Predictability of the Loop Current Variation and Eddy Shedding Process in the Gulf of Mexico Using an Artificial Neural Network Approach

XIANGMING ZENG, YIZHEN LI, AND RUOYING HE

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

(Manuscript received 15 September 2014, in final form 8 January 2015)

ABSTRACT

A novel approach based on an artificial neural network was used to forecast sea surface height (SSH) in the Gulf of Mexico (GoM) in order to predict Loop Current variation and its eddy shedding process. The empirical orthogonal function analysis method was applied to decompose long-term satellite-observed SSH into spatial patterns (EOFs) and time-dependent principal components (PCs). The nonlinear autoregressive network was then developed to predict major PCs of the GoM SSH in the future. The prediction of SSH in the GoM was constructed by multiplying the EOFs and predicted PCs. Model sensitivity experiments were conducted to determine the optimal number of PCs. Validations against independent satellite observations indicate that the neural network–based model can reliably predict Loop Current variations and its eddy shedding process for a 4-week period. In some cases, an accurate forecast for 5–6 weeks is possible.

1. Introduction

Originating at the Yucatan Channel and exiting through the Florida Straits, the Loop Current (LC) is a dominant circulation feature in the Gulf of Mexico (GoM) (Fig. 1). One of the most notable characteristics of LC is that it episodically sheds large high-speed eddies with diameters of about 200–300 km and swirl speeds of 1.8–2 m s^{-1}, which affect almost every aspect of the GoM, including ocean circulation, biochemical properties, larvae transport, and air–sea interaction (Oey et al. 2005b; Xue et al. 2013; Richards et al. 1993; Small et al. 2008). Between eddy shedding, variations of LC frontal position are also significant. The north and west edges of the LC can vary from about 25.5° to 27.5°N and from 86° to 90°W, respectively (Leben 2005; Gopalakrishnan et al. 2013; Zeng et al. 2015).

Daily operations of approximately 4000 oil and gas platforms in the northern GoM are significantly affected by the LC and its high-speed eddies, which make planning and scheduling a challenge for this expensive enterprise (Leben and Honaker 2006; Sammarco et al. 2004). Accurate prediction of the LC and LC eddies is of critical importance for both scientific research and societal benefit. For example, in order to mitigate the adverse impacts of the *Deepwater Horizon* oil spill in 2010, intensive research on the LC and LC eddies was performed during and after the incident (e.g., Liu et al. 2013). The LC and LC eddies also play an active role in the rapid intensification of GoM hurricanes, such as devastating Hurricanes Katrina and Rita (Leben and Honaker 2006), which caused extensive loss of life and property damage in many Gulf coastal communities.

Much effort has been expended to predict LC variation and its eddy shedding process using remote sensing observations and primitive equation numerical models. Oey et al. (2005a) performed a study to predict the LC and its eddy frontal position using the Princeton Ocean Model (POM). Yin and Oey (2007) applied the bred-ensemble forecast technique to estimate the locations and strengths of the LC and LC eddies from July to September 2005. Counillon and Bertino (2009) presented a small-ensemble forecast using the Hybrid Coordinate Ocean Model to predict LC eddy shedding. A semitheoretical basis was provided by Lugo-Fernández and Leben (2010) on the linear relationship between LC retreat latitude and eddy separation period. Forristall et al. (2010) showed the better skill of the statistical method on the LC prediction than most dynamical models. Mooers et al. (2012) evaluated several different ocean models' performance at LC eddy

Corresponding author address: Ruoying He, Dept. of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Campus Box 8208, 2800 Faucette Drive, Raleigh, NC 27695.
E-mail: rhe@ncsu.edu

DOI: 10.1175/JTECH-D-14-00176.1

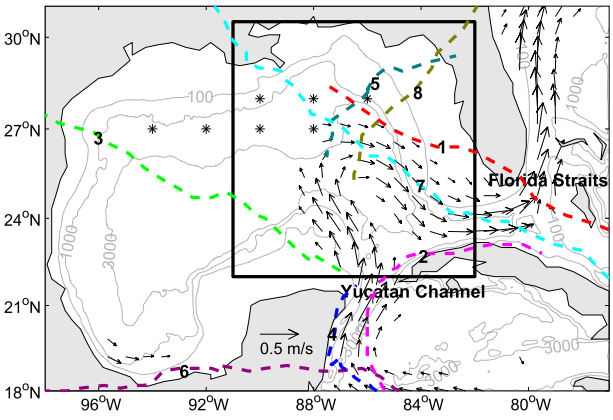

FIG. 1. Study domain. Black box is study area in the Gulf of Mexico. Stars are the seven reference stations for frontal position comparison. Gray lines are depth contours (m). Black arrows are geostrophic velocity calculated from long-term mean AVISO SSH data (only the velocities $> 0.1\ \mathrm{m\ s^{-1}}$ are plotted). The Loop Current is visible via the velocity arrows in the GoM box. Dash lines are tracks of hurricanes or tropical stroms from July 2010 to July 2013. 1: TS Bonnie, 22–25 Jul 2010; 2: Hurricane Paula, 11–15 Oct 2010; 3: TS Don, 27–30 Jul 2011; 4: Hurricane Rina, 22–29 Oct 2011; 5: TS Debby, 23–27 Jun 2012; 6: Hurricane Ernesto, 1–10 Aug 2012; 7: Hurricane Isaac, 20–30 Aug 2012; 8: TS Andrea, 5–10 Jun 2013. Tracks of hurricanes and tropical storms are from NOAA National Climatic Data Center and the Johns Hopkins University Applied Physics Laboratory.

shedding prediction using various prediction skill assessment methods in the report *Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP)*. More recently, with the four-dimensional variation method, Gopalakrishnan et al. (2013) tested the predictability of the LC eddy shedding process using the Massachusetts Institute of Technology General Circulation Model (MITgcm). Xu et al. (2013) applied the local ensemble transform Kalman filter with the parallel POM to estimate the states of the LC and LC eddies from April to July 2010. All these studies are based on either simple empirical relations or primitive equation ocean models focusing on a single LC eddy or a small number of LC eddy shedding events. The lack of generality makes the assessment of their model predictability very difficult (e.g., Mooers et al. 2012).

In this paper, we applied a novel method based on an artificial neural network (ANN) and empirical orthogonal function (EOF) analysis to the sea surface height (SSH) forecast in the GoM in order to predict LC variation and its eddy shedding process. EOF analysis was used to decompose the SSH data into spatial and temporal components. The ANN was used to predict future temporal components' variations. Future SSH was then constructed by combining the predicted temporal components with the spatial pattern. Various prediction skill assessment methods, such as prediction skill score, spatial correlation, and root-mean-square error, were conducted to evaluate this method's prediction skill on LC variation and the eddy shedding process. The structure of this paper is as follows: the data and method are introduced in section 2; parameter experiments, prediction results, skill assessment, and predictability analysis are presented in section 3; a summary is given in section 4.

2. Data and method

a. Dataset

This study used 21 years (1992–2013) of gridded, altimeter-based, absolute dynamic topography data. The altimeter products were produced by SSALTO/Data Unification and Altimeter Combination System (DUACS) and distributed by Archiving, Validation, and Interpretation of Satellite Oceanographic Data (AVISO; http://www.aviso.altimetry.fr/duacs/). Specifically, we used the ''Reference Series'' dataset following Chelton et al. (2011) and Mason et al. (2014) for the more stable satellite sampling (Collecte Localisation Satellites 2011). The data were constructed with two simultaneously operating altimeters, one in a 10-day exact repeat orbit (T/P, followed by *Jason-1* and presently by *Jason-2*) and the other in a 35-day exact repeat orbit (*ERS-1* followed by *ERS-2* and presently by Envisat; Chelton et al. 2011). The data have spatial and temporal resolutions of 1/3° and 7 days, respectively (Collecte Localisation Satellites 2011). We focus on the area of LC variation (indicated by the black box in Fig. 1) to reduce the data dimension and to increase the percentage of leading principal components in EOF analysis (see section 2d). For quality control, only the data located in the area with water depth greater than 100 m were chosen (Liu et al. 2008; Yin et al. 2014; Zeng et al. 2015). During the data selection process, the 2′ Gridded Global Relief Data (ETOP02), distributed by the National Oceanic and Atmospheric Administration (NOAA) National Geophysical Data Center, was interpolated to the SSH grid to obtain a topographic dataset. Other ancillary data, including tracks of hurricanes and tropical storms, were taken from NOAA's National Climatic Data Center and the Johns Hopkins University Applied Physics Laboratory.

b. ANN

The ANN is a computational model inspired by biological neural networks that is capable of solving a variety of problems in pattern recognition, time series prediction, and parameter optimization (e.g., Jain et al.

