Robust Ensemble Filtering With Improved Storm Surge Forecasting

U. Altaf, T. Buttler, X. Luo, C. Dawson,

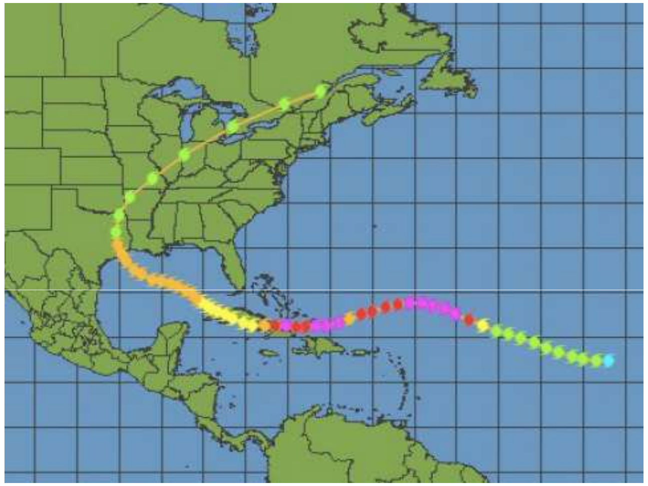
T. Mao, I.Hoteit

Meteo France, Toulouse,

Nov 13, 2012

Project

- *Ensemble data assimilation for storm surge forecasting*
- Joint project with Clint Dawson group – ICES, UT Austin
- *Area of interest:* "Gulf of Mexico"



- *Goal:* develop and implement a fully parallel nonlinear/ensemble filtering system for efficient storm surge forecasting

Motivations

- We implemented a variety EnKFs with ADCIRC with quite reasonable and comparable performances

- All filters exhibit some weakness during the surge associated with the change of regime: KFs are not well designed for such systems (Bennett, 2002; Hoteit et al., 2002):

  - *Look for ways to improve EnKFs during the surge*

  - *Give some sense to the " inflation trick" we are using in EnKFs*

Intro: Bayesian vs. Robust Filtering

- Bayesian filters update a prior with Bayes' rule to determine posterior, e.g. KFs, EnKFs, PFs, … Estimates are based on the minimum variance criterion

- All these filters make some assumptions on the statistical properties of the system, but these are often poorly known

- No guaranty that the RMS errors of these filters are "bounded", even though they are in some sense optimal

- Given all sources of poorly known uncertainties in the system, we opt for *using a <u>robust</u> instead of an <u>optimal</u> criterion*

Problem Formulation

- Consider the linear data assimilation problem

$$\begin{cases} \mathbf{x}_i = \mathcal{M}_{i,i-1}\left(\mathbf{x}_{i-1}\right) + \mathbf{u}_i \\ \mathbf{y}_i = \mathcal{H}_i\left(\mathbf{x}_i\right) + \mathbf{v}_i \end{cases}$$

  - $\mathbf{x}_i$ system state at time i
  - $\mathcal{M}_{i,i-1}$ transition matrix
  - $\mathbf{y}_i$ measurement of x_i
  - $\mathcal{H}_i$ Observation matrix
  - $\mathbf{u}_i$ dynamical and $\mathbf{v}_i$ observation Gaussian noise

Problem

- We are interested in estimating some linear combinations of the system states $\mathbf{z}_0^a, \cdots, \mathbf{z}_N^a$

$$\mathbf{z}_i = \mathbf{L}_i \mathbf{x}_i$$

given available observations

- If $\mathbf{L}_i$ the identity matrix, then the problem reduces to the estimation of the system state at every time

- Two ways to deal with this problem:
  - ✓ Direct estimation of $\mathbf{z}_i$
  - ✓ Indirect estimation: first estimate $\mathbf{x}_i$ then deduct $\mathbf{z}_i$

Kalman Filter Optimality

- The KF optimality is based on the minimum variance estimate

$$J_z^{KF}(\mathbf{z}_0^a, \cdots, \mathbf{z}_N^a) = \sum_{i=0}^{N} J_{z,i}^{KF} = \sum_{i=0}^{N} \mathbb{E}\|\mathbf{z}_i - \mathbf{z}_i^a\|_2^2$$

where

- $\mathbb{E}$ is the expectation operator
- $\mathbf{z}_i$ is the truth
- $\mathbf{z}_i^a$ is the posterior estimate

- KF solves the minimization problem sequentially

Kalman Filter (KF)

- For <u>linear Gaussian systems</u>, the Bayesian filter reduces to the KF which updates the mean and the covariance of the *pdf* as follows

