SST Assimilation and Assessment of the Predicted Temperature from the GoMOOS Nowcast/Forecast System

Huijie Xue*
Lei Shi
Stephen Cousins
School of Marine Sciences, University of Maine

Abstract

A circulation nowcast/forecast system was developed for the Gulf of Maine as an integral component of the Gulf of Maine Ocean Observing System (GoMOOS) technical program. The system has been used to generate nowcasts and short-term forecasts of the circulation and physical properties in the Gulf of Maine. One of the expectations is that the model can provide consistent SST to fill in AVHRR gaps and eventually produce reliable 3D temperature fields for fishery applications.

This paper first presents the framework of the nowcast/forecast system, which includes an algorithm to assimilate satellite derived SST. It then discusses the performance of the system be comparing the predicted and the observed temperatures (both *in situ* and satellite derived). In general, the assimilation algorithm is stable and produces SST patterns mimicking the AVHRR. However, the modeled temperature appears to be colder than the *in situ* temperature at buoy locations. Elimination of the exceptionally cold satellite SST at cloud edges before it is assimilated corrects for a large portion of the cold bias in the model. Furthermore, seasonal variation in temperature is well reproduced and the predicted synoptic temperature variation is significantly correlated with its counterpart from the mooring measured temperature.

1. Introduction

The Gulf of Maine is a productive, marginal sea (Figure 1). Its exchange with the open ocean is largely controlled by the geometry of banks and channels that characterize its open boundary. Inside the Gulf there are three basins separated at the 200 m depth, namely the Jordan Basin and the Georges Basin in the Eastern Gulf of Maine and the Wilkinson Basin in the Western Gulf of Maine. A circulation schematic based on the satellite-tracked drifters and hydrographic observations delineates two distinct cyclonic gyres centered over the two basins in the eastern Gulf.

* Corresponding author address: School of Marine Sciences, University of Maine, Orono, ME 04469-5741, USA
hxue@maine.edu, (207)581-4318
of Maine and a complex and well-developed cyclonic coastal current system encompassing the gyre pair. It has been documented repeatedly that the gulf-wide circulation is the strongest and most coherent in the summer, but lacks a recognizable pattern in the winter (e.g., Bumpus and Lauzier, 1965, Vermersch et al., 1979, Brown and Irish, 1992), most likely related to the evolving density structure as suggested by Brooks and Townsend (1989) and Brown and Irish (1992).

Processes that influence the density distribution inside the Gulf of Maine include surface heat flux, tidal mixing, river runoff, and the inflow of the Scotian Shelf Water (SSW) and the slope water. For example, Pettigrew et al. (1998) found that the thermohaline structure in the eastern Gulf of Maine was substantially modified by the cold and less saline SSW. Brooks (1990) observed that the slope water spread over Lindenkohl sill and moved from Georges Basin towards northwest. Xue et al. (2000) suggested that the surface heat flux plays an important role in regulating the annual cycle of the circulation in the Gulf of Maine by eroding the stratification in the upper water column due to winter cooling and reestablishing the stratification due to summer warming. The Gulf of Maine and Bay of Fundy system is well known for its nearly resonant semi-diurnal tidal responses and vigorous tidal stirring keeps the water vertically well mixed over Georges Bank, the western shelf of Nova Scotia (Loder and Greenberg 1986), and the eastern Maine coast (Pettigrew et al. 1998).

Figure 1. Topographic features of the Gulf of Maine region. GB – Georges Basin, JB – Jordan Basin, WB – Wilkinson Basin, NEC – the Northeast Channel, GSC – the Great South Channel. Also marked in asterisks are the approximate locations of GoMOOS buoys.
The GoMOOS circulation nowcast/forecast system is developed from the comprehensive Gulf of Maine circulation model of Xue et al. (2000). The goal is to establish an operational numerical prediction system for the Gulf of Maine, which produces forecasts of the ocean condition by coupling with available meteorological forecasts and makes them available via the web in real time. One of the expectations is that the model, in a long run, can provide short-term and long-term, three-dimensional temperature variations, which are known to affect the fisheries in the Gulf of Maine. For example, a correlation between warm temperature anomalies and lobster larval settlement has been found in the Penobscot Bay region.

Skillful forecasts are made often by using data to constrain the model in the assimilative manner. The Gulf of Maine circulation nowcast/forecast system now includes assimilation of the satellite SST, and it will also include in the future assimilation of the CODAR measured sea surface velocity, which is now under extensive tests and evaluation. In this paper, only the SST assimilation is discussed. The following section describes the current version of the GoMOOS circulation nowcast/forecast system. Section 3 discusses a SST assimilation scheme and comparisons of predicted and observed temperatures in the upper ocean of the GOM. Section 4 includes a summary of the operational system and a brief outline of the ongoing and planned activities.