1996). It has been widely used for variable prediction and mapping in the geosciences community (e.g., Maier and Dandy 2000; Maier et al. 2010; Krasnopolsky 2013). Hsieh and Tang (1998), for example, proposed the use of ANN in meteorology and oceanography, and then conducted a series of applications on Pacific sea surface temperature prediction (Tang et al. 2000; Hsieh 2001), Lorenz dynamical system forecast (Tang and Hsieh 2001), and El Niño–Southern Oscillation analysis (Hsieh 2004). With wind and tidal information as inputs, Lee (2006) predicted storm surge events around Taiwan using the ANN method. Wu et al. (2006) presented the advantage of the ANN over regression in predicting tropical Pacific sea surface temperature. Many applications also appear in pyrgeometer measurements, wind shear alerting, and climate change (e.g., Oliveira et al. 2006; Kwong et al. 2012; Yip and Yau 2012).

The ANN's basic elements are called neurons, which only execute summation over weighted input values, passing their results to a nonlinear transfer function to obtain neuron output values. A three-layer ANN with nonlinear transfer functions can be represented as a set of nonlinear equations used to calculate the output values from the input values (e.g., Maier and Dandy 2000). In the first layer (input layer), each input variable has its own neurons. The second, hidden layer is represented by several neurons, whose number can vary according to the complexity of the problem. Each neuron in the hidden layer receives inputs from all the neuron outputs of the first layer. This fully interconnected procedure is repeated again in the third, output layer. The output layer has one neuron for each output variable (Oliveira et al. 2006).

Each of the three layers has its own weighting factors, which are the ANN parameters determined during the training process. The training process is the determination of the proper interconnection of weighting factors for the ANN based on training dataset patterns, so that the output of the ANN can present the best fit with the output given by the testing dataset. In this way, the ANN learns the information given in the training dataset but still has a generalizing capability, not simply memorizing the patterns in the training dataset. The generalizing capabilities of the ANN guarantee that the trained model can give reasonable results for unknown patterns that differ from the training dataset (Oliveira et al. 2006). These procedures and structures give the ANN the ability of a universal approximator (Maier and Dandy 2000; Oliveira et al. 2006).

c. EOF analysis

In our application, the large spatial and temporal span of the SSH dataset would bring too many input variables to the ANN, which makes direct prediction using the ANN infeasible. As a result, the ANN requires some expensive optimization, and compression of the input dataset is necessary (Rixen et al. 2002). EOF analysis has been commonly used for decades in oceanographic and meteorological applications (e.g., Beckers and Rixen 2003; Hannachi 2004) and is applied in our study to provide a compressed dataset for ANN prediction. In oceanography, EOF analysis has been used for several purposes, such as objective analysis of in situ data (e.g., Holbrook and Bindoff 2000), statistical comparison between data and model results (e.g., Beckers et al. 2002), data reconstruction (e.g., Beckers and Rixen 2003; He et al. 2003; Miles and He 2010; Zhao and He 2012; Li and He 2014), variability analysis (e.g., Hendricks et al. 1996), filtering (e.g., Vautard et al. 1992), and data compression (e.g., Pedder and Gomis 1998; Rixen et al. 2002). EOF analysis can be done by applying the singular value decomposition (SVD) technique (e.g., Beckers and Rixen 2003). Let $\mathbf{X}$ be an $n \times m$ matrix such that the rows indicate temporal development and the columns are variables or spatial data points. The SVD technique can break up the matrix $\mathbf{X}$ into three matrices:

$$\mathbf{X} = \mathbf{UDV}^{\mathrm{T}}, \tag{1}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal and $\mathbf{D}$ is diagonal. Matrix $\mathbf{V}^{\mathrm{T}}$ is the spatial patterns (EOFs), and $\mathbf{UD}$ are the time-dependent principal components (PCs). Let λ_i be the diagonal part of $\mathbf{D}$ with $i = 1, \ldots, m$. Then, the ratio $f_i = \lambda_i^2 / \sum_{i=1}^{m} \lambda_i^2$ is a measure of the variance contained in spatial pattern i compared to the total variance. It is often said that PC i explains $(100 f_i)\%$ of the variance. The ratio is often the basis for deciding the number of PCs to retrain for data compression, and the ones with small ratios are usually discarded (e.g., Beckers and Rixen 2003). Usually, the ratios have been sorted in decreasing order, so that the first several PCs explain the major variance of the dataset.

d. Prediction procedure

The existing SSH of the GoM was chosen as a predictor of future SSH of the GoM using the ANN and EOF analysis. As we mentioned before, because the spatial area of this study leads to too large data dimension, direct SSH prediction is infeasible. To avoid the curse of dimension, EOF analysis was used to split the data into spatial EOFs and time-dependent PCs. From Eq. (1), we have

$$\mathbf{X} = \mathbf{PE}, \tag{2}$$

where $\mathbf{X}$ is our dataset, $\mathbf{P}$ is time-dependent PCs, and $\mathbf{E}$ is spatial EOFs. Equation (2) can be written in another form:

TABLE 1. Principal component number settings for the Gulf of Mexico SSH and corresponding time-averaged (2010–13) spatial CC (bold numbers) and RMSEs (m) between prediction and observation. PC percentage is the percentage variance accounted for by different PCs.

PC No.	PC percentage (%)	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
5	75	**0.889**	**0.884**	**0.867**	**0.838**	**0.800**	**0.758**
		0.101	0.106	0.112	0.130	0.145	0.157
9	85	**0.939**	**0.930**	**0.905**	**0.865**	**0.815**	**0.765**
		0.074	0.083	0.100	0.121	0.140	0.155
18	95	**0.974**	**0.960**	**0.925**	**0.874**	**0.817**	**0.765**
		0.049	0.065	0.092	0.119	0.140	0.155

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,m} \end{pmatrix} = \begin{pmatrix} p_{1,1} & \cdots & p_{1,m} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \cdots & p_{n,m} \end{pmatrix} \begin{pmatrix} e_{1,1} & \cdots & e_{1,m} \\ \vdots & \ddots & \vdots \\ e_{m,1} & \cdots & e_{m,m} \end{pmatrix}, \quad (3)$$

where $x_{i,j}$ is the jth spatial SSH point at time i, n is the time index, and m is the number of spatial points.

Because the first several leading PCs represent the majority of the dataset's variation (Table 1), we can just use the first k PCs to reconstruct the original dataset without losing much information (e.g., Beckers and Rixen 2003); that is,

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,m} \end{pmatrix} \approx \begin{pmatrix} p_{1,1} & \cdots & p_{1,k} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \cdots & p_{n,k} \end{pmatrix} \begin{pmatrix} e_{1,1} & \cdots & e_{1,m} \\ \vdots & \ddots & \vdots \\ e_{k,1} & \cdots & e_{k,m} \end{pmatrix}. \quad (4)$$

Similar to the methodology of Alvarez et al. (2000), Rixen et al. (2002), and Alvarez (2003), we can get the approximation of SSH at time $n + 1$ by

$$(x_{n+1,1} \quad \cdots \quad x_{n+1,m}) \approx (p_{n+1,1} \quad \cdots \quad p_{n+1,k}) \begin{pmatrix} e_{1,1} & \cdots & e_{1,m} \\ \vdots & \ddots & \vdots \\ e_{k,1} & \cdots & e_{k,m} \end{pmatrix} \quad (5)$$

if we can predict $(p_{n+1,1} \quad \cdots \quad p_{n+1,k})$.

The nonlinear autoregressive network with one hidden layer was chosen to do the PC prediction by following Maier and Dandy (2000). For a dataset, the PCs are independent of each other due to the property of EOF analysis (Hannachi 2004). As a result, training and forecasting were applied to the PCs independently; that is,

$$p_{n+1,j} = f(p_{n,j}, p_{n-1,j}, \ldots, p_{n-i,j}), \quad (6)$$

where $p_{n+1,j}$ is the value of the jth PC of the GoM SSH at the $(n + 1)$th record, and i is time delay. Because the time interval of SSH data is weekly, this prediction is one week ahead. Once we get $p_{n+1,j}$, we can put it into Eq. (6) and use it to predict $p_{n+2,j}$. The loop can continue until the accuracy decreases to the lowest acceptable level. The larger the time delay is, the longer the required training time is. We tried different time delays and used the correlation coefficient (CC) and root-mean-square error (RMSE) as the prediction performance criterion to evaluate the optimal time delay i. The time delay for the ANN is set to be $i = 5$, which gives the largest correlation coefficient and least RMSE. The optimal neuron number for each PC prediction was obtained in an iterative fashion based on Kaastra and Boyd's (1996) criteria on neuron number selection.