Prediction Step

$$\begin{cases} \mathbf{x}_i^b = \mathcal{M}_{i,i-1}\,\mathbf{x}_{i-1}^a\,, \\ \mathbf{P}_i^b = \mathcal{M}_{i,i-1}\,\mathbf{P}_{i-1}^a\,\mathcal{M}_{i,i-1}^T + \mathbf{Q}_i\,. \end{cases}$$

Analysis Step

$$\begin{cases} \mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}_i\left(\mathbf{y}_i - \mathcal{H}_i\mathbf{x}_i^b\right)\,, \\ \mathbf{P}_i^a = \mathbf{P}_i^b - \mathbf{K}_i\mathcal{H}_i\mathbf{P}_i^b\,, \\ \mathbf{K}_i = \mathbf{P}_i^b\mathcal{H}_i^T(\mathcal{H}_i\mathbf{P}_i^b\mathcal{H}_i^T + \mathbf{R}_i)^{-1}\,, \end{cases}$$

H_∞ Optimality

- First recognize that the sources of uncertainties are in the initial conditions, the model and the observations, so the "*total energy of uncertainties*" at a given time is

$$\|\mathbf{x}_i - \mathbf{x}_i^b\|^2_{(\mathbf{\Delta}_i^b)^{-1}} + \|\mathbf{u}_i\|^2_{\mathbf{Q}_i^{-1}} + \|\mathbf{v}_i\|^2_{\mathbf{R}_i^{-1}}$$

- $\mathbf{\Delta}_0, \mathrm{Q}_i, \ \mathrm{R}_i$ are "uncertainty weight matrices", and they are user-defined by design

- Per analogy to Kalman filtering, we consider them as the errors covariance matrices.

H_∞ Optimality

- H_∞ requires that the “energy” in estimation error to be less than the total energy of uncertainties in the system

$$||\mathbf{z}_i - \mathbf{z}_i^a||^2_{\mathbf{S}_i} \leq \frac{1}{\gamma_i} \left(||\mathbf{x}_i - \mathbf{x}_i^b||^2_{(\mathbf{\Delta}_i^b)^{-1}} + ||\mathbf{u}_i||^2_{\mathbf{Q}_i^{-1}} + ||\mathbf{v}_i||^2_{\mathbf{R}_i^{-1}} \right)$$

- $\mathbf{S}_i$ is another user-defined weight matrix

- To solve this problem, consider the cost function

$$J_{z,i}^{HF} = \frac{||\mathbf{z}_i - \mathbf{z}_i^a||^2_{\mathbf{S}_i}}{||\mathbf{x}_i - \mathbf{x}_i^b||^2_{(\mathbf{\Delta}_i^b)^{-1}} + ||\mathbf{u}_i||^2_{\mathbf{Q}_i^{-1}} + ||\mathbf{v}_i||^2_{\mathbf{R}_i^{-1}}}$$

we require $J_{z,i}^{HF} \leq \dfrac{1}{\gamma_i}$

H_∞ Optimality

- Optimality of H_∞ is achieved when $1/\gamma_i^*$ is "minimax point"

$$\frac{1}{\gamma_i} = \frac{1}{\gamma_i^*} \equiv \inf_{\mathbf{z}_i^a} \sup_{\mathbf{x}_i, \mathbf{u}_i, \mathbf{v}_i} J_{z,i}^{HF}$$

i.e. the minimum cost in the worst possible case

- Because it is difficult to evaluate γ_i^*, we choose γ_i

$$\frac{1}{\gamma_i^*} \leq \frac{1}{\gamma_i}$$

This guarantees existence of an H_∞ solution (Simon, 2006)

$$\sum_{i=0}^{N} \|\mathbf{z}_i - \mathbf{z}_i^a\|_{\mathbf{S}_i}^2 \leq \max_i \{\frac{1}{\gamma_i}\} \left(\sum_{i=0}^{N} \|\mathbf{x}_i - \mathbf{x}_i^b\|_{(\boldsymbol{\Delta}_i^b)^{-1}}^2 + \sum_{i=0}^{N} \|\mathbf{u}_i\|_{\mathbf{Q}_i^{-1}}^2 + \sum_{i=0}^{N} \|\mathbf{v}_i\|_{\mathbf{R}_i^{-1}}^2 \right)$$

The H_∞ Filter (HF)

- H_∞ filter updates a prior estimate to its posterior based on the minimax criterion as follows (Simon 2006)