2. The Gulf of Maine circulation nowcast/forecast system

The Gulf of Maine circulation nowcast/forecast system is based on the three-dimensional Princeton Ocean Model in a curvilinear grid (Figure 2). It is driven at the surface by heat, moisture, and momentum fluxes from the National Center for Environmental Prediction (NCEP)’s Eta mesoscale atmospheric forecast model. Boundary forcing includes daily river outflows from St. John, Penobscot, Kennebec, Androscoggin, Saco, and Merrimack, tidal (M$_2$, S$_2$, N$_2$, K$_1$, O$_1$, and P$_1$) and subtidal forcing from the open ocean, which is interpolated from the daily nowcast of the NCEP Regional Ocean Forecast System (ROFS).
The daily procedure, which starts at 6:00 am everyday, includes three consecutive jobs: preprocessing, model integration, and post processing. Preprocessing consists of a series of automated scripts, which download the river discharge, AVHRR, Eta and ROFS forecasts, and interpolate them to the Gulf of Maine grid. Handling missing data is a critical step. For short-term disruption of river discharge data, the last valid number is carried forward. Climatological monthly mean is used during extensive ice period during winter. For AVHRR, a composite of the last eight days is formed to minimize the cloud cover and, if there are still clouds, an 8-day climatology from the same Julian days is used to fill in the cloudy spots. Missing ROFS data is usually substituted using the ROFS output of the last available date with tidal correction, which work well for interruptions up to a week. Similarly, the last available Eta forecast is used to substitute for missing Eta data. This, however, can result in considerable errors.

Upon the completion of preprocessing, which supplies necessary boundary conditions, the automated daily procedure calls for the model integration session (figure 3). A 24-hour nowcast cycle assimilates the satellite SST (see details in section 3) and prepares the initial condition for the next 48-hour forecast.

Post-processing includes data storage, web interface, and error assessment (see details below). Model output of every 3 hours has been archived since 1 January 2001, and daily restart files have been saved incrementally for hindcasts of specific events. Temperature, salinity and velocity at three levels and surface elevation are shown graphically on the web at a 3-hour interval (http://gomoos.org). The GoMOOS circulation forecast system is currently running on both the SGI Origin 3200 and a dual processor PC. The daily procedure takes about 20 minutes on the SGI and about 2 hours on the PC. The system is robust, seldom needs human intervention over the last three years other than handling missing forcing data.
Understanding forecast errors is an integral component of the nowcast/forecast system. The 1st level of examination includes the comparison of the modeled and in situ temperature and velocity at GoMOOS buoy sites (see Figure 1 for buoy locations). In particular, an 8-day running mean is applied to the temperature time series to separate the synoptic and seasonal variations (see section 3b for details), while harmonic analysis is applied to the velocity time series to separate the tidal current from the residual flows. The quality of predicted tidal current varies from buoy to buoy. In general, the model performs better at three shelf buoys (B, E, and I) but rather poorly at buoy J, which is located in Cobscook Bay where the model does not have adequate resolution. Figure 4 compares the modeled and the observed surface velocity at buoy E and I. The modeled velocity has similar magnitude and variance as the observed velocity. Furthermore, the model appears to catch the timing of seasonal transition of the coastal current. The 2nd level of Quality Assessment (QA) subsystem is currently in the development, which includes the analysis of bias, variances, and spectral properties.

Figure 4. Comparisons of tidal residual velocity at buoy E and I.

3. SST assimilation

In order for a model to produce SST correctly, it is essential to have exact heat balance at the surface. The Eta heat fluxes appear to contain significant errors as
suggested by the discrepancy between the observed and the predicted air temperature as seen in Figure 5. As a result, the modeled SST is considerably lower in the summer months and does not show the typical east-west contrast often seen in the AVHRR (Figure 6a and 6b). Furthermore, heat fluxes affect not only the SST in the model, but also the subsurface temperature and the circulation, especially when and where strong mixing occurs. Secondly, since there are periods of extensive cloud covers especially during winter, satellite AVHRR data contains wide gaps. It is expected that the model can be used as a dynamically consistent interpreter to fill in the data gaps.

Figure 6. Without SST assimilation, the modeled sea surface temperature (a) appears to be colder and fails in reproducing the typical summertime SST pattern, which has higher temperature in the western Gulf of Maine as seen in the AVHRR (b).

