

# Improving Short-Range Ensemble Kalman Storm Surge Forecasting Using Robust Adaptive Inflation

M. U. ALTAF

*Delft University of Technology, Delft, Netherlands, and King Abdullah University of Science and Technology, Thuwal, Saudia Arabia*

T. BUTLER

*Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas*

X. LUO

*International Research Institute of Stavanger, Bergen, Norway*

C. DAWSON AND T. MAYO

*Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas*

I. HOTEIT

*King Abdullah University of Science and Technology, Thuwal, Saudia Arabia*

(Manuscript received 17 October 2012, in final form 29 January 2013)

## ABSTRACT

This paper presents a robust ensemble filtering methodology for storm surge forecasting based on the singular evolutive interpolated Kalman (SEIK) filter, which has been implemented in the framework of the  $H_\infty$  filter. By design, an  $H_\infty$  filter is more robust than the common Kalman filter in the sense that the estimation error in the  $H_\infty$  filter has, in general, a finite growth rate with respect to the uncertainties in assimilation. The computational hydrodynamical model used in this study is the Advanced Circulation (ADCIRC) model. The authors assimilate data obtained from Hurricanes Katrina and Ike as test cases. The results clearly show that the  $H_\infty$ -based SEIK filter provides more accurate short-range forecasts of storm surge compared to recently reported data assimilation results resulting from the standard SEIK filter.

## 1. Introduction

In recent years, a number of critical events have motivated the efforts to predict water levels accurately. Accurate predictions lead to improved management of public safety as well as improved management of environmental and economic interests. Storm surges from hurricanes and tropical cyclones are responsible for many deaths worldwide. For example, the infamous Bhola cyclone that made landfall in Bangladesh in

November 1970 is responsible for approximately 500 000 deaths, and the cyclone in the southern North Sea in February 1953 claimed thousands of lives (Murty et al. 1986; Gerritsen et al. 1995; McRobie et al. 2005). More recently, Hurricane Katrina made landfall in Louisiana and Mississippi in August 2005 (Knabb et al. 2006). The vast majority of deaths caused by hurricanes are as a result of storm surge and might be prevented with improved planning, warning systems, and emergency response. It is, therefore, imperative that accurate numerical forecasts of coastal flooding be delivered in real time to aid the coordination of evacuations and enhance the capabilities of existing warning systems.

Present-day computational models of storm surge are both sophisticated and make use of fine meshes, but the

---

*Corresponding author address:* M. U. Altaf, Department of Applied Mathematics, Delft University of Technology, Mekelweg 4, Delft, 2628CD, Netherlands.  
E-mail: m.u.altaf@tudelft.nl

use of finite computational resources and the requirement that forecasts be delivered in a timely manner result in significant epistemic uncertainties, and significant aleatory uncertainties will always exist (Brown et al. 2007). Major sources of aleatory uncertainty in the forecasting of storm surge include the meteorological forcing and meteorological effects outside of the model domain. One way to reduce the effect of these uncertainties that has been proven efficient in meteorology and oceanography is to update the state of storm surge forecast models with available data using data assimilation.

The Kalman filter (KF; Kalman 1960) is a standard data assimilation method. The KF was originally designed for linear systems in which the model is not only used to forecast the system state, but also to determine the uncertainty of the estimate (Maybeck 1979). Because of its relative simplicity in implementation, the KF is attractive for many data assimilation problems. Since geophysical models are often nonlinear, ensemble Kalman filters (EnKFs) were introduced as credible alternatives to the KF for nonlinear data assimilation problems. EnKF methods provide an efficient and effective framework to propagate filter uncertainties forward in time while avoiding the need for linearization of complicated model dynamics (Evensen 2003). These methods can be classified into two main categories: the stochastic EnKF (Burgers et al. 1998; Houtekamer and Mitchell 1998), which essentially updates each forecast ensemble member with perturbed data using the KF correction step; and the square root deterministic EnKF, which, in contrast, updates the ensemble mean and sample covariance matrix (Anderson 2001; Bishop et al. 2001; Pham 2001; Hoteit et al. 2002; Whitaker and Hamill 2002; Luo and Moroz 2009). In the context of storm surge, considerable efforts have been made in recent years to improve the forecasts of storm surge models using the KF and its variants (Heemink and Kloosterhuis 1990; Sorensen and Madsen 2006; El Serafy and Mynett 2008; Butler et al. 2012).

The KF-based schemes are equivalent to sequential Bayesian filters in that they adopt Bayes's rule to update forecasted state variables and statistics using sequentially acquired data to form an analysis state and error covariance. These types of data assimilation methodologies require basic assumptions on the statistical properties of both the dynamical and observational systems. However, these statistical properties are often poorly known, which could result in weak performance of the Bayesian filter (Schlee et al. 1967). For example, if implemented straightforwardly, an EnKF with a relatively small ensemble size may produce

inaccurate estimations of covariance matrices (Whitaker and Hamill 2002). Model error statistics are also generally poorly known and often crude approximations are used in the filter instead. This often degrades filter performance and may even cause filter divergence. In practice, it is customary to conduct covariance inflation (Pham et al. 1998; Anderson and Anderson 1999) and localization (Hamill et al. 2001, 2009) to mitigate these problems. Covariance inflation is often implemented by postmultiplication of the forecast or analysis error covariance matrix by a constant factor larger than 1. Other inflation schemes have been also proposed through, for example, relaxation to prior (Zhang et al. 2004) or using a multimodel multiphysics approach (Meng and Zhang 2007).

In contrast with the KF, the so-called robust filters emphasize the robustness of their error estimates, so that they may have better tolerances to possible uncertainties in assimilation. The estimation strategies of robust filters are not based on the Bayes's rule. The  $H_\infty$  filter (HF) (Francis 1987; Simon 2006) is one such example. The HF does not assume that the statistical properties of the system being assimilated are exactly known. Instead, it allows for the possibility that the user may have incomplete information of the system. Consequently, rather than looking for the best possible estimates based on Bayes's rule, the HF employs an optimal robust strategy, namely, the minimax rule (Burger 1985), to update its forecast (or background) statistics. This robustness is of interest in practical situations. For example, in storm surge forecasting, the system model exhibits vastly different behavior before and during the occurrence of a hurricane or other extreme event. This is known as the "change of model regime" phenomenon, which causes further uncertainties and complications in traditional Bayes's rule-type data assimilation schemes (Bennet 1992; Hoteit et al. 2005b). It was shown in Luo and Hoteit (2011) that the HF could capture such a change better than the KF. This was supported by the fact that in general the estimation error of the HF grows with the uncertainties in assimilation at a finitely bounded rate (except for the special case when the HF reduces to the KF itself), while the estimation error of the KF does not possess such a guarantee. Recently an ensemble time-local  $H_\infty$  filter (EnTLHF) was proposed in Luo and Hoteit (2011) as an analogy to the EnKF for high-dimensional data assimilation problems. The EnTLHF was constructed based on the EnKF, and that the computational complexity of the EnTLHF is in general comparable to that of the EnKF. It was further demonstrated that an EnKF with a certain covariance inflation technique is in fact an EnTLHF.

In this paper, we have developed and implemented a robust EnTLHF filter for storm surge forecasting based on the singular evolutive Kalman filter (SEIK). The data assimilation system uses the Advanced Circulation (ADCIRC; Luettich and Westerink 2005) model and available data from hindcast studies of Hurricanes Katrina and Ike. Recently, similar experiments were reported with the standard SEIK filter demonstrating improved water level forecast using ADCIRC (Butler et al. 2012). In those experiments validated hindcast studies over the western North Atlantic and Gulf of Mexico were used to generate the observations from Hurricanes Katrina and Ike, and then a coarser mesh of the Gulf of Mexico was used for the forecast. In those studies, the application of the SEIK filter was able to improve the water level prediction up to 0.5 m in the areas of interest during the landfall events. However, while the SEIK filter proved capable of improving forecasts by significant amounts, the predicted water levels still lacked accuracy as compared to the truth, which motivates the current study. Here, the same set of experiments are repeated with an EnTLHF and results are then compared with those recently reported using the standard SEIK filter for different inflation factors (Butler et al. 2012). The results suggest that the SEIK filter implemented in the framework of the  $H_\infty$  filter strongly improves the accuracy of the short range forecasts of storm surge resulting from hurricanes as compared to the traditional SEIK filter.

The paper is organized as follows. The storm surge prediction model is described in section 2. Section 3 and its subsections describe the SEIK filter with inflation factor and the EnTLHF in the framework of SEIK filter. In section 4, the numerical results from the two different storm events are presented. Conclusions follow in section 5 including remarks of ongoing research to improve long-range forecasts of storm surge.

## 2. ADCIRC coastal circulation and storm surge model

Computational models of storm surge have recently seen an increase in attention because of the devastating 2005 hurricane season. As a result, a multi-institutional research team has undertaken the further development and application of the state-of-the-art ADCIRC model (Luettich and Westerink 2005) to model the shallow-water equations that describe large-scale water motions and depth-integrated horizontal flow. ADCIRC discretizes the shallow-water equations, consisting of a coupled Generalized Wave Continuity Equation (GWCE; Lynch and Gray 1979; Kinnmark 1985) and

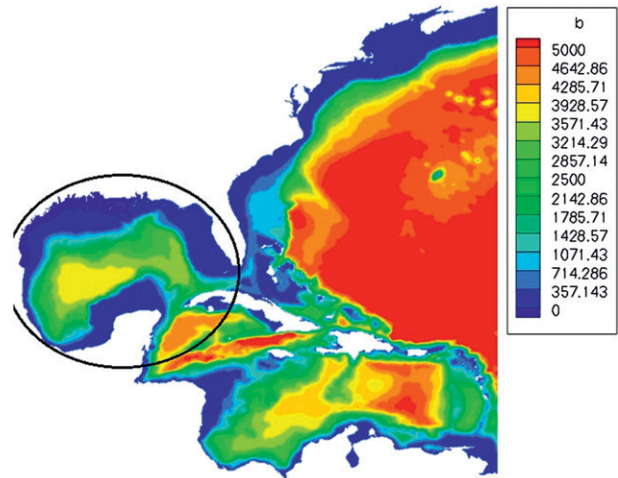


FIG. 1. Western North Atlantic domain and bathymetry (m). The Gulf of Mexico is circled in black.

momentum equations, using finite-element methods defined on unstructured meshes in space and finite-difference schemes in time. The resulting scheme is essentially explicit in time, except for the solution of a mass matrix in the GWCE discretization. Coupling of ADCIRC with a wind-wave model for capturing wave-induced initial states has recently been completed (Dietrich et al. 2011). The code has been parallelized for distributed memory, multicore computers, and has been demonstrated to achieve excellent scalability on these platforms (Tanaka et al. 2011). For the specific theory and numerical formulation of the ADCIRC model, we direct the interested reader to Luettich and Westerink (2005).