Prediction Step

$$\begin{cases} \mathbf{x}_i^b = \mathcal{M}_{i,i-1}\,\mathbf{x}_{i-1}^a\,, \\ \mathbf{\Delta}_i^b = \mathcal{M}_{i,i-1}\,\mathbf{\Delta}_{i-1}^a\,\mathcal{M}_{i,i-1}^T + \mathbf{Q}_i\,. \end{cases}$$

Analysis Step

$$\begin{cases} \mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{G}_i\left(\mathbf{y}_i - \mathcal{H}_i\mathbf{x}_i^b\right) \\ (\mathbf{\Delta}_i^a)^{-1} = (\mathbf{\Delta}_i^b)^{-1} + (\mathcal{H}_i)^T(\mathbf{R}_i)^{-1}\mathcal{H}_i - \gamma_i\mathbf{L}_i^T\mathbf{S}_i\mathbf{L}_i\,, \\ \mathbf{G}_i = \mathbf{\Delta}_i^a\mathcal{H}_i^T(\mathbf{R}_i)^{-1}\,, \end{cases}$$

subject to

$$(\mathbf{\Delta}_i^a)^{-1} = (\mathbf{\Delta}_i^b)^{-1} + (\mathcal{H}_i)^T(\mathbf{R}_i)^{-1}\mathcal{H}_i - \gamma_i\mathbf{L}_i^T\mathbf{S}_i\mathbf{L}_i \geq 0\,.$$

HF vs. KF

- $\|\mathbf{z}_i - \mathbf{z}_i^a\|_{\mathbf{S}_i}^2$ is bounded above by some finite value in HF. This is not necessarily true for KF!
- If $\gamma_i = 0$ then the HF reduces to KF
- The choice of $\mathbf{L}_i$ affects the estimate of HF, but not KF
- HF is more conservative; it tends to make its analysis uncertainties larger than that of the KF

$$(\mathbf{\Delta}_i^a)^{-1} = (\mathbf{\Sigma}_i^a)^{-1} - \gamma \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i < (\mathbf{\Sigma}_i^a)^{-1}$$

- KF is expected to perform better if system statistics are well known, but HF should be more "robust"

EnHF: A Hybrid HF - EnKF

- HF can be based on any EnKF, stochastic or deterministic
- The idea is to first use an EnKF to compute the uncertainty matrix Σ_i^a satisfying

$$(\boldsymbol{\Sigma}_i^a)^{-1} = (\boldsymbol{\Delta}_i^b)^{-1} + (\mathcal{H}_i)^T (\mathbf{R}_i)^{-1} \mathcal{H}_i$$

then "inflate" Σ_i^a to compute

$$(\boldsymbol{\Delta}_i^a)^{-1} = (\boldsymbol{\Sigma}_i^a)^{-1} - \gamma_i \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i \geq 0$$

with an appropriate/robust choice of γ_i

HF and Inflation in EnKFs

- By choosing different forms of $\gamma_i \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i$ in the EnHF update formula of the uncertainty matrix

$$(\boldsymbol{\Delta}_i^a)^{-1} = (\boldsymbol{\Delta}_i^b)^{-1} + (\mathcal{H}_i)^T (\mathbf{R}_i)^{-1} \mathcal{H}_i - \gamma_i \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i \geq 0$$

we can derive any EnKF with covariance inflation

- Case I-BG: If $\gamma_i \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i = c(\hat{\boldsymbol{\Delta}}_i^b)^{-1}$, we obtain the SEIK inflation in Pham et al. (1998)

$$(\hat{\boldsymbol{\Delta}}_i^a)^{-1} = (1-c)(\hat{\boldsymbol{\Delta}}_i^b)^{-1} + (\mathcal{H}_i)^T (\mathbf{R}_i)^{-1} \mathcal{H}_i$$

- Case I-ANA: If $\gamma_i \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i = c\ (\hat{\boldsymbol{\Sigma}}_i^a)^{-1}$, we derive the SR-EnKF inflation in Whitacker and Hamill (2002)

$$\hat{\boldsymbol{\Delta}}_i^a = (1-c)^{-1} \hat{\boldsymbol{\Sigma}}_i^a$$

HF with Modes Inflation

- Case I-MTX: If $\mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i = \mathbf{I}_{m_x}$, then

$$(\hat{\mathbf{\Delta}}_i^a)^{-1} = (\hat{\mathbf{\Sigma}}_i^a)^{-1} - \gamma_i \mathbf{I}_{m_x}$$