Here a 6-week leading prediction window was used to demonstrate the model prediction skill. We applied SSH data from 14 October 1994 to 23 June 2010 for the first 6-week prediction. Then, a weekly sliding window was applied from 30 June 2010 to 19 June 2013. During the ANN training procedure, 85% of the data were used for training and 15% were reserved for testing (Wu et al. 2006). To avoid overfitting, the Bayesian regulation method, which is coded as the program "trainbr" in the MATLAB Neural Network toolbox, was chosen for ANN training following Wu et al. (2006).

3. Results and discussion

a. Sensitivity of PC number selection

Figure 2 shows the correlation coefficients of predicted and observation-derived first leading PCs of the GoM SSH (which accounted for 30% of the variance) for different prediction weeks over 3 years (2010–13).

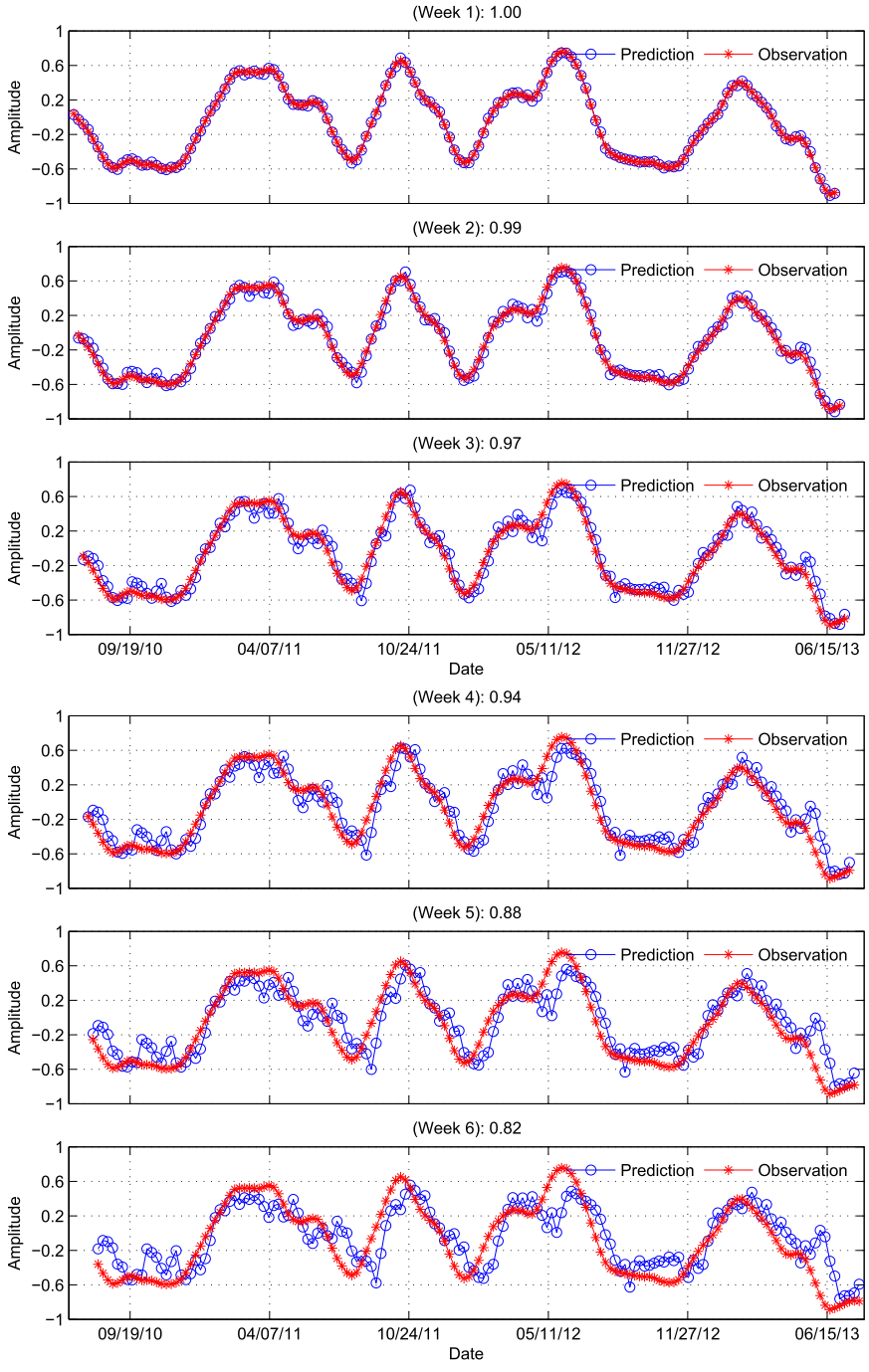

FIG. 2. Comparison between predicted (circles) and observation-derived (stars) first leading principal component of the GoM SSH from 2010 to 2013. Correlation coefficients are presented at the top of each figure.

From weeks 1 to 6, the correlation coefficients decrease from 1.00 to 0.82. For a 1-week prediction, the ANN can predict the PC values almost exactly. While remaining highly correlated, the accuracy decreases as the prediction time grows. As shown in the prediction method [Eq. (6)], small errors in each prediction step can propagate into the next and grow as the prediction time increases. For example, we can see obvious error increasing and propagating from weeks 1 to 6 near 11 May 2012 (Fig. 2). Other PCs' predictions have similar results (not shown).

The future SSH of the GoM can be constructed using Eq. (5) with the predicted PCs from Eq. (6). According

to the percentage variance accounted for by different PCs in EOF analysis, we chose three sets of PCs, corresponding to three percentage levels—75%, 85%, and 95%—for the sensitivity tests (Table 1). To quantify the accuracy of prediction, spatial correlation coefficients and RMSE between predicted and observed GoM SSH were calculated for each experiment as the criteria of prediction performance. Large correlation coefficients and small RMSE represent good prediction skill. According to Table 1, the experiment with 18 PCs has the largest mean spatial correlation coefficients and the least mean RMSE for the 6-week prediction window. Generally, the more PCs are used, the better the performance can be achieved. However, this performance difference decreases from weeks 1 to 6. Part of the reason is that the PCs that account for very small variance percentage are more noisy and difficult to predict. For example, the performance at week 6 is the same for 9 and 18 PCs. As expected, the prediction performance degrades from weeks 1 to 6, and it is because the time series prediction gets worse as the prediction time window extends (e.g., Fig. 2).

b. SSH prediction skill assessment

The following discussion is based on the experiment using 18 PCs, which gives us the best results among the three experiments (Table 1). From June 2010 to June 2013, the spatial correlation coefficients and RMSE between predicted and observed GoM SSH (Fig. 3) show obvious variation of prediction performance: the correlation coefficients are larger than 0.8 from weeks 1 to 3; starting from week 4, the correlation coefficients gradually decrease and oscillate. RMSEs show an opposite pattern, with small errors in weeks 1, 2, and 3 (generally less than 0.1 m) and larger oscillation from weeks 4 to 6. Although the prediction performance decreases from weeks 1 to 6, all the correlation coefficients of week 4 and some of weeks 5 and 6 are still larger than 0.7 (Fig. 3), which means the general pattern of LC variation is well captured at least 4 weeks ahead, sometimes even 5 or 6 weeks ahead.

The oscillation of prediction performance may be due to the highly nonlinear variation of the LC, error propagation, and transitory sea level variability caused by hurricane or tropical storm influence. In particular, the sudden change of atmospheric conditions may cause unusual variation of LC and LC eddies (Oey et al. 2006; Shay and Uhlhorn 2008). For example, at week 6, the least correlation coefficients in October 2011, August 2012, and June 2013 occur during the passage of Hurricane Rina, Hurricane Ernesto, and Tropical Storm Andrea, respectively. We can also see obvious error propagation and amplification from weeks 4 to 6 during these periods (Fig. 3). Figure 1 shows the tracks of hurricanes and tropical storms passing through our study area during prediction period.