A typical spatial domain on which ADCIRC is used is the Gulf of Mexico possibly including the western North Atlantic, as seen in Fig. 1. The ADCIRC model has undergone extensive verification and validation, in particular by comparison with data from previous storms [e.g., see Westerink et al. (2008); Bunya et al. (2010); Dietrich et al. (2010); and Kennedy et al. (2011) for hindcast studies of hurricanes ranging in time from 1965 to 2008]. The extensive data collected and fundamental knowledge gained from these hindcast studies has led to the recently developed real-time forecasting system based on the ADCIRC model, called the ADCIRC Surge Guidance System (ASGS; Luettich and Westerink 2005). In this model, data on the hurricane track and forward speed, and wind characteristics (wind speed, central pressure, and radius-to-maximum winds), are obtained every 6 h from the National Weather Service. This input is used to generate a parametric wind field, which provides forcing to ADCIRC. To be useful to emergency managers, the storm surge model must compute predicted water levels along the coast within

a 1–2-h time window. Typically the output desired is maximum water elevations, measured over a given coastal region during the entire storm event, and time histories of water levels and/or currents at critical locations along the coast. In the numerical results, data were obtained from ADCIRC hindcast studies while the data assimilation experiments used the forecast mode of ADCIRC to propagate the state variables forward in time (see section 4 for more details).

### 3. Data assimilation methodology

We first recall the algorithm of the SEIK filter before describing its implementation in the framework of robust  $H_\infty$  filtering.

#### a. The singular evolutive interpolated Kalman filter

The SEIK filter uses a square root ensemble Kalman formulation (Tippett et al. 2003) to solve the Bayesian filtering problem (Pham 2001; Hoteit et al. 2005a; Nerger et al. 2012). As any Bayesian-based filter, it operates in two steps as a succession of a forecast with the model and an analysis step to update the forecast step with the new data. At the forecast step, the SEIK filter uses a Monte Carlo approach integrating an ensemble of states forward in time with the numerical model as a way to propagate the state probability distribution function. Only the Gaussian part of the resulting forecast distribution is then used in the analysis step, so a Kalman correction step is applied on the mean and the covariance matrix of the ensemble to update the forecast distribution with the new incoming data. The essential difference between a square root filter (e.g., the ETKF) and the SEIK is the following. A square root filter decomposes a background covariance as the product of a certain square root matrix and its transpose, and then updates the square root matrix to its analysis counterpart at the correction step. In contrast, the SEIK filter decomposes a background covariance in the form of  $\mathbf{L}\mathbf{L}^T$  (with more details provided later). Here  $\mathbf{L}$  is a matrix constructed based on the background ensemble and needs no update, while  $\mathbf{U}$  is a symmetric, positive semidefinite, matrix in the dimension of the ensemble size  $N$  minus 1 (i.e.,  $N - 1$ ), and needs to be updated at the correction step. The equivalence between the square root filter and the SEIK was shown in the recent paper Nerger et al. (2012).

After every analysis step, a resampling step is needed in the SEIK to generate a new analysis ensemble to start a new filtering cycle. Similar to the deterministic square root ensemble Kalman filters (Tippett et al. 2003), the ensemble adjustment Kalman filter (EAKF; Anderson 2001), and the ensemble transform Kalman filter (ETKF;

Bishop et al. 2001; Wang et al. 2004, 2006), the SEIK filter samples the analysis ensemble so that the ensemble mean and sample covariance matrix exactly match the computed analysis mean and its covariance matrix. In contrast with the EAKF and ETKF, in which resampling is deterministic, the SEIK filter samples the analysis ensemble randomly (Pham 2001; Hoteit et al. 2002). Stochastic resampling has the advantage of providing representative ensembles while randomly redistributing the variance among the ensemble members (Leith 1974; Houdekamer et al. 1996; Sakov and Oke 2008).

To describe the SEIK filter algorithm, consider the following discrete dynamical system:

$$\mathbf{X}^t(t_k) = M(t_k, t_{k-1})\mathbf{X}^t(t_{k-1}) + \boldsymbol{\eta}_k, \quad (3.1)$$

where  $\mathbf{X}^t(t_k)$  denotes the vector representing the true state of the system at time  $t_k$ ,  $M(t, s)$  is the state transition operator that takes as inputs the state at time  $t$  and outputs the state at time  $s$ , and  $\boldsymbol{\eta}_k$  is the system noise with covariance matrix  $\mathbf{Q}_k$ . At time  $t_k$ , the observed data vector is given by

$$\mathbf{Y}_k = H_k \mathbf{X}^t(t_k) + \epsilon_k. \quad (3.2)$$

Here  $H_k$  is the observation operator and  $\epsilon_k$  is the observational noise with covariance matrix  $\mathbf{R}_k$ . Note that both the transition and the observational operators  $M$  and  $H$  can be nonlinear.

One can start the SEIK filter from a forecast or an analysis step, depending on whether data are available at the initial time or not. At the initial time, an analysis (or forecast) ensemble is needed. The procedure to generate an initial ensemble is discussed in section 4. Assuming now that an analysis ensemble of  $N$  members  $\mathbf{X}_i^a(t_{k-1})$  has been computed at time  $t_{k-1}$ , a forecast step is then applied as described below.

#### 1) MONTE CARLO FORECAST STEP

The numerical model is integrated  $N$  times from  $t_{k-1}$  to  $t_k$  starting from  $\mathbf{X}_i^a(t_{k-1})$  to compute the forecast ensemble members  $\mathbf{X}_i^f(t_k)$  as

$$\mathbf{X}_i^f(t_k) = M(t_k, t_{k-1})\mathbf{X}_i^a(t_{k-1}) + \boldsymbol{\eta}_k^i, \quad (3.3)$$

where  $\boldsymbol{\eta}_k^i$  are sampled from the distribution of the system noise.<sup>1</sup> The forecast state is then taken as the forecast ensemble average:

<sup>1</sup>The formula works for additive model error, as assumed in the present study. Nonadditive model error can be accounted for in a very similar way by inserting random noise simulated from the distribution of the model error, if known, as discussed by Evensen (2003).

$$\mathbf{X}^f(t_k) = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i^f(t_k), \quad (3.4)$$

and the associated forecast error covariance is approximated by

$$\mathbf{P}^f(t_k) = \frac{1}{N} \sum_{i=1}^N [\mathbf{X}_i^f(t_k) - \mathbf{X}^f(t_k)][\mathbf{X}_i^f(t_k) - \mathbf{X}^f(t_k)]^T. \quad (3.5)$$

For convenience in the correction step, one can rewrite (3.5) as

$$\mathbf{P}^f(t_k) = \mathbf{L}_k [N\mathbf{T}^T\mathbf{T}]^{-1} \mathbf{L}_k^T, \quad (3.6)$$

where  $\mathbf{L}_k$  is  $n \times (N - 1)$  with the  $i$ th column given by  $\mathbf{X}_i^f(t_k) - \mathbf{X}^f(t_k)$  and  $\mathbf{T}$  is a  $N \times (N - 1)$  full-rank matrix with zero column sums, given by Hoteit et al. (2002):

$$\mathbf{T} = \begin{pmatrix} \mathbf{I}_{N-1} \\ 0 \end{pmatrix} - \frac{1}{N} [\mathbf{1}_N, \dots, \mathbf{1}_N]_{N \times (N-1)},$$

with  $\mathbf{I}_{N-1}$  being the  $(N - 1)$ -dimensional identity matrix, and  $\mathbf{1}_N$  being the  $N$ -dimensional vector with all its elements being 1. With inflation, the forecast error covariance is taken as

$$\mathbf{P}^f(t_k) = \mathbf{L}_k [\lambda^{-1} N \mathbf{T}^T \mathbf{T}]^{-1} \mathbf{L}_k^T, \quad (3.7)$$

with  $\lambda$  the inflation factor<sup>2</sup> usually chosen slightly larger than 1.