To assimilate satellite AVHRR data into the model, an algorithm similar to that of Kelley et al. (1999) has been incorporated into the GoMOOS nowcast/forecast system since June 2001. The algorithm, chosen mainly because it is very efficient computationally, consists of three elements: an optimal interpolation scheme (Derber
and Rosati, 1989; Behringer *et al.*, 1998); a mixed-layer adjustment (Chalikov *et al.*, 1996); and a Newtonian nudging.

3a. Assimilation scheme

The first task is to produce observed daily temperature field. There are 4 passes per day including both ascending and descending tracks of n12 and n14. Data from all 4 passes in the past 8 days are combined to form a composite. It is not unusual that there is still considerable amount of cloud cover in the composite. Two different approaches have been used to fill the sea surface temperatures over the cloud-covered areas as seen in figure 7. One is to use the climatological 8-day composite, another is to use the modeled SST. The latter is effectively without data assimilation. Upon the deriving of the observation field (T_o), the optimal interpolation procedure is to determine a correction field for the model’s top layer temperature by simultaneously minimizing two difference fields. One is the difference between the corrected temperature field (T_c) and the model temperature (T_m), and the other is between the corrected temperature field and the observed temperature (T_o).

\[ I = (T_c - T_m)^T R(T_c - T_m) + (\nabla T_c - T_o)^T \Imath(\nabla T_c - T_o) \]  

(1)

where the superscript, t, denotes the transpose matrix. \( R \) and \( \Imath \) are the error covariance matrices for the model data and the observations, respectively. Following Kelley et al. (1999), \( \Imath \) is a diagonal matrix (assuming the observational error is uncorrelated). The value of the diagonal elements of \( \Imath \) is the observational error variance. A time factor is incorporated into \( \Imath \) such that more weight is placed on the observations with time closer to the assimilation date. The model error covariance is approximated, for any two grid-points, as \( R_{ij} = a \exp[-(r_{ij}/b)^2] \), where \( r_{ij} \) is the horizontal distance between two model grid points. \( a \) is the first guess error variance and \( b \) is the estimate of the correlation spatial scale of the model error, and they are set to 0.50 \( ^\circ \text{C}^2 \) and 60 km, respectively. \( \nabla \) is the transformation matrix that converts values at the model grid points to the observation locations. Minimization is achieved...
by using a preconditioned conjugate gradient algorithm (Gill et al., 1981, Golub and Van Loan, 1989), which finds the solution iteratively. Figure 8a is the resultant correction field corresponding to the composite shown in Figure 7.

SST, however, is not a strong dynamic constraint. Temperature below the surface needs to be adjusted accordingly. A simple procedure is currently employed in the GoMOOS nowcast/forecast system such that when the corrected surface temperature is warmer than the model surface temperature, the correction is distributed throughout the mixed layer, while when the corrected temperature is colder, the corrected temperature replaces the model temperatures down to the depth where they become equal. Newtonian nudging is used to slowly apply the correction field to the model temperature over the nowcast cycle. Figure 8b is the SST resulted from the nowcast cycle using the correction field shown in Figure 8a, which serves as the initial SST in the 9 June 2002 integration.

3b. The modeled temperature

The algorithm is steady and has produced robust sea surface temperature patterns since it was implemented. As seen from Figure 8b, the typical summertime east-west contrast in the sea surface temperature is well simulated, as well the cold temperature on Georges Bank. However, comparisons with in situ temperature show that the modeled temperature is colder (Figure 9). This is probably due to the artifact of the exceptionally cold temperature at cloud edges from the satellite SST (see Figure 7) and the assimilation scheme especially the mixed layer adjustment. Note also in this experiment the cloud-covered areas was filled in using the climatological 8-day composite derived from 15 years of AVHRR, which was found later to have a cold bias of ~2°C in the winter. Furthermore, vertical adjustment of the temperature was limited to the top 200m. Several experiments were carried out to test various aspects.
Figure 9. Time series of the modeled (dark curve), in situ (light curve), and satellite (Scattered dots) sea surface temperature at the GoMOOS buoys A, B, C, E, F, I, J, and M.
Figure 10. Similar to Figure 9 but for experiment 2. See text for differences between experiment 1 and 2.
Figure 11. Time series of the modeled and in situ temperature at 50m at the GoMOOS buoys B, E, I, and M. Left panels are for experiment 1, and right panels experiment 2. See text for differences between these two experiments.
of the assimilation scheme. Figure 10 shows similar comparisons of the sea surface temperature for the second experiment in which, 1) the SST correction was limited to 2°C so that the exceptionally cold temperature (artifact of cloud shadows) was excluded; 2) the modeled SST instead of the climatological satellite 8-day composite was used for cloud-covered area in the daily composite, and 3) the vertical adjustment is limited to the top 50m. It is clear that the modeled temperature matches much better with the observed temperature at all buoys. Figure 11 shows the comparisons of the temperature at 50m for both experiments at the buoys with at least 3 months of data in 2002. Although the 50m temperature is too warm at buoy B during summer months, the modeled temperature appears to agree better with the observed temperature at 50m in the second experiment even for buoy B, suggesting that this experiment produces better-fit vertical temperature profiles as well.