To better quantify the prediction skill of our method, we followed methods used by Oey et al. (2005a) and Mooers et al. (2012). First, we compared the SSH RMSE and skill score for prediction and persistence. The prediction RMSE of SSH at week n is defined to be $[\sum_{i=1}^{k}(\mathrm{SSH}m_{i,n} - \mathrm{SSH}o_{i,n})^2/k]^{1/2}$, where k is the total number of SSH data points in study domain, $\mathrm{SSH}m_{i,n}$ is the ith data point of model SSH at week n, $\mathrm{SSH}o_{i,n}$ is the ith data point of observed SSH at week n, and $n = 1, 2, \ldots, 6$. Similarly, the persistence RMSE of SSH at week n is defined to be $[\sum_{i=1}^{k}(\mathrm{SSH}m_{i,0} - \mathrm{SSH}o_{i,n})^2/k]^{1/2}$, where $\mathrm{SSH}o_{i,0}$ is the ith data point of observed SSH at week 0 for each prediction window. Figure 4 shows the SSH RMSE of prediction and persistence. The RMSE differences between prediction and persistence vary for each case. Generally, the RMSE grows with the increase of prediction time (from weeks 1 to 6) for both prediction and persistence. However, the RMSE of prediction increases at a much slower rate than that of persistence, indicating our prediction has much better skill and outperforms the persistence. One case when the prediction RMSE is larger than that of the persistence RMSE occurs around 19 August 2012, the period that Hurricanes Isaac and Ernesto passed through the LC area and northern Caribbean Sea, respectively (Fig. 1). The sudden change in atmospheric conditions may have caused the unusual variation of LC and LC eddies, which decrease the predictability of our model. The averaged RMSE also shows the better skill of prediction than that of persistence (Fig. 5).

Although the RMSE varies for each month, the monthly and 3-yr long-term mean RMSEs of prediction are generally smaller than their persistence counterparts. For example, for the first week, the 3-yr mean RMSE is almost the same for prediction and persistence. However, the 3-yr mean RMSE for persistence increases to about 0.18 m at week 6, whereas the predicted one is only about 0.13 m.

An alternative view of the model's prediction skill is shown in Fig. 6, which presents the same data as Fig. 4 in terms of the skill score. The prediction skill score of SSH at week n is defined to be

$$\mathrm{SS} = 1 - \left[\sum_{i=1}^{k}(\mathrm{SSH}m_{i,n} - \mathrm{SSH}o_{i,n})^2/k\right] \Big/ \left[\sum_{i=1}^{k}(\mathrm{SSH}o_{i,n} - \overline{\mathrm{SSH}o_{i,n}})^2/k\right], \quad (7)$$

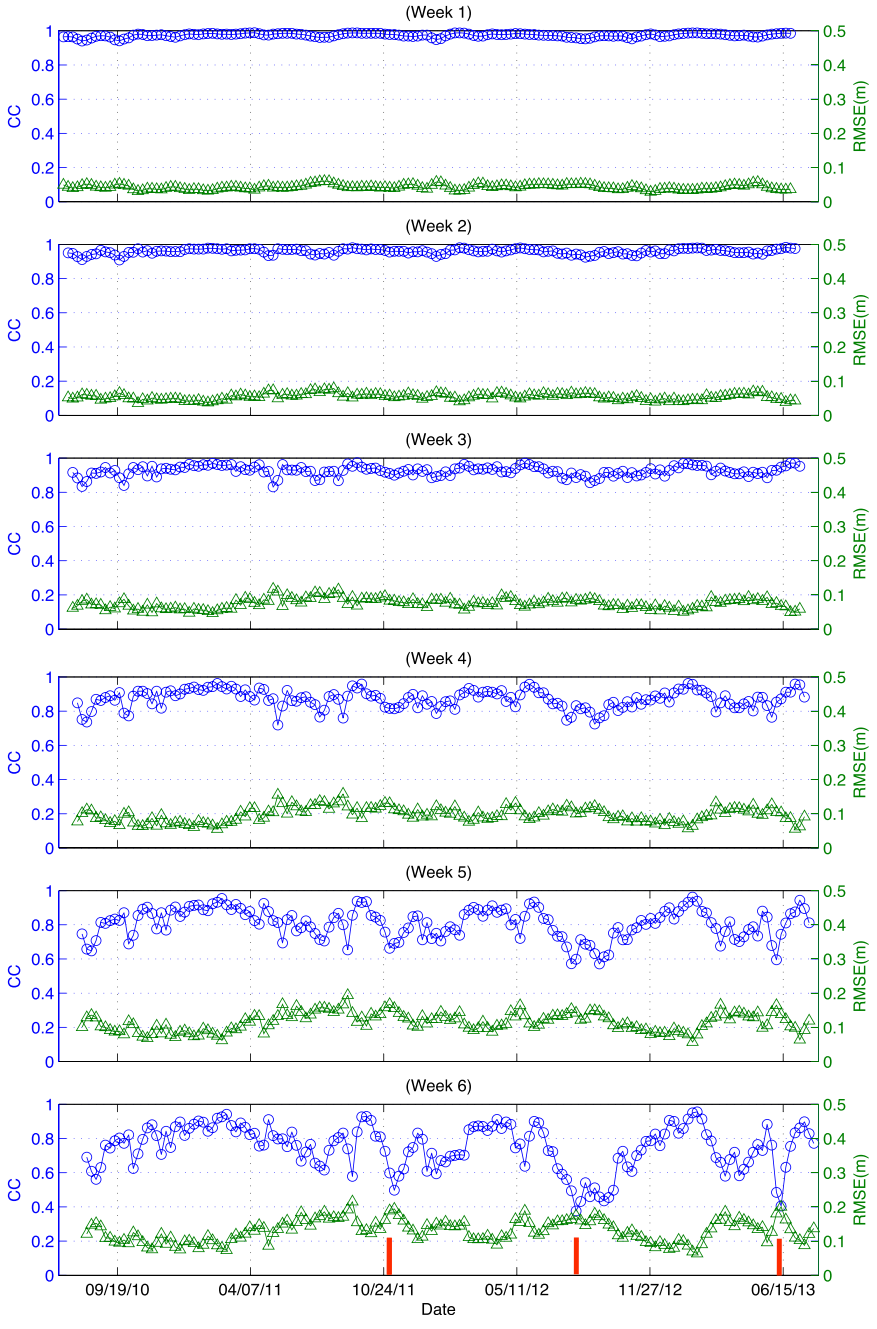

FIG. 3. Spatial CC and RMSEs of predicted and observed SSH in the study area from 2010 to 2013 with a 6-week ahead weekly sliding prediction window using 18 PCs. Circles are CC points, and triangles are RMSE points. Red lines in week 6 indicate the approximate passing time of Hurricane Rina in 2011, Hurricane Ernesto in 2012, and TS Andrea in 2013.

where the bar represents the arithmetical mean (Murphy 1988; Mooers et al. 2012). The persistence skill score of SSH at week n can be derived by simply replacing SSH$m_{i,n}$ with SSH$o_{i,0}$. The skill score is based on mean square, including both bias and variance, in order to facilitate the intercomparison with respect to a constant mean sea surface throughout the prediction time period (Mooers et al. 2012).

The skill scores of both prediction and persistence decrease with the increase of prediction time, while the amplitude and rate of decrease are different for each prediction case. Generally, the prediction skill score is

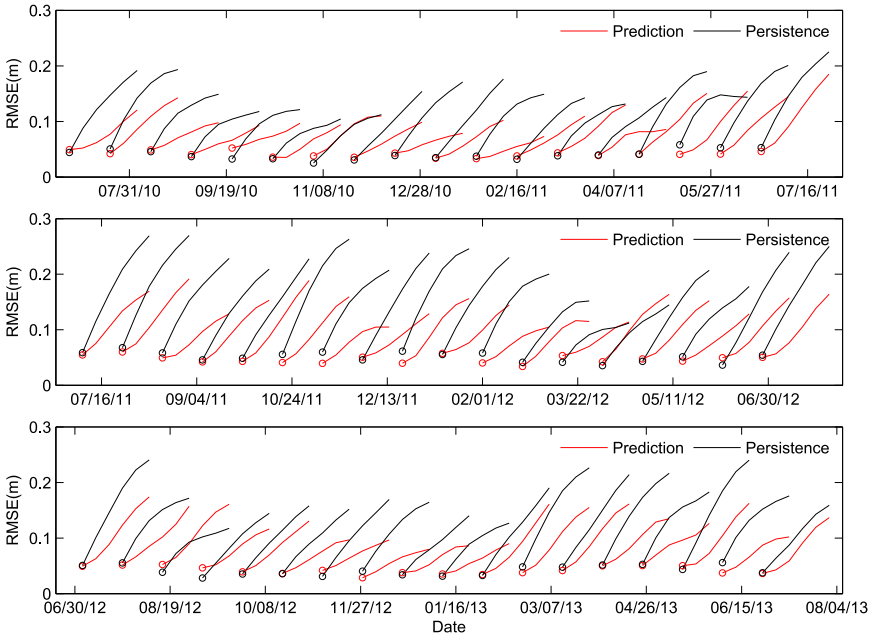

FIG. 4. RMSE comparison of SSH between prediction (red) and persistence (black). Circles represent the location of week 1. Values are plotted every 4 weeks.

greater than that of persistence and decays at a much slower rate. Similar to the RMSE case, the scenario when prediction skill was less than persistence skill occurred around 19 August 2012 when the variation of LC and LC eddies may be unusual due to the influences of hurricanes in or near the study domain (Fig. 1).