In this case, using an SVD on the EnKF analysis covariance matrix before inflation

$$\hat{\mathbf{\Sigma}}_i^a = \mathbf{E}_i^a \mathbf{D}_i^a (\mathbf{E}_i^a)^T$$, where $\mathbf{D}_i^a = \mathrm{diag}(\sigma_{i,1}, \cdots, \sigma_{i,m_x})$

Then after inflation,

$$\hat{\mathbf{\Delta}}_i^a = \mathbf{E}_i^a \mathbf{\Lambda}_i^a (\mathbf{E}_i^a)^T$$, with $\mathbf{\Lambda}_i^a = \mathrm{diag}\left(\frac{\sigma_{i,j}}{1 - c\,\sigma_{i,j}/\sigma_{i,1}}\right)$, $0 \le c < 1$

- Very similar to the ETKF inflation of Ott et al. (2004) who augmented the eigenvalues by a constant

A Simple Example

- Consider the model

$$x_{i+1} = 1 + 0.5x_i - 0.1x_i^2 + f(x_i; k, h, d) + u_i \,,$$

- Forecast model

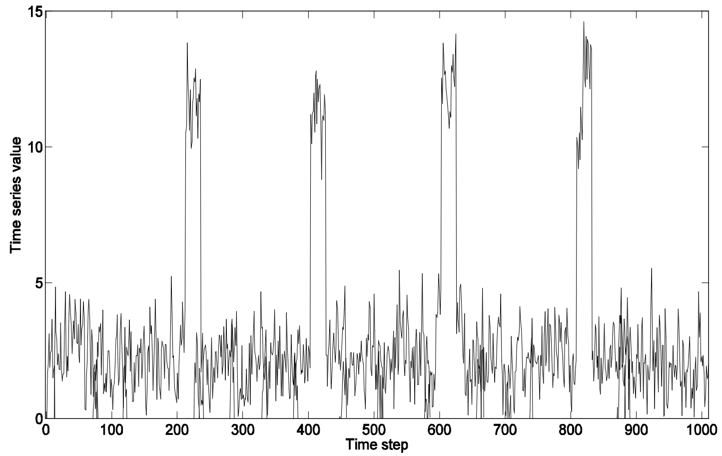
$$x_{i+1} = 1 + 0.5x_i + u_i$$

- Observation model

$$y_i = x_i + v_i$$

with $u_i \sim \mathrm{N}(u_i : 0, 1)$

$v_i \sim \mathrm{N}(v_i : 0, 1).$



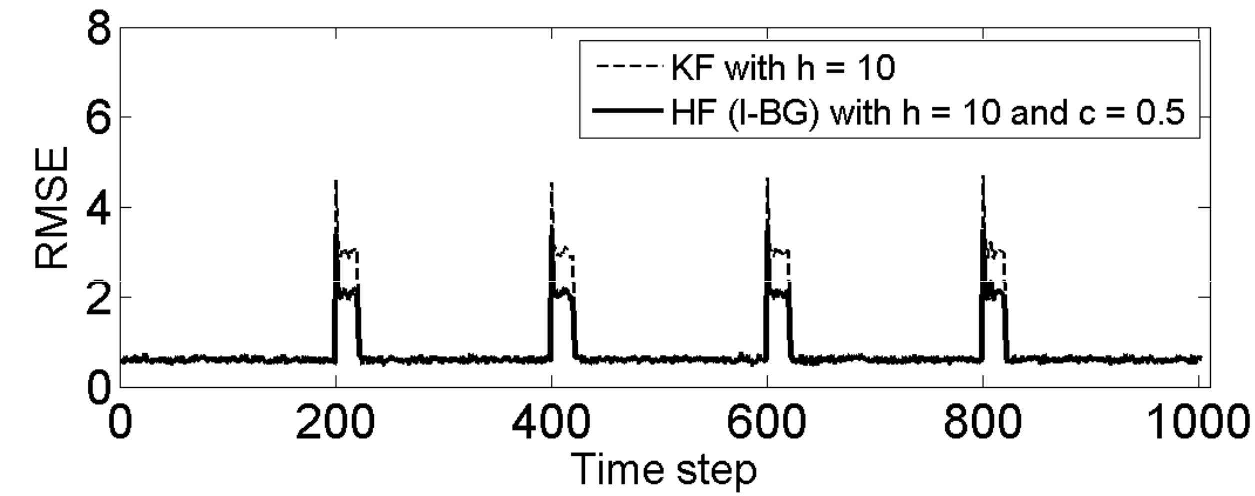
A time series with

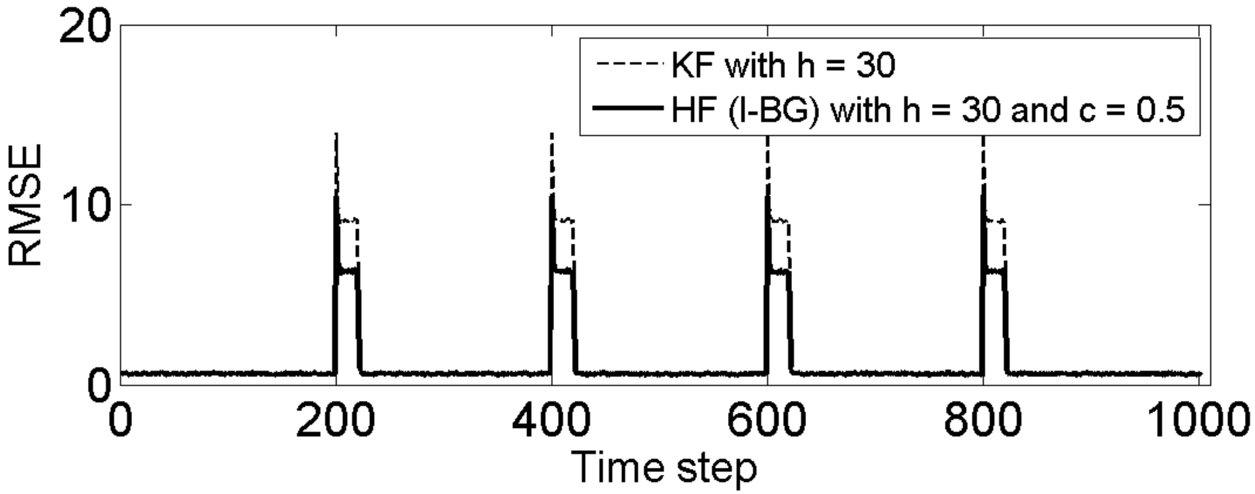
$$h = 10$$

$$k = 200,\ 400,\ 600 \text{ and } 800$$

A Simple Example – HF I-BG

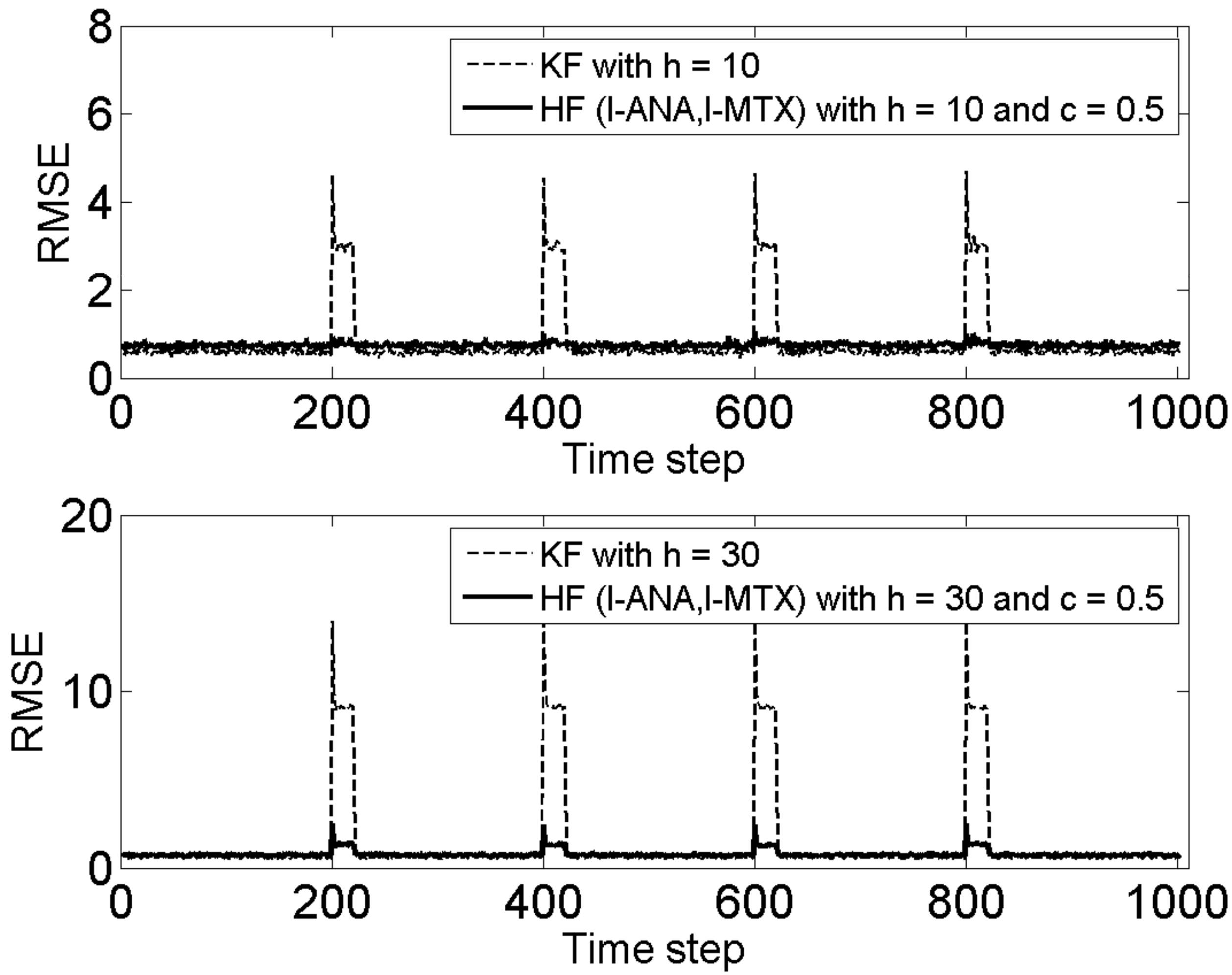
Assimilation results of I-BG HF:





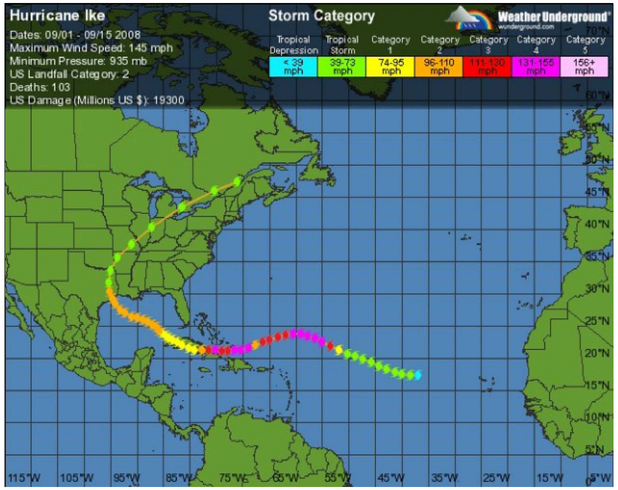
A Simple Example – HF I-ANA

- HF I-ANA and I-MTX are equivalent in 1D

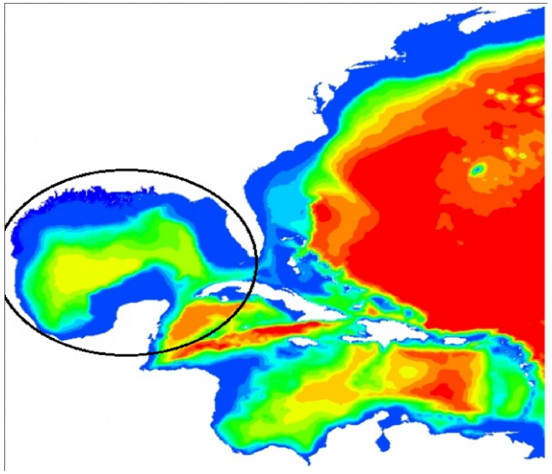


Application to Storm Surge Forecasting

- Interest of forecasting storm surge has dramatically increased since the devastating 2005 hurricane season



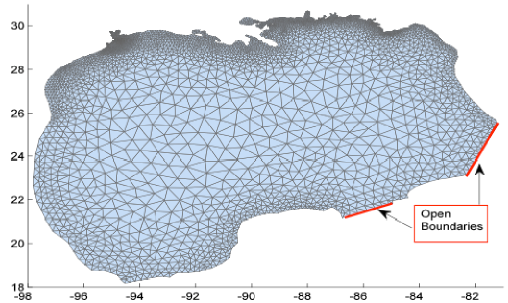
- Advanced Circulation (ADCIRC) discretizes shallow water equations using FEM on unstructured meshes
- A case study Hurricane Ike, which made landfall along the upper Texas coast on Sep. 13 2008

