric correction	RMS (cm)	σ_x (cm)	σ_y (cm)	R	SS	No. of buoys
JASL data vs free-running NCOM							
1998	No	7.94	10.80	6.71	0.71	0.45	189
	Yes	6.33	8.92	6.71	0.73	0.47	189
1999	No	7.11	9.06	6.27	0.63	0.34	181
	Yes	6.01	7.97	6.27	0.68	0.39	181
2000	No	6.61	8.11	5.81	0.63	0.33	151
	Yes	5.39	7.19	5.81	0.70	0.35	151
2001	No	7.13	10.83	6.58	0.77	0.47	91
	Yes	6.64	9.42	6.58	0.72	0.50	91
1998–2001	No	7.10	9.50	6.28	0.68	0.39	612
	Yes	5.96	7.98	6.28	0.72	0.45	612
JASL data vs data-assimilation NCOM							
1998	No	7.65	10.80	6.79	0.70	0.44	189
	Yes	6.25	8.92	6.79	0.76	0.51	189
1999	No	6.49	9.06	6.33	0.69	0.41	181
	Yes	5.38	7.97	6.33	0.75	0.50	181
2000	No	6.48	8.09	5.69	0.70	0.39	151
	Yes	5.21	7.19	5.69	0.74	0.51	151
2001	No	7.47	11.30	6.38	0.76	0.48	91
	Yes	6.32	9.42	6.38	0.77	0.53	91
1998–2001	No	7.09	9.52	6.22	0.70	0.42	612
	Yes	5.77	7.98	6.22	0.76	0.51	612