The averaged skill scores also show better skill of prediction than persistence (Fig. 7). Although the skill score varies for each month and decreases from weeks 1 to 6 for both prediction and persistence, the monthly and 3-yr mean skill scores of prediction are generally greater than their persistence counterparts. For example, the 3-yr mean skill score starts off the same for both prediction and persistence at week 1. As the prediction time increases to 6 weeks, it decreases to about 0.1 for persistence, whereas our model still has 0.5 prediction skill.

c. *Frontal position prediction skill assessment*

The frontal position of the LC and LC eddies is another metric to measure the prediction skill of our method. Here, frontal position predictability is evaluated using the methodology of Oey et al. (2005a). Because of the properties of the EOF method, the prediction tends to smooth the SSH contours at reduced values of SSH, especially near the study boundary area. As a result, it is difficult to find an optimal reference SSH contour line for frontal position comparison between prediction and observation. After a series of experiments, we chose 0.45-m SSH contour lines as the representation of the LC and LC fronts to better facilitate our skill assessment (Zeng et al. 2015). Using seven stations as reference positions (stars in Fig. 1), we define prediction error at week n as $E_{j,n} = dm_{j,n} - do_{j,n}$, where $dm_{j,n}$ is the shortest distance from predicted LC and LC eddies fronts to the station j at week n, $do_{j,n}$ is the corresponding observed distance, and $n = 1, 2, \ldots, 6$. We similarly define the persistence

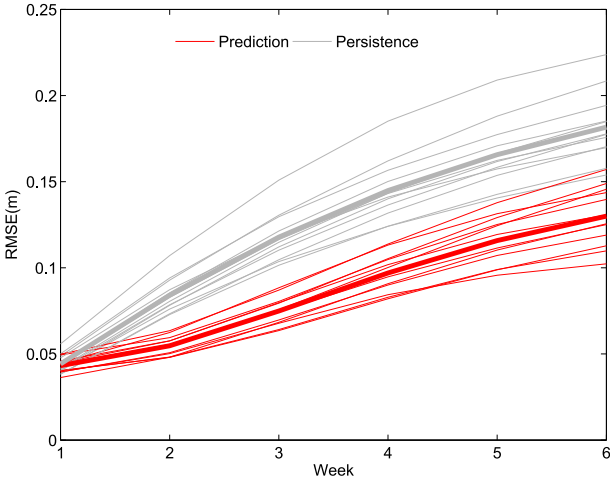

FIG. 5. Averaged RMSE of SSH for prediction (red) and persistence (gray). Thin lines are monthly means, and thick lines are means over the 3-yr prediction period.

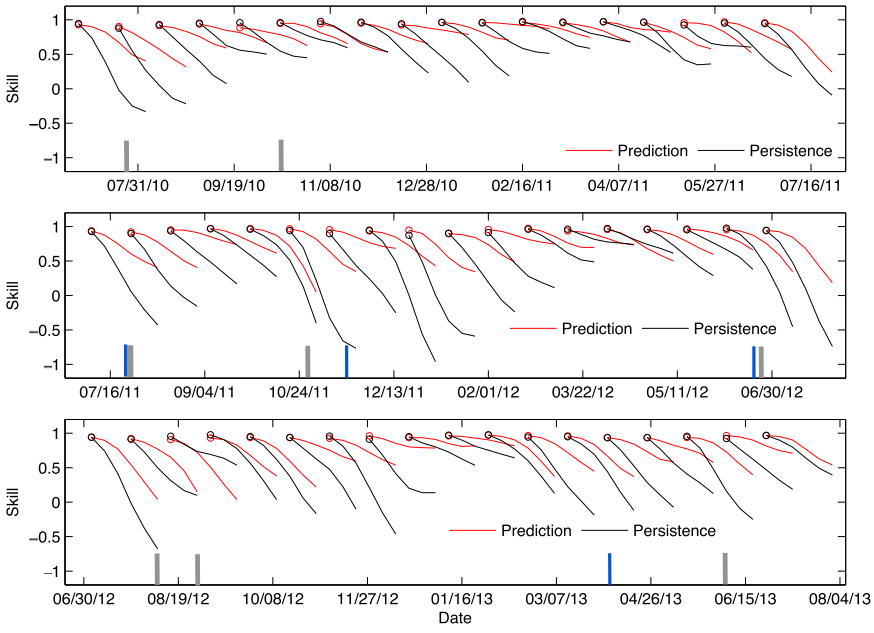

FIG. 6. Skill score comparison of SSH between prediction (red) and persistence (black). Circles represent the location of week 1. Values are plotted every 4 weeks. Gray short lines represent the approximate passing time of hurricanes or tropical storms in Fig. 1. Blue short lines indicate the time of LC eddy shedding.

error at week n to be $P_{j,n} = do_{j,0} - do_{j,n}$, where $do_{j,0}$ is the shortest distance from observed LC and LC eddies fronts to station j at week 0.Considering the limitation of the EOF technique, small-scale sea level features will not be reproduced well by the SSH reconstruction process. Small eddies (perimeter < 300 km) and data points within $\pm 0.5°$ of the domain boundary are therefore excluded for this comparison. Again, as prediction time increases, the RMSE of the frontal position grows for both prediction and persistence (Fig. 8). Although the values change for each prediction case, the RMSE of the predicted frontal position is generally smaller than that of persistence. The large jumps of persistence RMSE (Fig. 8) are the scenarios when the LC eddies move out of the study domain.

Because of the smoothing effect of EOF analysis, the model's prediction of the frontal position is not as good as that of SSH. However, generally, the frontal position prediction is much better than that of persistence. The monthly and 3-yr mean RMSE of the frontal position shows the model's advantage clearly (Fig. 9). To avoid the bias associated with the large jumps in Fig. 8, we excluded the large jump scenarios (>200 km) when calculating the means. Again, the RMSE of the frontal position is different each month for both prediction and persistence. The mean RMSE over 3 years of the predicted frontal position reaches about 60 km at week 6, while it grows to about 85 km for persistence. The 3-yr mean RMSE of the frontal position for prediction is less than that of persistence even for week 1.

d. One LC eddy shedding example

Figure 10 shows the comparison of predicted and observed GoM SSH for one LC eddy shedding event that occurred during the study time period. Again,

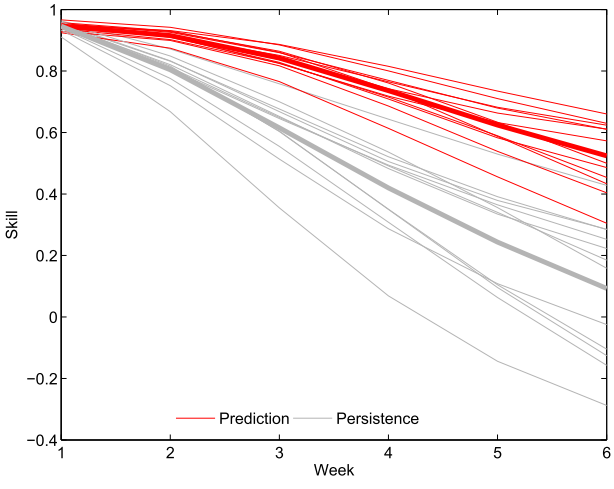

FIG. 7. Averaged skill score of SSH for prediction (red) and persistence (gray). Thin lines are monthly mean, and thick lines are means over the 3-yr prediction period.