Obviously, the largest variability in temperature is the annual cycle, which is then separated from the time series using a simple 8-day running mean. The low-pass filtered data (time series – its 8-day running mean) consists of higher frequency variations including those in the weather band and those with longer periods of several weeks. An example of this remnant time series is shown in Figure 12. Most of the events appear to be well reproduced in the model. Another way to measure the ability of the model in predicting the seasonal and synoptic variations in the temperature is the correlation coefficient (Table 1). For the annual cycle, $R^2$ is greater than 0.90 between the model temperature and the in situ temperature for all buoys except M. Note that the observed record at buoy M is significantly shorter. $R^2$ in the synoptic band ranges from 0.33 to 0.70, again with the exception of buoy M at 30m where $R^2$ was 0.0606. This low in correlation is due to large amplitude high frequency variations in the observed temperature at this location (see lower right panel in Figure 11).

Spatially, the correlation is higher at the 4 buoys located in the western Gulf of Maine (namely, A, B, C, and E), having to do with the eta model assimilating more information from the land-based weather stations, which is also probably why correlation is higher again at buoy I since it is near the C-MAN on Mount Desert Rock. Correlation is lower in the bays (F and J), which is also consistent with the
generally lower correlations for eta predictions in the bays. Finally, the correlation between the modeled temperature and the observed temperature decreases from the surface downward.

In general, the annual cycle in temperature is well reproduced with \( R^2 \) well above the 95% significance level. The correlation in the synoptic band is much lower for several reasons. First of all, the correlation between the buoy and the satellite temperature in the same frequency band is in the range of 0.3 to 0.4. Secondly, a daily composite of the satellite SST is formed and then assimilated in the model. As a result, the diurnal variation is greatly suppressed in the model. The objective correlation should thus be between the 24-hour lowpass filtered series. Thirdly, \( R^2 \) between the Eta forecasted and buoy measured wind and air temperature varies from 0.45 to 0.8 with the higher correlation attained in winter. This is probably due to the fact that northwesterly and northeasterly winds dominate in winter and more dense coverage of weather stations on the land supply enough observations to the Eta model and it hence produces more reliable forecasts. In spring and summer, however, a large percentage of the weather systems that affect the Gulf of Maine come from south and the small number of weather buoys in the sea provides only limited information to the Eta model, which results in less satisfactory meteorological forecast. The disassociation between the AVHRR and the in situ SST and the disassociation between the Eta predicted wind and the observed wind both affect the quality of model predicted sea surface temperature, much so in the synoptic frequency band. Nevertheless, the correlation coefficients are above the 95% significance level (except the 30m depth at buoy M) as the time series contain about 2900 data points with degree of freedom over 100, suggesting that the model has certain ability in predicting temperature variability in the synoptic frequency band.

### 4. Summary

The GoMOOS nowcast/forecast system based on the 3-dimensional Princeton Ocean Model generates daily and short-term forecasts of the circulation and physical properties in the Gulf of Maine. With the realtime meteorological forcing and river discharge, as well as the model (i.e., ROFS) predicted open ocean boundary