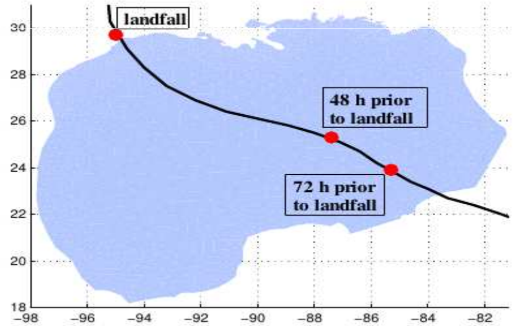


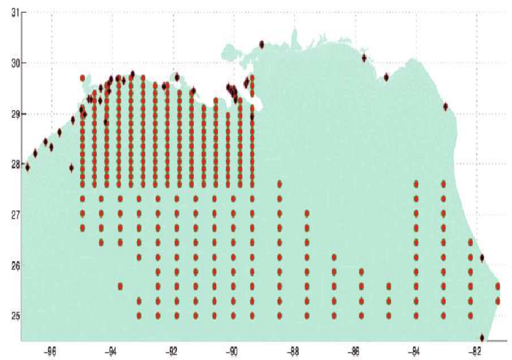
- Observations of water levels are taken from a high-resolution hindcast of Ike
- Forecast model uses a low-resolution configuration with different winds and ICs

Experiments Design

- Assimilation experiments setup
  - Time step: 10 s
  - Grid of 8006 nodes for U, V, Eta and 14,269 elements
  - 5 tidal constituents: M2, S2, K1, O1, P1
  - Measurement Stations: 350
  - Analysis: Every 2 hours
  - Assimilations steps: 48
  - HF based on SEIK
  - Ensemble size: 10





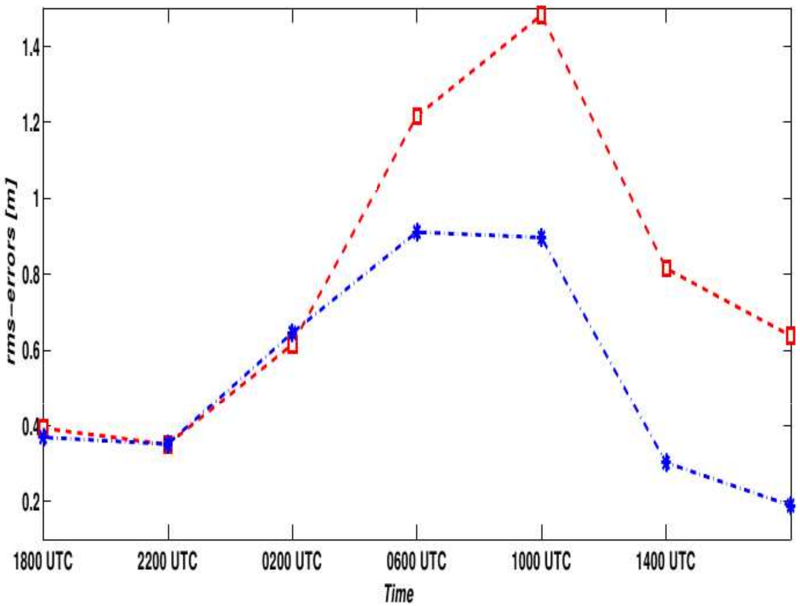


Inflation factor λ	Coastal rms-error	Rms-error > 3m
ND	1.92	1.91
1.0	1.68	1.65
1.1	1.59	1.62
1.2	1.45	1.46
1.4	1.58	1.61
1.5	1.62	1.65
1.6	**1.38**	**1.42**
1.7	1.47	1.51
2.	1.52	1.54