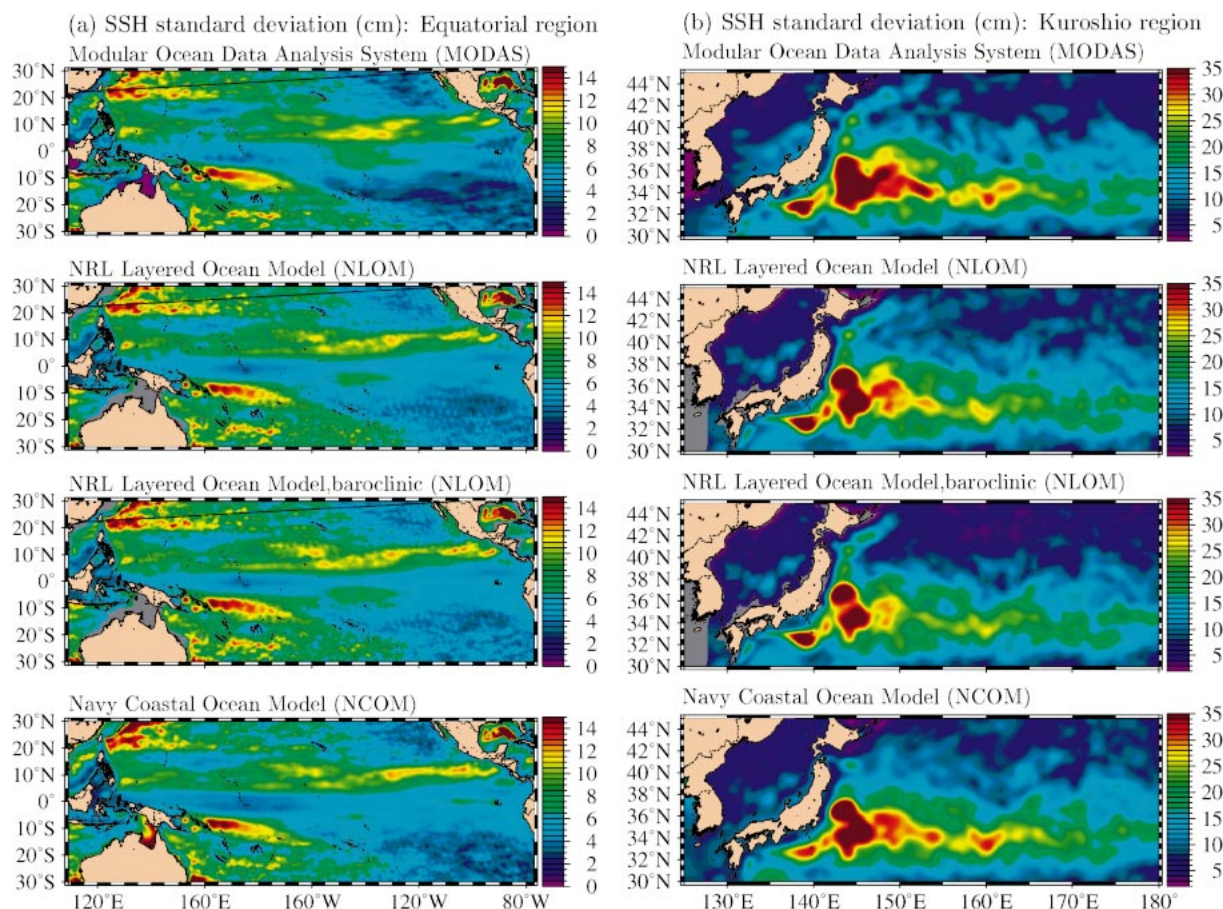


FIG. 12. Comparisons of the 3-yr mean SSH standard deviation calculated from SSH fields at the (a) equatorial ocean and (b) Kuroshio regions during 1998–2000. From the top: $\frac{1}{8}^\circ$ MODAS, $\frac{1}{16}^\circ$ NLOM, $\frac{1}{16}^\circ$ NLOM baroclinic height (NLOM SSH minus the height equivalent to the NLOM pressure anomaly in its deepest layer), and $\frac{1}{8}^\circ$ NCOM. Note that while the NCOM plots show equal coverage to that of MODAS, MODAS SSH is not assimilated directly into NCOM. Thus, these three in-house systems allow for examination and analysis in addition to assimilation.

justment, we find that the impact of the barometric correction is 1.5 cm on the standard deviation of daily mean SSH but only 0.2 cm on the standard deviation of 30-day mean SSH. Thus, the relative impact of the barometer correction on conclusions depends largely on the accuracy of the model results. If the rms difference is much larger than 1.5 cm for daily comparisons or 0.2 cm for monthly comparisons, then the correction would not significantly alter conclusions. In this study, the impact of the inverse barometer adjustment is significant only for the comparisons using the daily time series, where it is similar to the impact of ocean data assimilation. The effects of assimilation are more evident after the barometric correction is applied.

Overall, we find indication that the future operational $\frac{1}{8}^\circ$ global NCOM will have skill over most of the global ocean in representing variations of SSH. The approach of assimilating SSH via the synthetic profiles of temperature and salinity is shown to generally improve skill in predicting SSH. The assimilation did not lead to a loss or unrealistic distribution of SSH variability, as

confirmed by regional comparisons of mean SSH standard deviation values in the equatorial ocean and Kuroshio region. Some regions with relatively low skill, such as 20°S in the central or western Pacific Ocean, and even the equatorial Pacific may be the product of relative deficiencies in the climatological database as much as relative inadequacies in methodology. Data coverage in the equatorial regions has increased greatly since the development of statistics relating temperature $[T(z)]$ and salinity $[S(z)]$ to SSH and sea surface temperature (SST) using the MODAS 2.1 hydrographic climatological database. The moored equatorial arrays [Tropical Atmosphere Ocean (TAO) and Pilot Research Moored Array (PIRATA)] and the rapidly growing ARGO float database of quality $T(z)$ and $S(z)$ profiles to ~ 1500 m provide a potential for significant improvement in formerly data-sparse regions. Additional improvements to the assimilation methodology should also improve model skill in SSH predictions. However, some of the discrepancy between model and tide gauge is doubtless due to the local configuration and dynamics

in the vicinity of the tide gauge and factors requiring finer resolution than $\frac{1}{8}^\circ$ global NCOM ($\sim 14\text{--}16$ km at midlatitudes). Nested models with increasingly finer scale and more detailed forcing are needed to resolve the processes active within nearshore regions. Global NCOM is designed to serve as a bridge between the open ocean and shallow water regimes, providing skill over the global ocean sufficient to support models appropriate for a given application. This paper provides an assessment of that skill for SSH and indications of how that skill may be similarly quantified for other OGCMs. Operational NCOM is available online (http://www.ocean.nrlssc.navy.mil/global_ncom) for real-time predictions.

Acknowledgments. The numerical simulations were performed under the Department of Defense High Performance Computing Modernization Program, on an IBM SP3 at the Naval Oceanographic Office, Stennis Space Center, Mississippi. We would like to extend our special thanks to T. Townsend and O. M. Smedstad for their help in data processing. Tide gauge data used in this paper were obtained from the Joint Archive for Sea Level (JASL) project which is funded by the National Ocean Service of the National Oceanic and Atmospheric Administration (NOAA). Additional thanks go to P. Caldwell of JASL for numerous discussions regarding the tide gauge dataset. This work was funded as part of the NRL 6.4 Large-scale Models and the 6.4 Ocean Data Assimilation projects, managed by the Space and Naval Warfare Systems Command PMW-155 under Program Element 603207N. It received additional funding from the 6.1 Dynamics of Low latitude Western Boundary Currents project funded by the Office of Naval Research (ONR) under Program Element 601153N.