### 2) KALMAN CORRECTION STEP

New incoming data are used to update the forecast state  $\mathbf{X}^f(t_k)$  with the standard Kalman correction:

$$\mathbf{X}^a(t_k) = \mathbf{X}^f(t_k) + \mathbf{G}_k [\mathbf{Y}_k - H_k \mathbf{X}^f(t_k)], \quad (3.8)$$

where the Kalman gain matrix  $\mathbf{G}_k$  is given by

$$\mathbf{G}_k = \mathbf{L}_k \mathbf{U}_k (\mathbf{HL})_k^T \mathbf{R}_k^{-1}, \quad (3.9)$$

with  $(\mathbf{HL})_k^T = [H_k(\mathbf{I}_k^1), \dots, H_k(\mathbf{I}_k^{N-1})]$  and  $\mathbf{I}_k^i$  the columns of  $\mathbf{L}_k$ . The matrix  $\mathbf{U}_k$  in (3.9) is updated using

$$\mathbf{U}_k^{-1} = \lambda^{-1} N \mathbf{T}^T \mathbf{T} + (\mathbf{HL})_k^T \mathbf{R}_k^{-1} (\mathbf{HL})_k. \quad (3.10)$$

The analysis filter error covariance matrix is then

$$\mathbf{P}^a(t_k) = \mathbf{L}_k \mathbf{U}_k \mathbf{L}_k^T. \quad (3.11)$$

### 3) RESAMPLING STEP

To start a new filtering cycle, one has to resample an ensemble of  $N$  analysis states from the estimated analysis mean  $\mathbf{X}^a(t_k)$  and covariance matrix  $\mathbf{P}^a(t_k)$ . Rewriting  $\mathbf{P}^a(t_k)$  as

$$\mathbf{P}^a(t_k) = \mathbf{L}_k (\mathbf{C}_k^{-1})^T \mathbf{\Omega}_k^T \mathbf{\Omega}_k \mathbf{C}_k^{-1} \mathbf{L}_k^T. \quad (3.12)$$

Here  $\mathbf{C}_k$  is the Cholesky decomposition of  $\mathbf{U}_k^{-1} = \mathbf{C}_k \mathbf{C}_k^T$  and  $\mathbf{\Omega}_k$  is a random orthonormal matrix  $\mathbf{\Omega}_{k-1}$  with zero column sums,<sup>3</sup> which plays the same role as the ‘‘centring matrix’’ generated by the spherical simplex method in Wang et al. (2004, 2006). The new analysis ensemble members are then taken as

$$\mathbf{X}_i^a(t_k) = \mathbf{X}^a(t_k) + \sqrt{N} \mathbf{L}_k (\mathbf{\Omega}_{k,i} \mathbf{C}_k^{-1})^T, \quad 1 \leq i \leq N, \quad (3.13)$$

where  $\mathbf{\Omega}_{k,i}$  denotes the  $i$ th row of  $\mathbf{\Omega}_k$ . One can see from (3.12) that sampling the analysis members in this way ensures that the sample mean and sample covariance of the ensemble are exactly  $\mathbf{X}^a(t_k)$  and  $\mathbf{P}^a(t_k)$ , respectively.

### b. The $H_\infty$ SEIK filter

In a recent work, Luo and Hoteit (2011) proposed the EnTLHF for large-scale nonlinear data assimilation problems. The EnTLHF is an ensemble implementation of the time-local  $H_\infty$  filter (TLHF) derived in Luo and Hoteit (2011). In the TLHF, one aims at finding an analysis  $\tilde{\mathbf{X}}^a(t_k)$  at the  $k$ th assimilation cycle that satisfies the inequality

$$\begin{aligned} \|\mathbf{X}^t(t_k) - \tilde{\mathbf{X}}^a(t_k)\|_{\mathbf{S}_k}^2 &\leq \frac{1}{\gamma_k} (\|\mathbf{X}^t(t_k) - \tilde{\mathbf{X}}^f(t_k)\|_{[\tilde{\mathbf{P}}^f(t_k)]^{-1}}^2 \\ &+ \|\boldsymbol{\eta}_k\|_{\mathbf{Q}_k^{-1}}^2 + \|\epsilon_k\|_{\mathbf{R}_k^{-1}}^2) \end{aligned} \quad (3.14)$$

in the worst possible cases. Here,  $\mathbf{S}_k$  is a positive definite weight matrix that is chosen by the filter designer,  $\tilde{\mathbf{X}}^f(t_k)$  and  $\tilde{\mathbf{P}}^f(t_k)$  are the estimated background state and covariance matrix [similar to  $\mathbf{X}^f(t_k)$  and  $\mathbf{P}^f(t_k)$  in the SEIK],

<sup>2</sup> Inflation was originally introduced in the form of a forgetting factor in the SEIK filter, which is the inverse of the inflation factor, in the context of the SEIK filter.

<sup>3</sup> A procedure for computing an  $\mathbf{\Omega}_k$  using Householder matrices can be found in Hoteit et al. (2002).

and  $\gamma_k$  is a scalar, called the local performance level. Thus, the three terms,  $\|\mathbf{X}^i(t_k) - \tilde{\mathbf{X}}^f(t_k)\|_{[\tilde{\mathbf{P}}^f(t_k)]^{-1}}^2$ ,  $\|\boldsymbol{\eta}_k\|_{\mathbf{Q}_k^{-1}}^2$ , and  $\|\epsilon_k\|_{\mathbf{R}_k^{-1}}^2$ , represent the “energy” of the uncertainties in specifying the background (analogous to the uncertainty in specifying the initial conditions), the model error, and the observation error, respectively. By assuming that the analysis  $\tilde{\mathbf{X}}^a(t_k)$  is a linear update of the background  $\tilde{\mathbf{X}}^f(t_k)$ , one obtains a filter solution that resembles the celebrated Kalman filter [see Simon (2006) for the deduction]. The filter is then robust in the sense that it bounds the state estimate error by the total uncertainties in the system.

Luo and Hoteit (2011) showed that when the system is linear, the KF and the TLHF have basically the same algorithms except that it introduces the extra term  $-\gamma_k \mathbf{S}_k$  in the inverse covariance update formulas, which becomes

$$[\tilde{\mathbf{P}}^a(t_k)]^{-1} = [\tilde{\mathbf{P}}^f(t_k)]^{-1} + (H_k)^T (\mathbf{R}_k)^{-1} H_k - \gamma_k \mathbf{S}_k, \quad (3.15)$$

subject to the constraint  $[\tilde{\mathbf{P}}^a(t_k)]^{-1} \geq 0$ . When  $\gamma_k = 0$ , then the TLHF simply reduces to the Kalman filter.

For analogy to the EnKF, the EnTLHF is an ensemble implementation of the TLHF. When  $\gamma_k > 0$ , then  $-\gamma_k \mathbf{S}_k < 0$  in (3.15), so that the estimated covariance matrix  $\tilde{\mathbf{P}}^a(t_k)$  is larger than that with  $\gamma_k = 0$ . In this sense, the presence of the extra term  $-\gamma_k \mathbf{S}_k$  in (3.15) introduces inflation to the covariance matrix, similar to that in the EnKF. In fact, as shown in Luo and Hoteit (2011), the EnTLHF provides a robust strategy for conducting covariance inflation, while an EnKF with certain covariance inflation is essentially an EnTLHF.

In the context of the SEIK filter, one can estimate the background mean  $\tilde{\mathbf{X}}^f(t_k)$  and covariance  $\tilde{\mathbf{P}}^f(t_k)$  in the EnTLHF according to (3.4) and (3.6), respectively. Let  $\Delta_k = \mathbf{L}_k [\mathbf{N}^T \mathbf{T}^T]^{-1} \mathbf{L}_k^T$  such that  $\tilde{\mathbf{P}}^f(t_k) = \Delta_k$  by (3.6). If in (3.15) we set

$$\gamma_k \mathbf{S}_k = (\Delta_k)^{-1} - (\lambda \Delta_k)^{-1},$$

with  $\lambda$  being the inflation factor as in (3.7), and write  $\tilde{\mathbf{P}}^a(t_k)$  in the form of  $\mathbf{L}_k \mathbf{U}_k \mathbf{L}_k^T$  [with the dimension of  $\mathbf{U}_k$  being  $(N-1) \times (N-1)$ ], then it can be shown that (3.15) reduces to the update in (3.7) in the SEIK filter. However, based on (3.15), we can also derive other update schemes with inflation. For instance, by choosing  $\mathbf{S}_k = (\mathbf{L}_k \mathbf{L}_k^T)^{-1}$  and  $\boldsymbol{\Psi}_k = [\mathbf{N}^T \mathbf{T}^T]^{-1}$ , we have

$$(\mathbf{U}_k)^{-1} = \boldsymbol{\Psi}_k^{-1} + (H_k \mathbf{L}_k)^T \mathbf{R}_k^{-1} (H_k \mathbf{L}_k) - \gamma_k \mathbf{I}, \quad (3.16)$$

where  $\mathbf{I}$  is the identity matrix with a suitable dimension. Let  $\boldsymbol{\Phi}_k^{-1} = \boldsymbol{\Psi}_k^{-1} + (H_k \mathbf{L}_k)^T \mathbf{R}_k^{-1} (H_k \mathbf{L}_k)$ , and suppose that through spectral decomposition,  $\boldsymbol{\Phi}_k^{-1}$  can be decomposed as  $\boldsymbol{\Phi}_k^{-1} = \mathbf{V}_k \mathbf{D}_k \mathbf{V}_k^T$ , where  $\mathbf{V}_k$  is a matrix consisting of the normalized eigenvectors of  $\boldsymbol{\Phi}_k^{-1}$ , and  $\mathbf{D}_k$  is a diagonal matrix with  $k$ th diagonal entry corresponding to the  $k$ th eigenvalue  $\sigma_{k,i}$  ( $\sigma_{k,j} \leq \sigma_{k,i}$  if  $j > i$ ). To make sure that  $(\mathbf{U}_k)^{-1}$  is positive semidefinite, we require  $\gamma_k \leq \sigma_{k,N-1}$ . One can choose  $\gamma_k = c \sigma_{k,N-1}$ , with  $c \in [0, 1)$  being a scalar, so that  $(\mathbf{U}_k)^{-1} = \mathbf{V}_k [\mathbf{D}_k - c \sigma_{k,N-1} \mathbf{I}] \mathbf{V}_k^T$ . As a result,

$$\begin{aligned} \mathbf{U}_k &= \mathbf{V}_k (\mathbf{D}_k - c \sigma_{k,N-1} \mathbf{I})^{-1} \mathbf{V}_k^T \\ &= \mathbf{V}_k \text{diag} \left( \frac{1}{\sigma_{k,1} - c \sigma_{k,N-1}}, \dots, \frac{1}{(1-c) \sigma_{k,N-1}} \right) \mathbf{V}_k^T, \end{aligned} \quad (3.17)$$

and the eigenvalues  $1/(\sigma_{k,i} - c \sigma_{k,N-1})$  ( $i = 1, \dots, N-1$ ) of  $\mathbf{U}_k$  are larger than  $1/\sigma_{k,j}$ , their counterparts of  $\boldsymbol{\Phi}_k$ . Therefore, the choice  $\mathbf{S}_k = (\mathbf{L}_k \mathbf{L}_k^T)^{-1}$  and  $\gamma_k = c \sigma_{k,N-1}$  leads to the inflation of the eigenvalues of  $\mathbf{U}_k$ . In this regard, our proposed inflation scheme is similar to that in Ott et al. (2004), in which the authors essentially introduce covariance inflation to the analysis covariance, say  $\mathbf{P}^a$ , by replacing  $\mathbf{P}^a$  by  $\mathbf{P}^a + \delta \mathbf{I}$  [cf. (42) of Ott et al. (2004)], where  $\delta > 0$  plays a role similar to the covariance inflation factor, and  $\mathbf{I}$  is the identity matrix with a suitable dimension. In doing so, all the eigenvalues of  $\mathbf{P}^a$  are increased at the same pace, while in (3.17) the eigenvalue increments of  $\mathbf{U}_k$  are nonuniform and actually adaptive with time (as it is so for the eigenvalue  $\sigma_{k,N-1}$  used for inflation). In fact, with some algebra, it can be shown that the eigenvalues of  $\boldsymbol{\Phi}_k$  receive inflation proportional to magnitude. The eigenvectors corresponding to larger eigenvalues may represent more influential directions of the model in data assimilation. As a result (at the next assimilation cycle), the observation corresponding to these directions may receive relatively more weights in update. If the true model dynamics undergoes a certain change that is not well captured by the model in data assimilation, then by putting more (relative) weights onto the observation that correspond to the relatively influential directions, the update formula will rely more on the corresponding observation, and tend to reduce the impact of model change by giving less (relative) weight to the possibly not-so-good background. For the specific problem of storm surge forecasting, the nonuniform inflation is expected to improve the filter performance during the period of hurricane landfall.