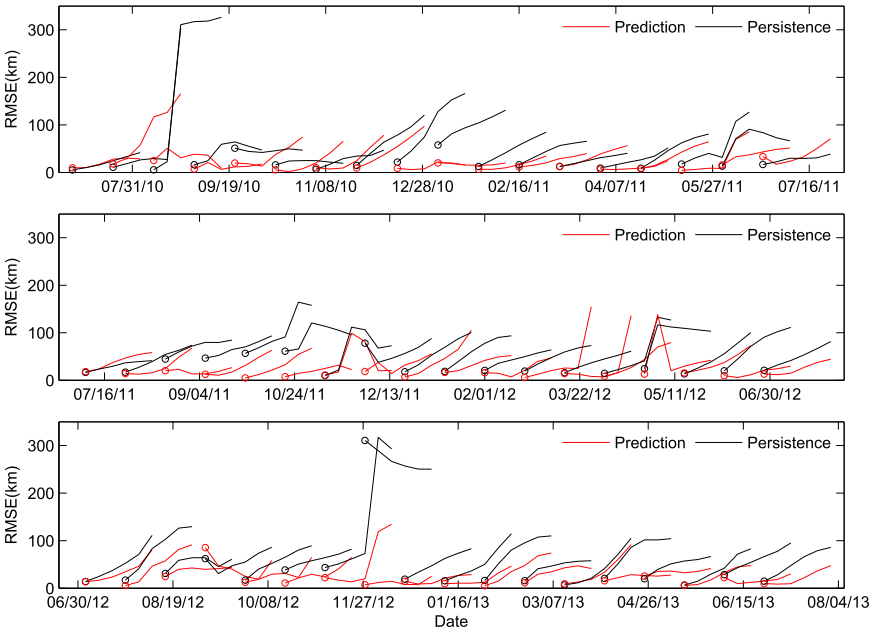

FIG. 8. Frontal position RMSE comparison of LC and LC eddies between prediction (red) and persistence (black). Circles represent the location of week 1. Values are plotted every 4 weeks.

because of the smoothing property of the EOF method, the SSH between LC and LC eddies tends to be larger than observation, which makes it difficult to find an optimal contour line to represent the edges of LC and LC eddies for both prediction and observation. After several experiments, we choose 0.45- and 0.51-m SSH contours as the required edges for observation and prediction, respectively, to demonstrate the eddy shedding process. Here, eddy detachment is defined as the separation of two SSH contour lines from one entire LC SSH contour line (week 4 in Fig. 10), which occurred simultaneously in the prediction and observation. We note that the choices of different SSH contour lines for prediction and observation are only for clear demonstration of the LC eddy shedding process.

The variation of LC and LC eddies occurs at the same time in both prediction and observation: at weeks 1 and 2, LC extends to the 1000-m isobaths near 28°N, and the SSH on both sides of the extended LC is low; from week 3, the LC eddy starts to shed from the LC; at week 4, the SSH contour lines clearly show that the LC eddy is separated from the LC; from weeks 5 to 6, the distance between the LC and shed eddy is further separated. Because of the statistical property of EOF analysis, there are some lower-predicted SSH areas as compared to the observation, and the shed eddy tends to be smaller than the observed one. The influence of Tropical Storm (TS) Don during 27–30 July 2011 on the GoM SSH may also have reduced the accuracy of our prediction (Fig. 10). Although the track of TS Don does not pass directly over the LC, its influence on LC frontal eddies may still play an important role on LC variation and the eddy shedding process (e.g., Le Hénaff et al. 2012; Androulidakis et al. 2014). In this case, LC eddy shedding can still be predicted 4 weeks ahead (Fig. 10). Overall, although there were some differences between predicted and observed SSH

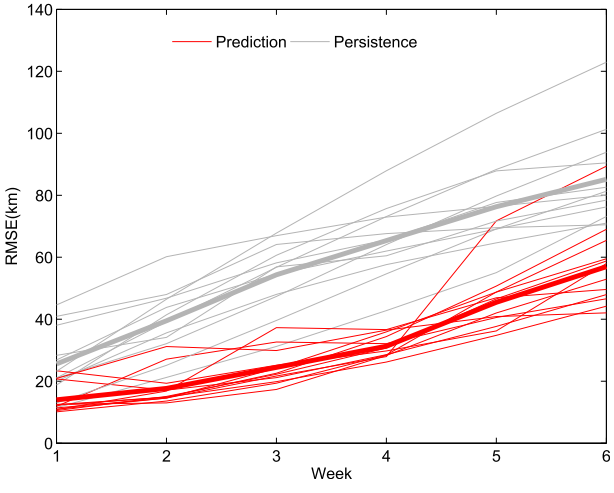

FIG. 9. Averaged frontal position RMSE for prediction (red) and persistence (gray). Sudden jumps in Fig. 8 were excluded when the average was calculated. Thin lines are monthly means, and thick lines are means over the 3-yr period.

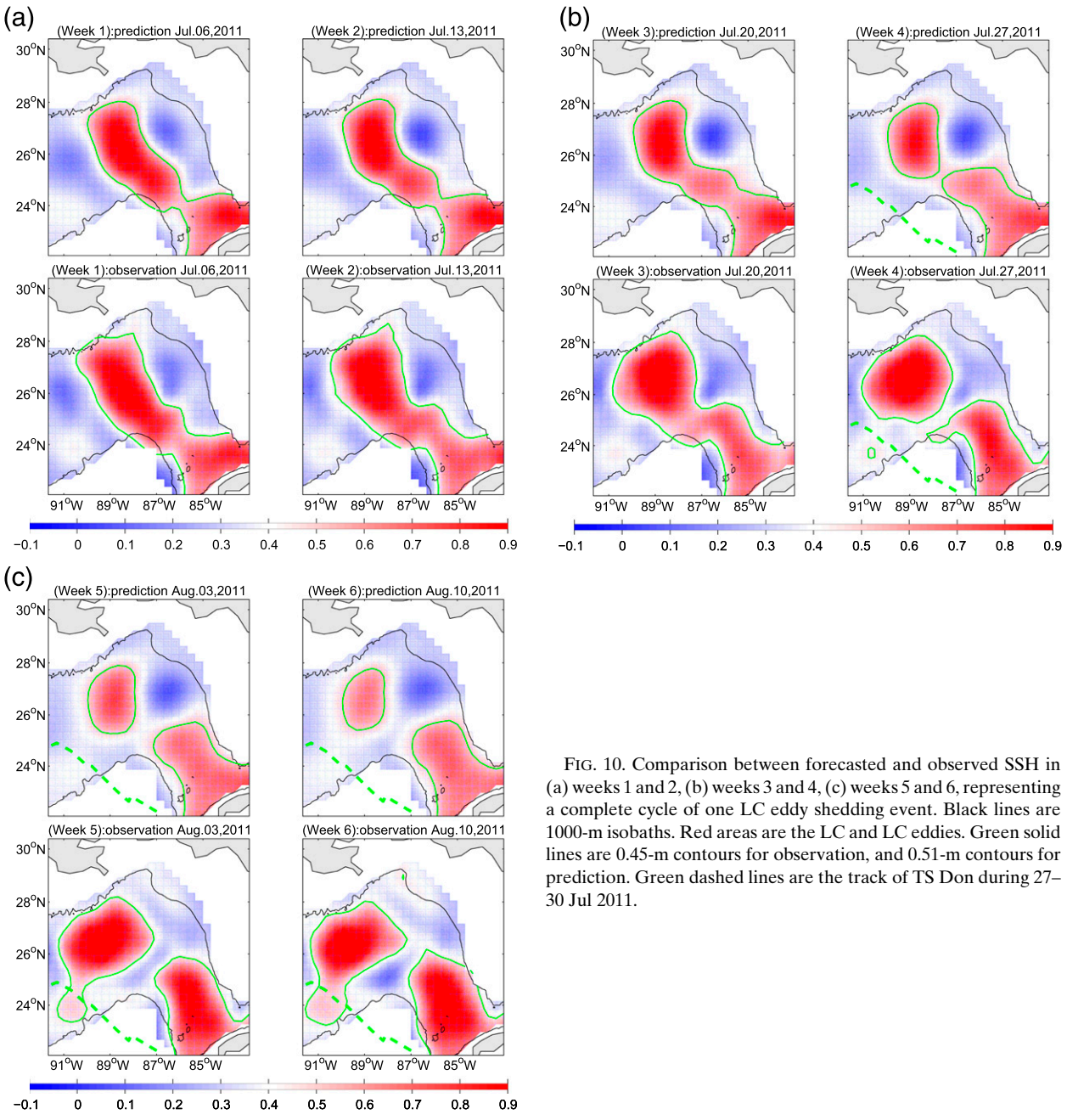

FIG. 10. Comparison between forecasted and observed SSH in (a) weeks 1 and 2, (b) weeks 3 and 4, (c) weeks 5 and 6, representing a complete cycle of one LC eddy shedding event. Black lines are 1000-m isobaths. Red areas are the LC and LC eddies. Green solid lines are 0.45-m contours for observation, and 0.51-m contours for prediction. Green dashed lines are the track of TS Don during 27–30 Jul 2011.

values, the general evolution of the LC and LC eddies was well captured.