<table>
<thead>
<tr>
<th>Buoy</th>
<th>BUOY LOCATION</th>
<th>( R^2 )-seasonal</th>
<th>( R^2 )-synoptic band</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buoy A</td>
<td>1m</td>
<td>0.9824</td>
<td>0.6339</td>
</tr>
<tr>
<td>Buoy B</td>
<td>1m</td>
<td>0.9847</td>
<td>0.6984</td>
</tr>
<tr>
<td>Buoy C</td>
<td>1m</td>
<td>0.9698</td>
<td>0.6444</td>
</tr>
<tr>
<td>Buoy E</td>
<td>1m</td>
<td>0.9710</td>
<td>0.5168</td>
</tr>
<tr>
<td>Buoy F</td>
<td>1m</td>
<td>0.9859</td>
<td>0.3780</td>
</tr>
<tr>
<td>Buoy I</td>
<td>1m</td>
<td>0.9856</td>
<td>0.4278</td>
</tr>
<tr>
<td>Buoy J</td>
<td>1m</td>
<td>0.979</td>
<td>0.2951</td>
</tr>
<tr>
<td>Buoy M</td>
<td>1m</td>
<td>0.8593</td>
<td>0.4638</td>
</tr>
<tr>
<td>Buoy B</td>
<td>50m</td>
<td>0.9039</td>
<td>0.3274</td>
</tr>
<tr>
<td>Buoy E</td>
<td>50m</td>
<td>0.9615</td>
<td>0.3722</td>
</tr>
<tr>
<td>Buoy I</td>
<td>50m</td>
<td>0.9885</td>
<td>0.4794</td>
</tr>
<tr>
<td>Buoy M</td>
<td>30m</td>
<td>0.8054</td>
<td>0.0606</td>
</tr>
</tbody>
</table>
condition, the nowcast/forecast system is able to produce realtime, 3-dimension distributions of the circulation and water properties for the Gulf of Maine, Bay of Fundy and the adjoining Georges Bank region. Preliminary analyses of the results in 2002 indicate that the model produces realistic seasonal variations of the surface temperature and the coastal current. In the synoptic frequency band, the correlation coefficients between the modeled temperature and the in situ temperature are higher in the western Gulf of Maine than in the eastern Gulf of Maine, higher in the shelf than in the bays. More importantly, they are above the 95% significance level and generally comparable or higher than the correlation coefficients between the satellite SST and the in situ SST at the same location. As such, it is concluded that the model generated surface temperature can be used to fill in the data gaps of the satellite derived sea surface temperature.

We are continuing to examine the assimilation scheme and the model produced subsurface temperature and stratification. For example, the subsurface temperature in the western Gulf of Maine in particular at buoy B tends to be too warm in summer months. It is noted that the strength of winter convection is stronger in the western Gulf of Maine and the deep convection in winter could have lingering effects on the subsurface temperature in summer. Moreover, buoy M was the only mooring with CTD at greater depths (80, 130, 180 and 240m), and there were about 2 month data in June and July and about 1 month data in October 2002. It appears that the model has a larger seasonal variation at depths. The model results will be compared with NMFS hydrographic data (Taylor et al., 2003). A spatially and temporally varying adjustment depth might be needed and a viable approach may be the feature model of Gangopadhyay et al. (2002). We hope to reach a similar benchmark for the subsurface temperature.

Another ongoing task is to establish the benchmark for velocity comparisons in different frequency bands, i.e., the tidal, synoptic, and seasonal variability. It is important to understand the source of errors and their attributions to the errors in the open boundary condition, surface wind forcing, and the modeled hydrographic structure and mixing processes. It is hoped that these analyses can lead to new insights in terms of predictability of the model and guide the assimilation of the sea surface velocity measured by CODARs.

CODARs are the high frequency radar units that can be used to map ocean surface currents. To assimilate the CODAR data into numerical ocean models, Lewis et al. (1998) used a shear stress approach, in which the modeled velocity is nudged towards the CODAR measured velocity by imposing an additional shear stress. It was noted that errors in the CODAR data could cause unrealistic horizontal divergence and sea level in the model. More recently, Lipphardt et al. (2000) found that the unrealistic divergence could be limited by first filtering the CODAR velocity field. A CODAR data assimilation scheme similar to that of Oke et al. (2000) has been developed. It is a simplified Kalman filter data assimilation system that assimilates
low-pass filtered CODAR velocity. The non-homogeneous and non-isotropic forecast error covariance was empirically derived from an ensemble of typical model simulations.

Coastal ocean forecasting is one of the major challenges that the oceanography community faces. Using data to constrain models has been recognized as a key element for successful forecasts. On the other hand, benchmarks for successful predictions need to be established, although they may vary from region to region as limited by the available open boundary condition from larger scale models and by the meteorological forecasts of different resolutions. It is hoped that overall quality of coastal ocean forecasts would be improved as the GODAE products become available and more standardized and as we accumulate the experiences with coastal ocean forecasts.

Acknowledgments. The authors would like to thank Prof. Andrew Thomas and Mr. Ryan Weatherbee at the University of Maine for providing the AVHRR data and Prof. Neal Pettigrew and his group at the University of Maine for providing the buoy data. This study is supported by the Gulf of Maine Ocean Observing System and by the NASA grant (NAG5-10620) to the University of Maine.

4. References
http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0305/