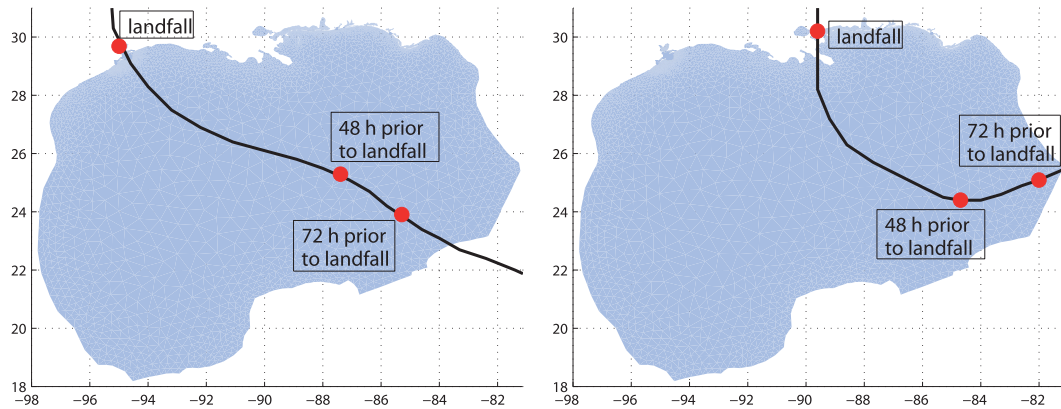


FIG. 2. (left) Track of Hurricane Ike through the Gulf of Mexico. The circles with annotations are the locations of landfall at 0710 UTC 13 Sep 2008 and the locations of the hurricane at approximately 48 and 72 h before landfall. (right) Track of Hurricane Katrina through the Gulf of Mexico. The circles with annotations are the locations of landfall at 1110 UTC 29 Aug 2005 and the locations of the hurricane at approximately 48 and 72 h before landfall.

#### 4. Numerical experiments

We consider Hurricanes Ike and Katrina as two particular test cases in our numerical experiments. Hurricane Ike became a category 4 hurricane as it traveled through the Atlantic, Caribbean, and Gulf of Mexico in September 2008, and was a category 2 hurricane upon making landfall along the upper Texas coast in the early morning hours of 13 September, see left plot in Fig. 2. Ike was a devastating storm causing massive financial damages, and was also responsible for nearly 200 deaths (Berg 2009). Hurricane Katrina was the most devastating hurricane in recent years as a category 3 hurricane upon its second landfall and is held responsible for more than 1800 deaths (Knabb et al. 2006). Katrina formed in August of 2005 over the Bahamas, made its first landfall in southern Florida, traveled through the Gulf of Mexico, and made its second landfall in southeast Louisiana on the morning of 29 August (see Fig. 2).

The measurement data were obtained from high-resolution hindcast studies. These studies have been shown to match the actual measured data from gauges located throughout the Gulf for a number of recent hurricanes. The ADCIRC hindcast runs use a time step of 1 s on high-resolution grids of the same domain covering the Gulf of Mexico and the western north Atlantic seaboard (see Fig. 1). The hindcasts use data-assimilated winds and atmospheric pressure fields provided by Ocean Weather, Inc. (OWI). The Hurricane Ike hindcast was run on a grid of 3 322 439 nodes corresponding to 6 615 381 elements. The Hurricane Katrina hindcast was run on a grid of 5 035 113 nodes corresponding to 9 945 623 elements. The measurement data are then obtained every 2 h from these hindcast studies and is

used for the assimilation experiments. In all the experiments, we set the standard deviation of the measurement noise to produce an assumed 95% confidence interval of  $\pm 0.01$  m.

The data assimilation experiments are run using a time step of 10 s on a grid of 8006 nodes and 14 269 elements covering the Gulf of Mexico (see Fig. 3). Separate observation stations are used for tracking Ike and Katrina (see Figs. 4–5). This is done mainly because hurricanes Ike and Katrina have distinct tracks (see plots in Fig. 2). By using separate domains, discretizations, wind forcing, and time steps for the simulation of true data versus the data assimilation experiments, we

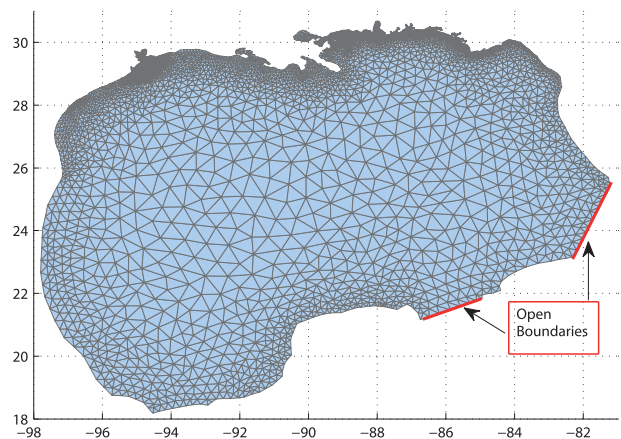


FIG. 3. Discretization of the Gulf of Mexico domain containing 8006 nodes and 14 269 elements. Open boundaries are denoted by bold boundary lines. All other boundaries are land. The open boundaries are forced by the five tidal constituents: K1, O1, P1, M2, and S2. The land boundaries are reflective.

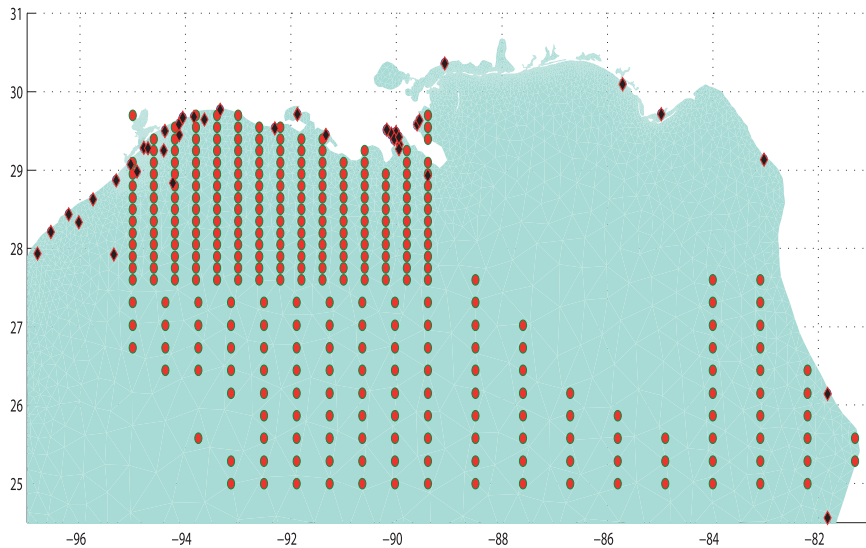


FIG. 4. The 371 hypothetical observation stations used for Hurricane Ike simulations. The red-filled circles denote an array of stations where data are simulated from the hindcast study. The black-filled circles correspond to 43 existing measurement sites. These 43 observation stations collected data during the storm and were used to validate the hindcast study.

avoid simulating observation data within the same numerical framework as used in the forecasting/data assimilation framework. The main differences between the hindcast simulations taken as truth and the simulations used in the data assimilation runs are summarized in Table 1. For all the simulations the data are assimilated after every 2-h forecasts of the ensemble members. A detailed description of these data assimilation experiments can be found in Butler et al. (2012).

We make use of an empirical orthogonal function (EOF) analysis as presented in Hoteit et al. (2001). Driving ADCIRC only by tidal forcing terms eliminates all transient behavior in the state within a few days. We, therefore, run the ADCIRC model driven only by tidal forcing terms for 60 days and store the state every 5 h. The perturbation of these states from their mean value is used to define a sample covariance matrix from which the ensemble members can be initialized. In this way, we

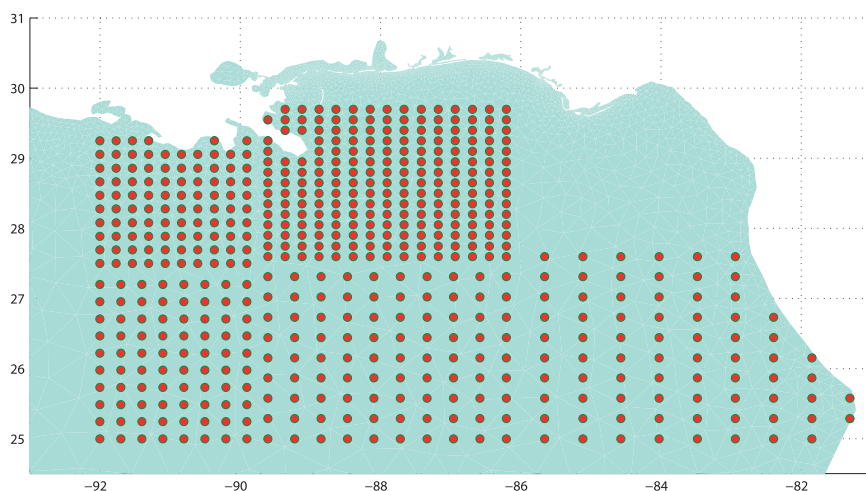


FIG. 5. The 559 hypothetical observation stations used for Hurricane Katrina simulations. The red-filled circles denote an array of stations where data are simulated from the hindcast study.