4. Summary

This study used 21 years of satellite data to build and test a nonlinear statistical model for predicting the GoM SSH, from which LC variation and the eddy shedding process can be predicted. To reduce the data dimension, we first applied the EOF analysis to decompose existing SSH into spatial patterns (EOFs) and time-dependent principal components (PCs). The nonlinear, autoregressive neural network was then used to predict leading PCs of the GoM SSH 6 weeks ahead. Finally, the future SSH of the GoM was constructed by multiplying the spatial EOFs of the GoM SSH and the predicted PCs. Having tested the sensitivity of prediction results to different variance percentage levels, the 95% level (with 18 PCs) was selected for skill assessment and analysis.

To evaluate the prediction skill of this method, a 6-week weekly sliding prediction window was performed over 3 years. We calculated the spatial correlation coefficients and RMSEs between predicted and observed SSH during the 3 years and then compared the skill assessment metrics between prediction and persistence, such as the SSH skill score and RMSE, and the frontal position RMSE of the LC and LC eddies. For SSH prediction, the 3-yr mean RMSE of prediction is about 30% less than that of persistence at week 6. The 3-yr mean skill score of SSH prediction is about 0.5 compared to 0.1 for persistence at week 6. The RMSE of the frontal position also shows better skill of prediction (60 km) over persistence (80 km).

Although there are some performance oscillations in the later weeks of the prediction window, the generally high correlation coefficients and low RMSEs indicate the high accuracy of our method for LC variation and eddy shedding prediction. The skill comparison between prediction and persistence also validates the prediction skill of our methods. Generally, the model can capture the LC variation and eddy shedding process 4 weeks ahead, and in some cases, 5 and 6 weeks ahead is possible.

The prediction errors mainly come from the ANN PC prediction and EOF SSH construction. Because of the complexity of LC variation, the influence of atmospheric conditions, and the limit of available datasets, the input information may not be enough for the ANN to capture the underlying relation of PC variation. For example, the accuracy of prediction decreases when hurricanes or tropical storms happen in or near the study area. Part of the reason is that air–sea interaction was not built into our model, and these extreme events cannot be well predicted. Because only the leading PCs were used for prediction and SSH construction, some information was omitted. However, because of the high dimensionality of the datasets and computational efficiency, EOF decomposition and construction techniques are necessary for our prediction.

Several advantages of the neural network method over primitive equation dynamical ocean models for LC prediction are 1) the nonlinear temporal variation can be better captured by a nonlinear statistical approach, 2) the prediction accuracy has more generality as demonstrated by continuous LC predictions over a period of 3 years, 3) the computation is usually faster than primitive ocean models, and 4) no boundary and forcing conditions or partial differential equation discretization are needed. For phenomena such as the GoM LC and the LC eddy shedding process that cannot be well predicted by modern dynamical ocean models, this method provides an alternative means for an accurate forecast.

Acknowledgments. The research support provided by Gulf of Mexico Research Initiative/GISR Grant 02-S130202, NOAA Grant NA11NOS0120033, and NASA Grants NNX12AP84G and NNX13AD80G is much appreciated. We also thank Dr. E. Zaron and Dr. F. Semazzi for the helpful discussions, and J. Warrillow for the editorial assistance. The constructive comments from two anonymous reviewers are also appreciated.

REFERENCES

Alvarez, A., 2003: Performance of satellite-based ocean forecasting (SOFT) systems: A study in the Adriatic Sea. *J. Atmos. Oceanic Technol.,* **20,** 717–729, doi:10.1175/1520-0426(2003)20<717:POSBOF>2.0.CO;2.

——, C. Lopez, M. Riera, E. Hernandez-Garcia, and J. Tintore, 2000: Forecasting the SST space-time variability of the Alboran Sea with genetic algorithms. *Geophys. Res. Lett.,* **27,** 2709–2712, doi:10.1029/1999GL011226.

Androulidakis, Y. S., V. H. Kourafalou, and M. Le Hénaff, 2014: Influence of frontal cyclones evolution on the 2009 (Ekman) and 2010 (Franklin) Loop Current Eddy detachment events. *Ocean Sci. Discuss.,* **11,** 1949–1994, doi:10.5194/osd-11-1949-2014.

Beckers, J.-M., and M. Rixen, 2003: EOF calculations and data filling from incomplete oceanographic datasets. *J. Atmos. Oceanic Technol.,* **20,** 1839–1856, doi:10.1175/1520-0426(2003)020<1839:ECADFF>2.0.CO;2.

——, and Coauthors, 2002: Model intercomparison in the Mediterranean: MEDMEX simulations of the seasonal cycle. *J. Mar. Syst.,* **33–34,** 215–251, doi:10.1016/S0924-7963(02)00060-X.

Chelton, D. B., M. G. Schlax, and R. M. Samelson, 2011: Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.,* **91,** 167–216, doi:10.1016/j.pocean.2011.01.002.

Collecte Localisation Satellites, 2011: SSALTO/DUACS user handbook: (M)SLA and (M)ADT near-real time and delayed time products. Version 2rev4, Rep. SALP-MU-P-EA-21065-CLS, 49 pp.

Counillon, F., and L. Bertino, 2009: High-resolution ensemble forecasting for the Gulf of Mexico eddies and fronts. *Ocean Dyn.,* **59,** 83–95, doi:10.1007/s10236-008-0167-0.

Forristall, G. Z., R. R. Leben, and C. A. Hall, 2010: SS: Metocean: A statistical hindcast and forecast model for the Loop Current. *Proc. Offshore Technology Conf.,* Houston, TX, OTC, Paper OTC-20602-MS, 12 pp., doi:10.4043/20602-MS.

Gopalakrishnan, G., B. D. Cornuelle, I. Hoteit, D. L. Rudnick, and W. B. Owens, 2013: State estimates and forecasts of the loop current in the Gulf of Mexico using the MITgcm and its adjoint. *J. Geophys. Res. Oceans,* **118,** 3292–3314, doi:10.1002/jgrc.20239.

Hannachi, A., 2004: A primer for EOF analysis of climate data. University of Reading Rep., 33 pp. [Available online at http://www.met.rdg.ac.uk/~han/Monitor/eofprimer.pdf.]

He, R., R. H. Weisberg, H. Zhang, F. E. Muller-Karger, and R. W. Helber, 2003: A cloud-free, satellite-derived, sea surface temperature analysis for the West Florida Shelf. *Geophys. Res. Lett.,* **30,** 1811, doi:10.1029/2003GL017673.

Hendricks, J. R., R. R. Leben, G. H. Born, and C. J. Koblinsky, 1996: Empirical orthogonal function analysis of global TOPEX/POSEIDON altimeter data and implications for detection of global sea level rise. *J. Geophys. Res.,* **101,** 14 131–14 145, doi:10.1029/96JC00922.

Holbrook, N. J., and N. L. Bindoff, 2000: A statistically efficient mapping technique for four-dimensional ocean temperature data. *J. Atmos. Oceanic Technol.,* **17,** 831–846, doi:10.1175/1520-0426(2000)017<0831:ASEMTF>2.0.CO;2.

Hsieh, W. W., 2001: Nonlinear principal component analysis by neural networks. *Tellus,* **53A,** 599–615, doi:10.1034/j.1600-0870.2001.00251.x.

——, 2004: Nonlinear multivariate and time series analysis by neural network methods. *Rev. Geophys.,* **42,** RG1003, doi:10.1029/2002RG000112.

——, and B. Tang, 1998: Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. Amer. Meteor. Soc.,* **79,** 1855–1870, doi:10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO;2.

Jain, A. K., J. Mao, and K. M. Mohiuddin, 1996: Artificial neural networks: A tutorial. *Computer,* **29,** 31–44, doi:10.1109/2.485891.

Kaastra, I., and M. Boyd, 1996: Designing a neural network for forecasting financial and economic time series. *Neurocomputing,* **10,** 215–236, doi:10.1016/0925-2312(95)00039-9.

Krasnopolsky, V., 2013: *The Application of Neural Networks in the Earth System Sciences: Neural Networks Emulations for Complex Multidimensional Mappings.* Atmospheric and Oceanographic Sciences Library, Vol. 46, Springer Science & Business, 205 pp.

Kwong, K. M., M. H. Y. Wong, J. N. K. Liu, and P. W. Chan, 2012: An artificial neural network with chaotic oscillator for wind shear alerting. *J. Atmos. Oceanic Technol.,* **29,** 1518–1531, doi:10.1175/2011JTECHA1501.1.