REFERENCES

- Blumberg, A. F., and G. L. Mellor, 1987: A description of a three-dimensional coastal ocean circulation model. *Three-Dimensional Coastal Ocean Models*, N. Heaps, Ed., Amer. Geophys. Union, 208 pp.
- , and H. Herring, 1987: Circulation modelling using orthogonal curvilinear coordinates. *Three-Dimensional Models of Marine and Estuarine Dynamics*, J. Nihoul and B. Jamart, Eds., Oceanography Series, Vol. 45, Elsevier, 55–88.
- Caldwell, P., 1992: Building an archive of tropical sea level data. *Earth Syst. Monitor*, **3**, 3–6.
- , and M. Merrifield, 2000: Joint Archive for Sea Level annual data report: November 2000. JIMAR Contribution No. 335, Data Rep. 18, SOEST, University of Hawaii, 48 pp.
- Fox, D. N., C. N. Barron, M. R. Carnes, M. Booda, G. Peggion, and J. V. Gurley, 2002a: The Modular Ocean Data Analysis System. *Oceanography*, **15**, 22–28.
- , W. J. Teague, C. N. Barron, M. R. Carnes, and C. M. Lee, 2002b: The Modular Ocean Data Analysis System (MODAS). *J. Atmos. Oceanic Technol.*, **19**, 240–252.
- Gibson, J. K., P. Källberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA description. ECMWF Re-Analysis Project Report Series, No. 1, 72 pp. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom.]
- Gill, A. E., 1982: *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.
- Holland, C. L., and G. T. Mitchum, 2003: Interannual volume variability in the tropical Pacific. *J. Geophys. Res.*, **108**, 3369, doi:10.1029/2003JC001835.
- Jacobs, G. A., C. N. Barron, D. N. Fox, K. R. Whitmer, S. Klingenberg, D. May, and J. P. Blaha, 2002: Operational altimeter sea level products. *Oceanography*, **15**, 13–21.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kara, A. B., H. E. Hurlburt, and P. A. Rochford, 2002: Air–sea flux estimates and the 1997–1998 ENSO event. *Bound.-Layer Meteor.*, **103**, 439–458.
- , A. J. Wallcraft, and H. E. Hurlburt, 2003: Climatological SST and MLD predictions from a global layered ocean model with an embedded mixed layer. *J. Atmos. Oceanic Technol.*, **20**, 1616–1632.
- , H. E. Hurlburt, P. A. Rochford, and J. J. O’Brien, 2004: The impact of water turbidity on the interannual sea surface temperature simulations in a layered global ocean model. *J. Phys. Oceanogr.*, **34**, 345–359.
- Martin, P. J., 2000: Description of the Navy Coastal Ocean Model Version 1.0. NRL Rep. NRL/FR/7322/00/9962, 45 pp. [Available from NRL, Code 7322, Bldg. 1009, Stennis Space Center, MS 39529-5004.]
- , G. Peggion, and K. J. Yip, 1998: A comparison of several coastal ocean models. NRL Rep. NRL/FR/7322/97/9692, 96 pp. [Available from NRL, Code 7322, Bldg. 1009, Stennis Space Center, MS 39529-5004.]
- Mitchell, J. L., W. J. Teague, G. A. Jacobs, and H. E. Hurlburt, 1996: Kuroshio Extension dynamics from satellite altimetry and a model simulation. *J. Geophys. Res.*, **101**, 1045–1058.
- Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417–2424.
- , 1992: Climatology, persistence, and their linear combination as standards of reference in skill scores. *Wea. Forecasting*, **7**, 692–698.
- , and H. Daan, 1985: Forecast evaluation. *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, A. H. Murphy and R. W. Katz, Eds., Westview Press, 379–437.
- , and E. S. Epstein, 1989: Skill scores and correlation coefficients in model verification. *Mon. Wea. Rev.*, **117**, 572–581.
- Qiu, B., 2003: Kuroshio Extension variability and forcing of the Pacific decadal oscillations: Responses and potential feedback. *J. Phys. Oceanogr.*, **14**, 104–113.
- Rosmond, T. E., J. Teixeira, M. Peng, T. F. Hogan, and R. Pauley, 2002: Navy Operational Global Atmospheric Prediction System (NO-GAPS): Forcing for ocean models. *Oceanography*, **15**, 99–108.
- Smedstad, O. M., H. E. Hurlburt, E. J. Metzger, R. C. Rhodes, J. F. Shriver, A. J. Wallcraft, and A. B. Kara, 2003: An operational eddy-resolving $1/16^\circ$ global ocean nowcast/forecast system. *J. Mar. Syst.*, **40–41**, 341–361.
- Wallcraft, A. J., A. B. Kara, H. E. Hurlburt, and P. A. Rochford, 2003: The NRL Layered Global Ocean Model (NLOM) with an embedded mixed layer sub-model: Formulation and tuning. *J. Atmos. Oceanic Technol.*, **20**, 1601–1615.