TABLE 1. Summary of differences between simulations for hindcast (truth) simulation used to generate data and simulations used in data assimilation forecasting experiments.

	Data assimilation	Truth
Domain	Gulf of Mexico	Western North Atlantic
Avg mesh element size	98 km <sup>2</sup>	1.34 km <sup>2</sup>
Time step	10 s	1 s
Wind field	Dynamic Holland	OWI
Bottom friction formulation	Chezy	Hybrid

initialize the ensemble members using a covariance  $\mathbf{P}$  from a physically meaningful feature space. The ratio  $\sum_{j>N-1} \sigma_j / \text{Tr}(\mathbf{P})$  (with  $\sigma_j$  being the  $j$ th eigenvalue of  $\mathbf{P}$ ) represents the relative error in the square  $L^2$  norm of using approximations to the state in the  $N - 1$ -dimensional feature space and is useful in determining the ensemble size  $N$  given a prescribed error tolerance, which we also refer to as the percentage of inertia retained by the EOFs. In the experiments below, we chose to retain about 90% of the spatial variability of this sequence of states using the EOF analysis, resulting in an ensemble size of  $N = 10$ .

To quantify performances, a root-mean-square error (rmse) metric is used. We are interested in the ability to forecast the maximum coastal surge. We are also interested in forecasts of water elevations at particular times, specifically the times leading up to landfall. We also present figures plotting the pointwise errors for both maximum water level forecasts and water level forecasts during the landfall event for each hurricane. These figures illustrate the important improvement in errors obtained using the  $H_\infty$  SEIK compared to the traditional SEIK filter for 2-h forecasts of the storm surge resulting from these hurricanes.

#### a. Hurricane Ike

For the Hurricane Ike simulations, the spinup period is set to 1 day starting at 0000 UTC 9 September 2008 and ending at 0000 UTC 10 September 2008. The Ike simulations end at 0000 UTC 14 September 2008, almost 18 h after landfall occurred at 0710 UTC 13 September 2008. The “spinned up” state is taken as the mean state of the system. The data are assimilated using 371 assimilation stations (see Fig. 4) after every 2-h forecasts of the ensemble members starting at 0200 UTC 10 September 2008 until 0000 UTC 14 September 2008 resulting in a total of 48 assimilation cycles being computed.

Tables 2–3 give average rmse of the maximum water level forecasts for the Ike simulations using SEIK and  $H_\infty$  SEIK filters with different values of inflation factor  $\lambda$

TABLE 2. The average rmse of the maximum water level forecasts are shown for Ike simulations using SEIK filter with different values of inflation factor  $\lambda$ . The second column shows the average rmse of the maximum water levels forecasted in the landfall area. The third column shows the average rmse of the maximum forecasted water levels where the hindcast data (truth) showed maximum water elevations greater than 3 m (2927 out of 8006 nodes). The boldface font indicates the best values of rmse corresponding to inflation factor.

Inflation factor $\lambda$	Coastal rmse	Rmse > 3 m
—	1.92	1.91
1.0	1.68	1.65
1.1	1.59	1.62
1.2	1.45	1.46
1.4	1.58	1.61
1.5	1.62	1.65
<b>1.6</b>	<b>1.38</b>	<b>1.42</b>
1.7	1.47	1.51
2.	1.52	1.54

and factor  $c$ , respectively. The second column in both tables presents the average rmse of the maximum water levels forecasted in the landfall area (29°–29.8°N, 94.4°–95.25°W, see Fig. 6). The rmse near the landfall areas with the SEIK filter varies from within 1.38 to 1.68 m, with the lowest rmse performances obtained using  $\lambda = 1.6$ . Overall, the SEIK filter is able to reduce the rmse by more than 27% as compared to the initial forecasted average rmse when no data are assimilated. Improvements are more pronounced in the  $H_\infty$  SEIK filter, which significantly reduces the average rmse of the maximum forecasted water elevations to a value of 0.80 m with the factor  $c = 0.7$ . This is an overall improvement of more than 58% as compared to the initial value of the rmse without data assimilation. The third column in Tables 2–3

TABLE 3. The average rmse of the maximum water level forecasts are shown for Ike simulations using EnTLHF with different values of inflation factor  $c$ . The second column shows the average rmse of the maximum water levels forecasted in the landfall area (see Fig. 6). The third column shows the average rmse of the maximum forecasted water levels where the hindcast data (truth) showed maximum water elevations greater than 3 m (2927 out of 8006 nodes). The boldface font indicates the best values of rmse corresponding to factor  $c$ .

Factor $c$	Coastal rmse	Rmse > 3 m
—	1.92	1.91
0.1	1.43	1.38
0.2	1.40	1.42
0.3	1.47	1.50
0.4	1.34	1.36
0.5	1.30	1.33
0.6	1.17	1.10
<b>0.7</b>	<b>0.80</b>	<b>0.87</b>
0.8	1.35	1.38

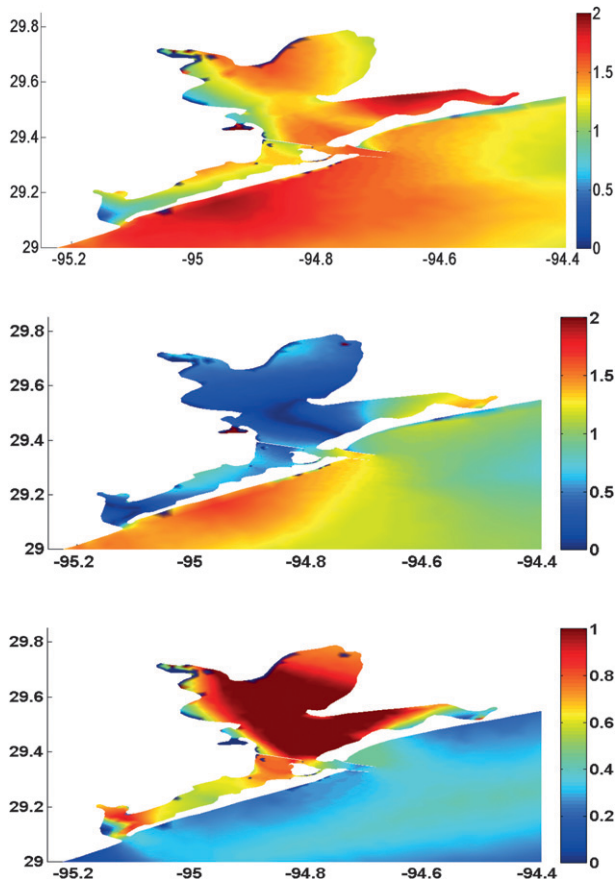


FIG. 6. Plots of errors in maximum free surface water level predictions (m) from true state for (top) simulation using SEIK filter with ( $\lambda = 1.6$ ) and (middle) simulation using  $H_\infty$  SEIK filter with ( $c = 0.7$ ). (bottom) Plot of differences between maximum free surface water level predictions (m) of simulation using SEIK filter with ( $\lambda = 1.6$ ) and simulation using  $H_\infty$  SEIK filter with ( $c = 0.7$ ). Note the different color bar range in (bottom).

shows the average rmse of the maximum water elevation forecasts at all nodes that recorded up to 60% of the maximum surge from the true forecast (i.e., we consider the rmse at roughly 36.5% of the nodes where the

maximum water elevation forecast of the hindcast study exceeded 3 m). A similar trend is observed and  $H_\infty$  with factor  $c = 0.7$  leads to an overall improvement of 56%.

The inflation factor  $c$  plays a role similar to that of the inflation factor  $\lambda$  in the standard SEIK filter. While a good value can improve filter performance by imposing more weight on the assimilated observations, which could be critical in times of regime change, a bad one can degrade the filter performance. The focus of this work is to demonstrate that the combination of the  $H_\infty$  filter within the SEIK framework can produce more accurate results. We, therefore, compare the performances of the best-case scenarios obtained with these filters when their inflation factors are chosen through a systematic trial-and-error search. We focus the remaining discussion on the best configurations of the SEIK and  $H_\infty$  SEIK filters (i.e., with  $\lambda = 1.6$  and  $c = 0.7$ ).

Figure 6 plots the errors from the true forecast of maximum water elevation in the SEIK and  $H_\infty$  SEIK filters. Figure 6 also shows the improvements to the SEIK filter predictions of maximum water elevation computed from  $H_\infty$  SEIK filter. We see in the plot of Fig. 6 that the  $H_\infty$  SEIK filter offers a significant improvement to forecasts up to 1 m at many locations as compared to the SEIK filter.

Figure 7 plots the averaged rmse of water elevations (in meters) in the landfall area (see Fig. 8) as obtained from the SEIK and  $H_\infty$  SEIK filters. The time interval here is chosen to clearly show the errors during the landfall period (i.e., from 1800 UTC 12 September to 1800 UTC 13 September 2008). The  $H_\infty$  SEIK filter was able to significantly reduce the errors, keeping them at an acceptable level at the landfall period (i.e., from 0600 UTC 13 September to 1000 UTC 13 September 2008). This is an important improvement indicating that the  $H_\infty$  SEIK filter can produce more accurate short-range forecasts of extreme events.