Leben, R. R., 2005: Altimeter-derived Loop Current metrics. *Circulation in the Gulf of Mexico: Observations and Models, Geophys. Monogr.,* Vol. 161, Amer. Geophys. Union, 181–201.

——, and D. J. Honaker, 2006: What do we know and what can we predict about the timing of Loop Current Eddy separation? *Proceedings of the Symposium on 15 Years of Progress in Radar Altimetry,* D. Danesy, Ed., ESA Special Publ. SP-614, Paper 19.

Lee, T.-L., 2006: Neural network prediction of a storm surge. *Ocean Eng.,* **33,** 483–494, doi:10.1016/j.oceaneng.2005.04.012.

Le Hénaff, M., V. H. Kourafalou, Y. Morel, and A. Srinivasan, 2012: Simulating the dynamics and intensification of cyclonic Loop Current Frontal Eddies in the Gulf of Mexico. *J. Geophys. Res.,* **117,** C02034, doi:10.1029/2011JC007279.

Li, Y., and R. He, 2014: Spatial and temporal variability of SST and ocean color in the Gulf of Maine based on cloud-free SST and chlorophyll reconstructions in 2003–2012. *Remote Sens. Environ.,* **144,** 98–108, doi:10.1016/j.rse.2014.01.019.

Liu, Y., R. H. Weisberg, and Y. Yuan, 2008: Patterns of upper layer circulation variability in the South China Sea from satellite altimetry using the Self-Organizing Map. *Acta Oceanol. Sin.,* **27** (Suppl.), 129–144.

——, A. MacFadyen, Z.-G. Ji, and R. H. Weisberg, 2013: *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record Breaking Enterprise.* John Wiley & Sons, 280 pp.

Lugo-Fernández, A., and R. R. Leben, 2010: On the linear relationship between Loop Current retreat latitude and eddy separation period. *J. Phys. Oceanogr.,* **40,** 2778–2784, doi:10.1175/2010JPO4354.1.

Maier, H. R., and G. C. Dandy, 2000: Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environ. Modell. Software,* **15,** 101–124, doi:10.1016/S1364-8152(99)00007-9.

——, A. Jain, G. C. Dandy, and K. P. Sudheer, 2010: Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environ. Modell. Software,* **25,** 891–909, doi:10.1016/j.envsoft.2010.02.003.

Mason, E., A. Pascual, and J. C. McWilliams, 2014: A new sea surface height–based code for oceanic mesoscale eddy tracking. *J. Atmos. Oceanic Technol.,* **31,** 1181–1188, doi:10.1175/JTECH-D-14-00019.1

Miles, T. N., and R. He, 2010: Temporal and spatial variability of Chl-*a* and SST on the South Atlantic Bight: Revisiting with cloud-free reconstructions of MODIS satellite imagery. *Cont. Shelf Res.,* **30,** 1951–1962, doi:10.1016/j.csr.2010.08.016.

Mooers, C. N. K., E. D. Zaron, and M. K. Howard, 2012: Final report for phase I: Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP). Final Rep. to Research Partnership to Secure Energy for America, 149 pp.

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, doi:10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2.

Oey, L. Y., T. Ezer, G. Forristall, C. Cooper, S. DiMarco, and S. Fan, 2005a: An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico. *Geophys. Res. Lett.,* **32,** L12611, doi:10.1029/2005GL023253.

——, ——, and H.-C. Lee, 2005b: Loop Current, rings and related circulation in the Gulf of Mexico: A review of numerical models and future challenges. *Circulation in the Gulf of Mexico: Observations and Models, Geophys. Monogr.,* Vol. 161, Amer. Geophys. Union, 31–56.

——, ——, D.-P. Wang, S.-J. Fan, and X.-Q. Yin, 2006: Loop Current warming by Hurricane Wilma. *Geophys. Res. Lett.,* **33,** L08613, doi:10.1029/2006GL025873.

Oliveira, A. P., J. Soares, M. Z. Božnar, P. Mlakar, and J. F. Escobedo, 2006: An application of neural network technique to correct the dome temperature effects on pyrgeometer measurements. *J. Atmos. Oceanic Technol.,* **23,** 80–89, doi:10.1175/JTECH1829.1.

Pedder, M., and D. Gomis, 1998: Applications of EOF analysis to the spatial estimation of circulation features in the ocean sampled by high-resolution CTD soundings. *J. Atmos. Oceanic Technol.,* **15,** 959–978, doi:10.1175/1520-0426(1998)015<0959:AOEATT>2.0.CO;2.

Richards, W. J., M. F. McGowan, T. Leming, J. T. Lamkin, and S. Kelley, 1993: Larval fish assemblages at the Loop Current boundary in the Gulf of Mexico. *Bull. Mar. Sci.,* **53,** 475–537.

Rixen, M., J.-M. Beckers, A. Alvarez, and J. Tintore, 2002: Results on SSH neural network forecasting in the Mediterranean Sea. *Remote Sensing of the Ocean and Sea Ice 2001,* M. Rixen et al., Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 4544), 24, doi:10.1117/12.452757.

Sammarco, P. W., A. D. Atchison, and G. S. Boland, 2004: Expansion of coral communities within the Northern Gulf of Mexico via offshore oil and gas platforms. *Mar. Ecol.: Prog. Ser.,* **280,** 129–143, doi:10.3354/meps280129.

Shay, L. K., and E. W. Uhlhorn, 2008: Loop Current response to Hurricanes Isidore and Lili. *Mon. Wea. Rev.,* **136,** 3248–3274, doi:10.1175/2007MWR2169.1.

Small, R. J., and Coauthors, 2008: Air–sea interaction over ocean fronts and eddies. *Dyn. Atmos. Oceans,* **45,** 274–319, doi:10.1016/j.dynatmoce.2008.01.001.

Tang, B., W. W. Hsieh, A. H. Monahan, and F. T. Tangang, 2000: Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea

surface temperatures. *J. Climate,* **13,** 287–293, doi:10.1175/1520-0442(2000)013<0287:SCBNNA>2.0.CO;2.

Tang, Y., and W. W. Hsieh, 2001: Coupling neural networks to incomplete dynamical systems via variational data assimilation. *Mon. Wea. Rev.,* **129,** 818–834, doi:10.1175/1520-0493(2001)129<0818:CNNTID>2.0.CO;2.

Vautard, R., P. Yiou, and M. Ghil, 1992: Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D,* **58,** 95–126, doi:10.1016/0167-2789(92)90103-T.

Wu, A., W. W. Hsieh, and B. Tang, 2006: Neural network forecasts of the tropical Pacific sea surface temperatures. *Neural Networks,* **19,** 145–154, doi:10.1016/j.neunet.2006.01.004.

Xu, F.-H., L.-Y. Oey, Y. Miyazawa, and P. Hamilton, 2013: Hindcasts and forecasts of Loop Current and eddies in the Gulf of Mexico using local ensemble transform Kalman filter and optimum-interpolation assimilation schemes. *Ocean Modell.,* **69,** 22–38, doi:10.1016/j.ocemod.2013.05.002.

Xue, Z., R. He, K. Fennel, W.-J. Cai, S. Lohrenz, and C. Hopkinson, 2013: Modeling ocean circulation and biogeochemical variability in the Gulf of Mexico. *Biogeosciences,* **10,** 7219–7234, doi:10.5194/bg-10-7219-2013.

Yin, X.-Q., and L.-Y. Oey, 2007: Bred-ensemble ocean forecast of loop current and rings. *Ocean Modell.,* **17,** 300–326, doi:10.1016/j.ocemod.2007.02.005.

Yin, Y., X. Lin, Y. Li, and X. Zeng, 2014: Seasonal variability of Kuroshio intrusion northeast of Taiwan Island as revealed by self-organizing map. *Chin. J. Oceanol. Limnol.,* **32,** 1435–1442, doi:10.1007/s00343-015-4017-x.

Yip, Z. K., and M. K. Yau, 2012: Application of artificial neural networks on North Atlantic tropical cyclogenesis potential index in climate change. *J. Atmos. Oceanic Technol.,* **29,** 1202–1220, doi:10.1175/JTECH-D-11-00178.1.

Zeng, X., Y. Li, R. He, and Y. Yin, 2015: Clustering of Loop Current patterns based on the satellite-observed sea surface height and self-organizing map. *Remote Sens. Lett.,* **6,** 11–19, doi:10.1080/2150704X.2014.998347.

Zhao, Y., and R. He, 2012: Cloud-free sea surface temperature and colour reconstruction for the Gulf of Mexico: 2003–2009. *Remote Sens. Lett.,* **3,** 697–706, doi:10.1080/01431161.2012.666638.