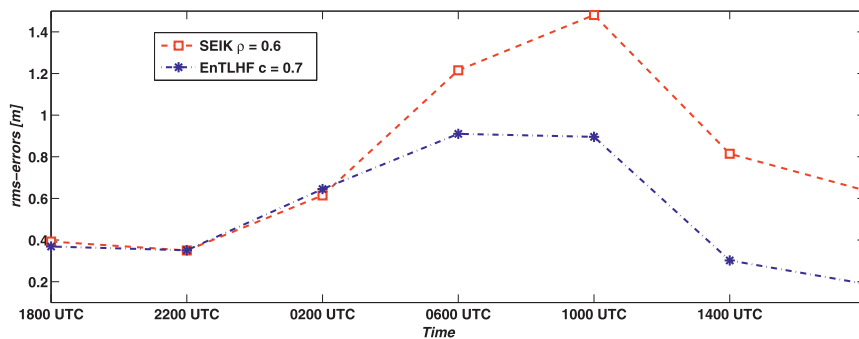


FIG. 7. The plot of averaged rmse of water elevations (m) for Ike simulations in the landfall area (see Fig. 8) using SEIK filter with  $\lambda = 1.6$  and  $H_\infty$  SEIK filter with  $c = 0.7$ . The time interval for these plots is from 1800 UTC 12 Sep to 1800 UTC 13 Sep 2008.

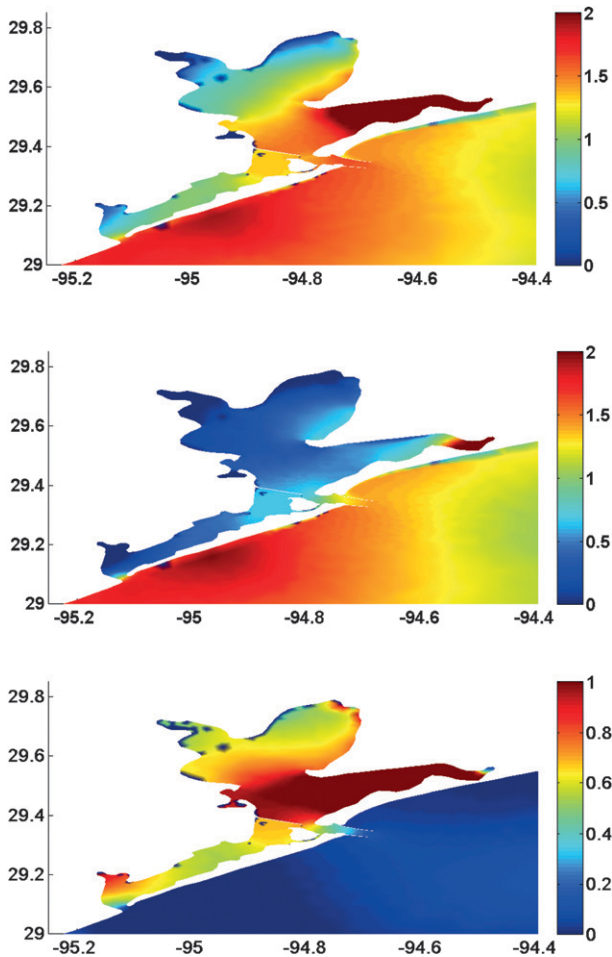


FIG. 8. Plots of free surface elevation error (m) at 0800 UTC 13 Sep 2008 from truth for (top) SEIK filter using  $\lambda = 1.6$  and (middle)  $H_\infty$  SEIK filter ( $c = 0.7$ ). (bottom) Plot of differences (m) of free surface elevation forecasts at 0800 UTC 13 Sep 2008 between  $H_\infty$  SEIK ( $c = 0.7$ ) and SEIK filter ( $\lambda = 1.6$ ). Note the different color bar range in (bottom).

Figure 8 plots the errors of the true forecast of water elevations at 0800 UTC 13 September 2008, which is an hour after Ike made landfall at 0710 UTC on the same day as they result from both filters. These graphs clearly show that  $H_\infty$  SEIK filter simulation is in better agreement with the data, predicting up to approximately 0.6–1.0 m of higher water elevation than the SEIK filter in majority of the areas across the coastline.

Finally, Fig. 9 plots the hydrographs of simulated data from the Ike hindcast results at four stations. The stations shown in this figure correspond to just four of the many existing stations that collected true data during Hurricane Ike. In these hydrographs the stars denote the true measurements at the assimilation times, the plus signs denote the forecasted results, and the circles are the analyzed states data for the  $H_\infty$  SEIK filter with

$c = 0.7$ . We also plot the 95% confidence intervals of the forecasted data, which are computed using (3.6). We observe that the analysis steps bring the model closer to the truth in majority of the cases. In general, the filter underestimates the level of the surge at the early assimilation window. This is most likely due to the coarse scale of the discretization so the dissipation in the model is more pronounced. Overall, the estimated uncertainties are quite reasonable.

### b. Hurricane Katrina

For Hurricane Katrina simulations, the spinup is set to 6 h starting at 0000 UTC 25 August 2005 and ending at 0600 UTC 25 August 2005. The spinned-up state is again taken as the mean state of the system. The first data assimilation cycle is computed on a 2-h forecast of the initial ensemble members. The data are assimilated using 559 assimilation stations (see Fig. 5) every 2 h until 0600 UTC 30 August 2005 resulting in a total of 60 assimilation cycles being computed. The Katrina storm surge was significantly stronger, from 4 m to approximately 6.6 m over a smaller area as determined by the hindcast. Hurricane Ike, on the other hand, had a 3–5-m storm surge over a larger geographic area. For hurricane Ike, 36.5% of the nodes recorded water elevations within 60% of the overall maximum value, but for Katrina, only 5.1% of the nodes recorded water elevations within 60% of the overall maximum value. Accurate estimation of extreme storm surge over geographically small areas is an important and numerically challenging problem of great interest.

For the Katrina simulations we will investigate the improvements using similar rmse metrics. Tables 4–5 give the average rmse of the maximum water level forecasts for the Katrina simulations using the SEIK filter and the  $H_\infty$  SEIK filter with different values of inflation factors  $\lambda$  and factor  $c$ , respectively. The second column in Tables 4–5 presents the average rmse of the maximum water elevation forecasts at all 405 out of 8006 nodes where forecasts of the hindcast study exceeded 4 m. The rmse of the  $H_\infty$  SEIK filter varies within the range 0.72–1.68 m as compared to 1.42–2.09 m in the SEIK filter. Using  $c = 0.6$ ,  $H_\infty$  SEIK is able to reduce the rmse by 70% as compared to the initial forecasted average rmse when no data are assimilated. This is an important improvement over the best performance obtained with the SEIK filter with standard inflation. The third column of Tables 4–5 shows the average rmse of the maximum water elevation forecasts at all nodes on which the hindcast exceeded 5 m (163 out of 8006). The improved performance of the  $H_\infty$  SEIK using  $c = 0.6$  is even more pronounced here with the average rmse reduced to a value of 0.45 as compared to 2.90 when no data assimilation

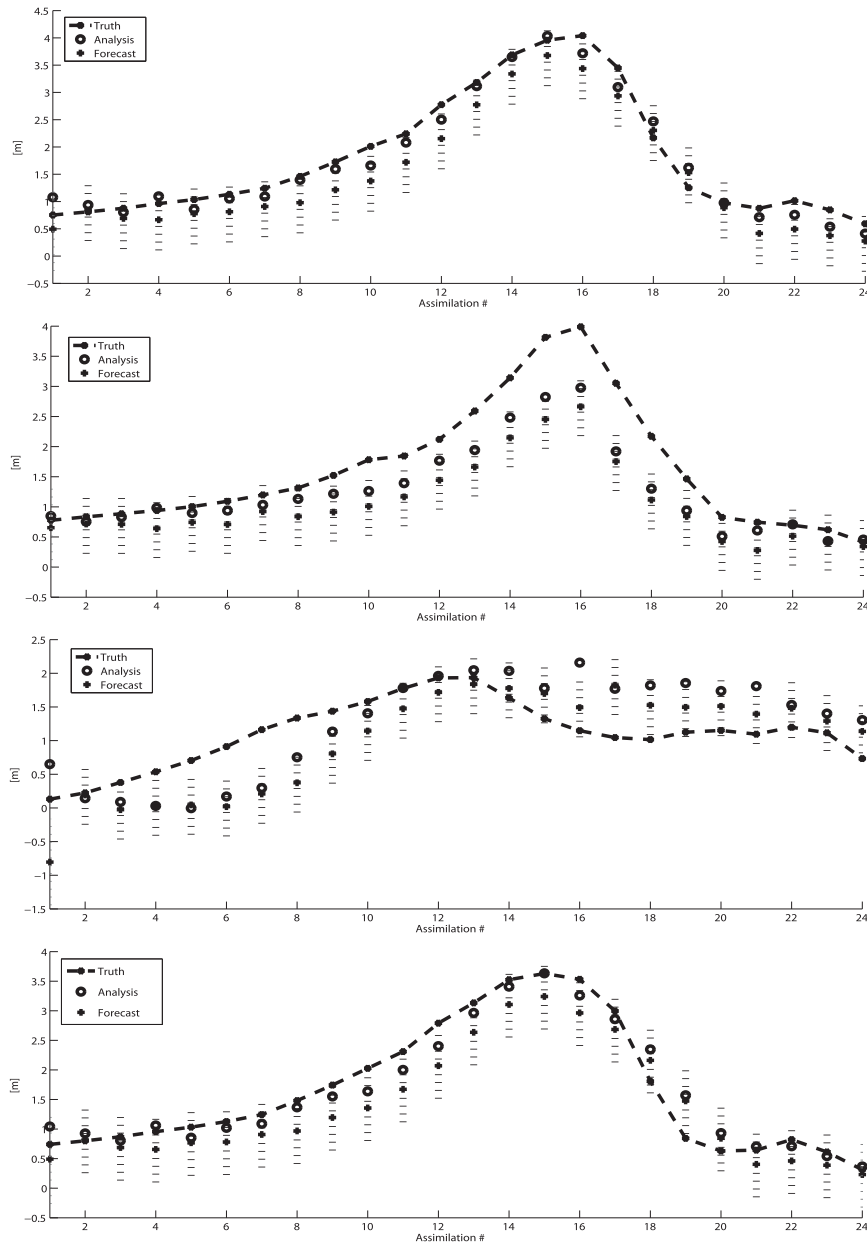


FIG. 9. Plots of station data (m) vs assimilation cycle where data are used to update the forecast. The stars mark data that are simulated (and then assimilated every 2 h beginning at 0200 UTC 12 Sep 2008 and ending at 0000 UTC 14 Sep 2008 using  $H_\infty$  SEIK filter), the analyzed state (circles), and forecast state (plus signs) with 95% confidence intervals (vertical dashed lines centered at plus signs).

was performed. In our experiments we noticed that the results of the  $H_\infty$  SEIK filter depend on the parameter  $c$ . Several runs could, therefore, be required in practice to find a good value of  $c$ .

Figure 10 shows the plots of the errors from the true forecast of maximum water elevation in the SEIK filter with an inflation factor  $\lambda = 1.2$ , and  $H_\infty$  SEIK with  $c = 0.6$  for forecasts of maximum water elevations

in the nearshore areas. We see from these plots that in the nearshore areas where the storm surge error in the SEIK simulation was around 2 m, the  $H_\infty$  SEIK simulation often improves the error to between 0.5 and 1.0 m.

We also plot in Fig. 11 the errors from the true forecast of water elevations on 29 August 2008, at around the landfall time in the SEIK and  $H_\infty$  SEIK filters. Again  $H_\infty$

TABLE 4. The average rmse of the maximum water level forecasts are shown for Katrina simulations using SEIK filter with different values of inflation factor  $\lambda$ . The second column shows the average rmse at the 405th of 8006 nodes where the hindcast data (truth) showed maximum water elevations greater than 4 m. The third column shows the average rmse of the maximum forecasted water levels where the hindcast data (truth) showed maximum water elevations greater than 5 m (163 out of 8006 nodes). The boldface font indicates the best values of rmse corresponding to inflation factor.

Inflation factor $\lambda$	Rmse for surge >4 m	Rmse for surge >5 m
—	2.38	2.87
1.0	1.52	1.82
1.1	1.70	2.04
<b>1.2</b>	<b>1.42</b>	<b>1.53</b>
1.3	1.58	1.71
1.4	1.95	2.02
1.6	1.61	1.75
2.	Numerically unstable	Numerically unstable

SEIK simulation offers a significant improvement to forecasts up to 1 m at some locations as compared to the SEIK filter particularly in the area around the landfall event.

## 5. Conclusions

The robustness in the framework of  $H_\infty$  filtering has a natural connection to the covariance inflation technique. Compared to existing works on covariance inflation in the ensemble filtering methodologies, the  $H_\infty$  filtering theory provides a theoretical framework that unifies various inflation techniques in the literature, and establishes the connection between covariance inflation and robustness. The  $H_\infty$  filtering theory also provides an explicit definition of robustness and the associated mathematical description.

We have developed an ensemble algorithm for efficient implementation of the HF based on the singular evolutive interpolated Kalman (SEIK) filter with large-scale nonlinear data assimilation problems. It is shown that applying the optimality criteria of the HF on the SEIK filter leads to an algorithm very similar to that of the SEIK filter, but with an adaptive inflation scheme. We referred to this filter as the  $H_\infty$  SEIK filter. Through numerical experiments, we verified the relative robustness of the EnTLHF in comparison with the SEIK filter. Validated hindcast studies over the western North Atlantic and Gulf of Mexico were used to generate the true states and observations, and a coarser resolution of the Gulf of Mexico was used for the data assimilation simulations. The numerical model used in this study is the Advanced Circulation (ADCIRC) model (Luettich and Westerink 2005). We have used data and simulations

TABLE 5. The average rmse of the maximum water level forecasts are shown for Katrina simulations using  $H_\infty$  SEIK filter (EnTLHF) with different values of  $c$ . The second column shows the average rmse at the 405th of 8006 nodes where the hindcast data (truth) showed maximum water elevations greater than 4 m. The third column shows the average rmse of the maximum forecasted water levels where the hindcast data (truth) showed maximum water elevations greater than 5 m (163 out of 8006 nodes). The boldface font indicates the best values of rmse corresponding to factor  $c$ .

Factor $c$	Rmse for surge >4 m	Rmse for surge >5 m
—	2.38	2.87
0.1	1.54	1.70
0.2	1.61	1.82
0.3	1.59	1.93
0.4	1.72	1.97
0.5	1.50	1.73
<b>0.6</b>	<b>0.72</b>	<b>0.45</b>
0.7	1.17	1.28
0.8	Numerically unstable	Numerically unstable

corresponding to two different storm events—Hurricane Katrina and Hurricane Ike—as test cases. The results suggest that the  $H_\infty$  SEIK filter provides more accurate storm surge forecasts than the traditional SEIK filter. This suggests that the adaptive inflation scheme in the  $H_\infty$  SEIK filter was proven quite efficient in enhancing the filter behavior during the period of water surge, within which the system uncertainties can be important. Although the results shown are only for two different test cases and focuses on short-range forecasts, they suggest that the  $H_\infty$ -based filter is more robust for storm surge forecasting problems.

We studied the relevance of a robust  $H_\infty$ -based ensemble filter in the context of improving short-range storm surge predictions. Short-range forecasting problems are mainly controlled by the initial conditions, providing a compelling test case to evaluate the behavior of a state-estimation  $H_\infty$ -based data assimilation technique. Meteorological forcing, which is the main drive behind long-range storm surge variability, was considered perfect in this study. Accounting for uncertainties in the wind forcing is essential for practical short- and long-range storm surge forecast operations and this problem will be addressed in future work. One straightforward way to do that is to use the meteorological uncertainties from weather ensemble analysis and forecasts, which are now produced in real time (Zhang et al. 2011), to force ensemble storm surge forecasts. Other input parameters to ADCIRC, such as bathymetry and bottom friction, could also change the trajectory of the surge under identical forcing terms. These parameters are typically not known with great accuracy or fidelity. We believe that using the  $H_\infty$  SEIK filter in a framework where state variables are coupled

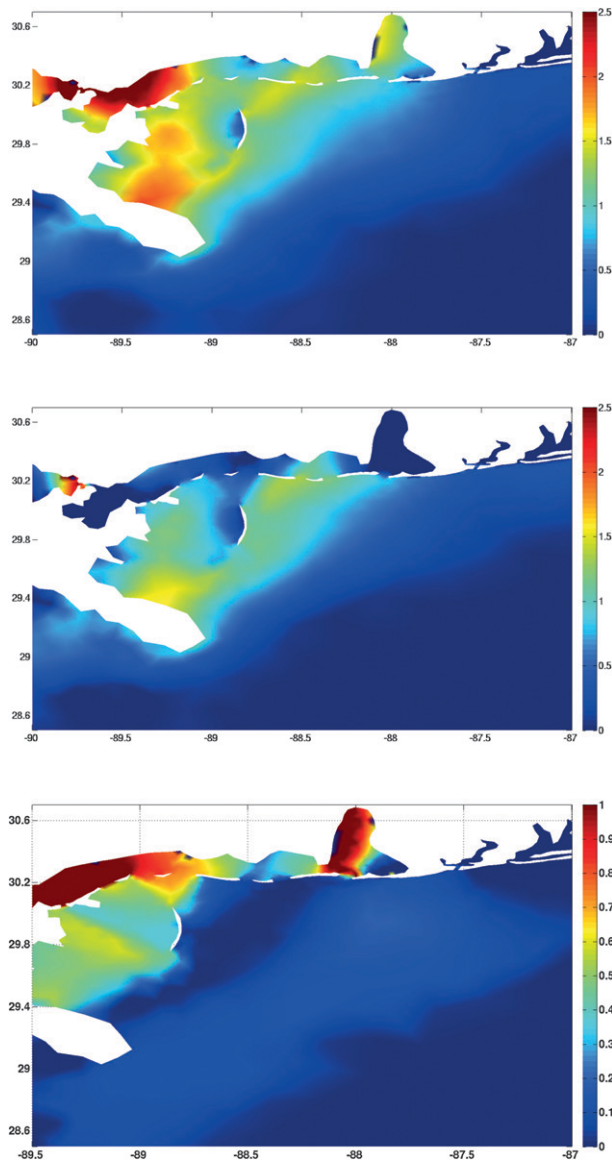


FIG. 10. Plots of errors in maximum free surface water level predictions (m) from true state for (top) simulation using SEIK filter with ( $\lambda = 1.2$ ) and (middle) simulation using  $H_\infty$  SEIK filter with ( $c = 0.6$ ). (bottom) Plot of differences between maximum free surface water level predictions (m) of simulation using SEIK filter with ( $\lambda = 1.2$ ) and simulation using  $H_\infty$  SEIK filter with ( $c = 0.6$ ). Note the different color bar range in (bottom).

with parameter estimation will result in improved storm surge forecasts. Covariance localization was also not implemented in this work, although it is likely to enhance filters performances, particularly in real forecasting exercises. This was not essential for the present work as localization is expected to have similar impact on the SEIK and  $H_\infty$ -SEIK filters solution. Furthermore, ensemble sampling errors are not expected to be a major

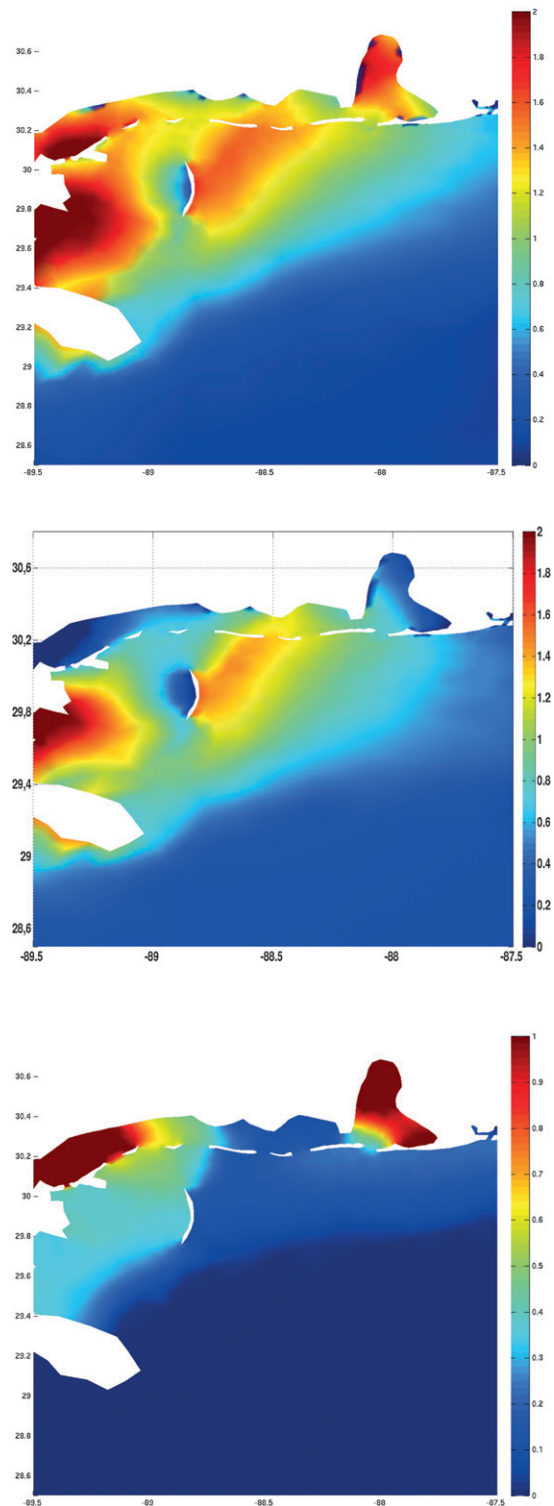


FIG. 11. Plots of free surface elevation error (m) at 1200 UTC 29 Aug 2008 from truth for (top) SEIK filter using  $\lambda = 1.2$  and (middle)  $H_\infty$  SEIK filter using  $c = 0.6$ . (bottom) Plot of Differences (m) of free surface elevation forecasts at 0800 UTC 13 Sep 2008 between  $H_\infty$  SEIK ( $c = 0.6$ ) and SEIK filter ( $\lambda = 1.2$ ). Note the different color bar range in (bottom).

source of uncertainties because of the dissipative nature of storm surge systems. All these research questions are being studied by the current coauthors.

*Acknowledgments.* Luo acknowledges partial financial support from the Research Council of Norway and industrial partners through the project “Transient well flow modelling and modern estimation techniques for accurate production allocation.”

#### REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.
- , and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.
- Bennet, A., 1992: *Inverse Methods in Physical Oceanography*. Cambridge University Press, 346 pp.
- Berg, R., 2009: Tropical cyclone report: Hurricane Ike. National Hurricane Center, Rep. AL092008, 55 pp.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.
- Brown, J. D., T. Spencer, and I. H. Moller, 2007: Modeling storm surge flooding of an urban area with particular reference to modelling uncertainties: A case study of Canvey Island, United Kingdom. *Water Resour. Res.*, **43**, W06402, doi:10.1029/2005WR004597.
- Bunya, S., and Coauthors, 2010: A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part I: Model development and validation. *Mon. Wea. Rev.*, **138**, 345–377.
- Burger, J. O., 1985: *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, 624 pp.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.
- Butler, T., M. Altaf, C. Dawson, I. Hoteit, X. Luo, and T. Mayo, 2012: Data assimilation within the advanced circulation (ADCIRC) modeling framework for hurricane storm surge forecasting. *Mon. Wea. Rev.*, **140**, 2215–2231.
- Dietrich, J., and Coauthors, 2010: A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part II: Synoptic description and analyses of Hurricanes Katrina and Rita. *Mon. Wea. Rev.*, **138**, 378–404.
- , and Coauthors, 2011: Hurricane Gustav (2008) waves and storm surge: Hindcast, synoptic analysis, and validation in southern Louisiana. *Mon. Wea. Rev.*, **139**, 2488–2522.
- El Serafy, G. Y. H., and A. E. Mynett, 2008: Improving the operational forecasting system of the stratified flow in Osaka Bay using an ensemble Kalman filter-based steady state Kalman filter. *Water Resour. Res.*, **44**, W06416, doi:10.1029/2006WR005412.
- Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
- Francis, B. A., 1987: *A Course in  $H_\infty$  Control Theory*. Springer-Verlag, 156 pp.
- Gerritsen, H., J. de Fries, and M. Philippart, 1995: The Dutch continental shelf model. *Quantitative Skill Assessment for Coastal Ocean Models*, D. Lynch and A. Davies, Eds., Coastal and Estuarine Studies, Vol. 47, Amer. Geophys. Union, 425–467.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.
- , —, J. L. Anderson, and C. Snyder, 2009: Comments on “Sigma-point Kalman filter data assimilation methods for strongly nonlinear systems.” *J. Atmos. Sci.*, **66**, 3498–3500.
- Heemink, A. W., and H. Kloosterhuis, 1990: Data assimilation for non-linear tidal models. *Int. J. Numer. Methods Fluids*, **11**, 1097–1112.
- Hoteit, I., D. T. Pham, and J. Blum, 2001: A semi-evolutionary partially local filter for data assimilation. *Mar. Pollut. Bull.*, **43**, 164–174.
- , —, and —, 2002: A simplified reduced order Kalman filtering and application to altimetric data assimilation in the tropical Pacific. *J. Mar. Syst.*, **36**, 101–127.
- , G. Korres, and G. Triantafyllou, 2005a: Comparison of extended and ensemble based Kalman filters with low and high resolution primitive equation ocean models. *Nonlinear Processes Geophys.*, **12**, 755–765.
- , G. Triantafyllou, and G. Patihakis, 2005b: Efficient data assimilation into a complex 3D physical-biogeochemical model using a partially local Kalman filter. *Ann. Geophys.*, **23**, 1–15.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- , L. Lefaire, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble predictions. *Mon. Wea. Rev.*, **124**, 1225–1242.
- Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D, J. Basic Eng.*, **82**, 35–45.
- Kennedy, A., and Coauthors, 2011: Origin of the Hurricane Ike forerunner surge. *Geophys. Res. Lett.*, **38**, L08608, doi:10.1029/2011GL047090.
- Kinmark, I., 1985: *The Shallow Water Wave Equations: Formulation, Analysis and Application*. Lecture Notes in Engineering, Vol. 15, Springer-Verlag, 216 pp.
- Knabb, R., J. Rhome, and D. Brown, 2006: Tropical cyclone report: Hurricane Katrina: 23–30 August 2005. National Hurricane Center, 43 pp.
- Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418.
- Luettich, R., and J. Westerink, cited 2005: ADCIRC: A parallel advanced circulation model for oceanic, coastal and estuarine waters. User’s manual for version 48. [Available online at [http://www.unc.edu/ims/adcirc/documentv48/ADCIRC\\_title\\_page.html](http://www.unc.edu/ims/adcirc/documentv48/ADCIRC_title_page.html).]
- Luo, X., and I. M. Moroz, 2009: Ensemble Kalman filter with the unscented transform. *Physica D*, **238**, 549–562.
- , and I. Hoteit, 2011: Robust ensemble filtering and its relation to covariance inflation in the ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 3938–3953.
- Lynch, D., and W. Gray, 1979: A wave equation model for finite element tidal computations. *Comput. Fluids*, **7**, 207–228.
- Maybeck, P., 1979: *Stochastic Models, Estimation, and Control*. Academic Press, 450 pp.

- McRobie, A., T. Spencer, and H. Gerritsen, 2005: The big flood: North Sea storm surge. *Philos. Trans. Roy. Soc. London*, **363A**, 1263–1270.
- Meng, Z., and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423.
- Murty, T. S., R. A. Flather, and R. F. Henry, 1986: The storm surge problem in the Bay of Bengal. *Prog. Oceanogr.*, **16**, 195–233.
- Nerger, L., T. Janjic, J. Schroter, and W. Hiller, 2012: A unification of ensemble square root Kalman filters. *Mon. Wea. Rev.*, **140**, 2335–2345.
- Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428.
- Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, **129**, 1194–1207.
- , J. Verron, and M. C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.*, **16**, 323–340.
- Sakov, P., and P. R. Oke, 2008: Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.*, **136**, 1042–1053.
- Schlee, F. H., C. J. Standish, and N. F. Toda, 1967: Divergence in the Kalman filter. *AIAA J.*, **5**, 1114–1120.
- Simon, D., 2006: *Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches*. Wiley-Interscience, 552 pp.
- Sorensen, J. V. T., and H. Madsen, 2006: Parameter sensitivity of three Kalman filter schemes for assimilation of water levels in shelf sea models. *Ocean Modell.*, **11**, 441–463.
- Tanaka, S., S. Bunya, J. Westerink, C. Dawson, and R. Luettich, 2011: Scalability of an unstructured grid continuous Galerkin based hurricane storm surge model. *J. Sci. Comput.*, **46**, 329–358.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.
- Wang, X., C. H. Bishop, and S. J. Julier, 2004: Which is better, an ensemble of positive-negative pairs or a centered simplex ensemble? *Mon. Wea. Rev.*, **132**, 1590–1605.
- , T. Hamill, J. Whitaker, C. Bishop, and X. Wang, 2006: A comparison of hybrid ensemble transform Kalman filter–optimum interpolation and ensemble square root filter analysis schemes. *Mon. Wea. Rev.*, **135**, 1055–1076.
- Westerink, J. J., and Coauthors, 2008: A basin- to channel-scale unstructured grid hurricane storm surge model applied to southern Louisiana. *Mon. Wea. Rev.*, **136**, 833–864.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.
- Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1053.
- , Y. Weng, J. F. Gamache, and F. D. Marks, 2011: Performance of convection-permitting hurricane initialization and prediction during 2008–2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophys. Res. Lett.*, **38**, L15810, doi:10.1029/2011GL048469.