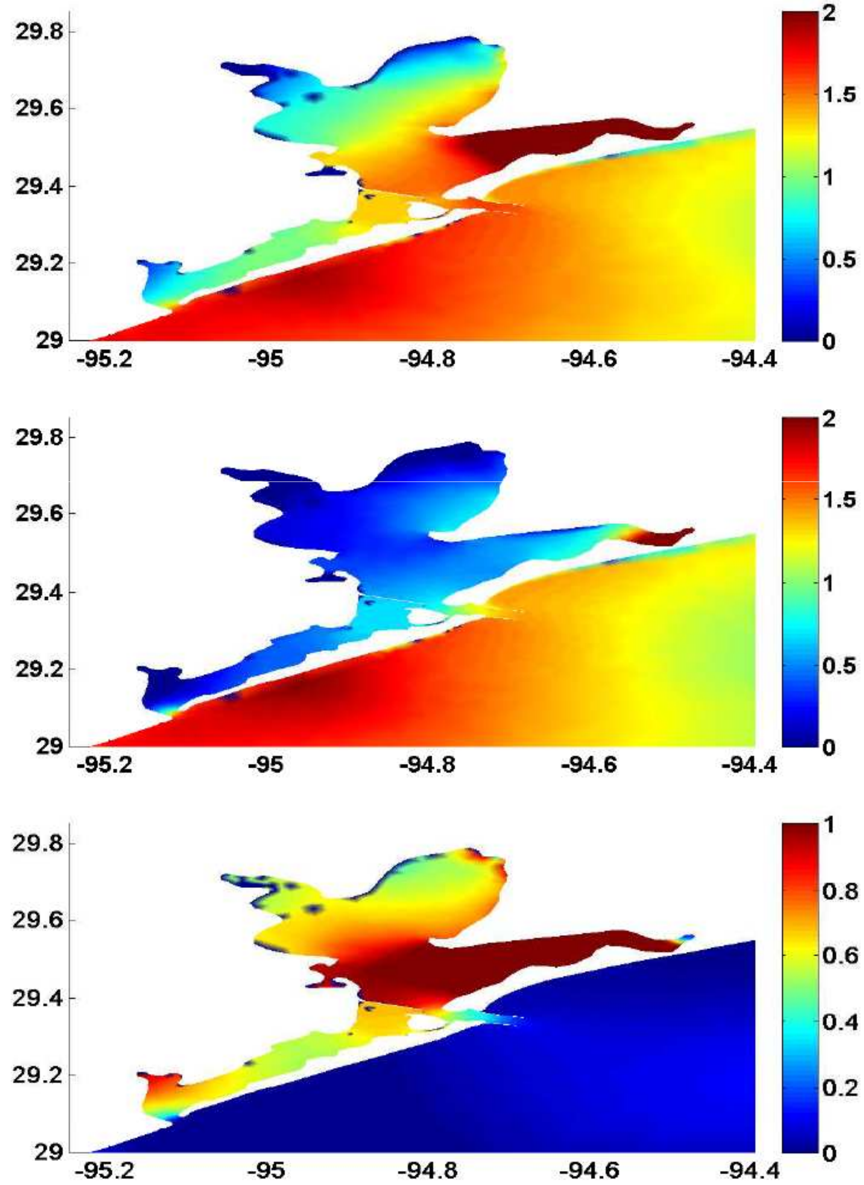
Factor c	Coastal rms-error	Rms-error > 3m
ND	1.92	1.91
0.1	1.43	1.38
0.2	1.40	1.42
0.3	1.47	1.50
0.4	1.34	1.36
0.5	1.30	1.33
0.6	1.17	1.10
0.7	**0.80**	**0.87**
0.8	1.35	1.38

Average rms-errors of the maximum water level forecasts in Ike simulations using 1) SEIK and 2) HF-SEIK with different inflation

Free surface elevation error on 13/9/2008 at 0800 UTC from truth SEIK, HF-SEIK, and differences



Averaged rms-error of water elevations in the landfall area best cases with SEIK and HF-SEIK between 9/12/2008 and 9/13/2008



Discussion

- H_∞ provides a unified framework for inflation in EnKFs
- H_∞ is more robust for systems with fast varying regimes

- *Develop "optimal" adaptive inflation schemes based on HF: one still need to add an optimal criterion to define "optimal inflation"*
- *Include parameters and inputs, such as bathymetry and winds, in the estimation process*
- *Assimilation with coupled wave - storm Surge models*

References

- I. Hoteit, D.-T. Pham, and J. Blum, 2002. *A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific*. Journal of Marine Systems, 36, 101-127.
- D. Simon, 2006. *Optimal state estimation: Kalman, H-infinity, and nonlinear approaches*. Wiley-Intersciences, pp.522.
- X. Luo X, I. Hoteit , 2011. *Robust filtering and its relation to covariance inflation in the ensemble Kalman filter*. Monthly Weather Review, 139, 3938-3953.
- T. Butller, U. Altaf, C. Dawson, I. Hoteit, X. Luo, and T. Mayo, 2011: *Data Assimilation within the advanced circulation (ADCIRC) modeling framework for hurricane storm surge forecasting*. Monthly Weather Review, 140, 2215-2231.

THANK YOU

THANK YOU

KAUST

King Abdullah University of Science and Technology

Participants

KAUST	UT-Austin
Ibrahim Hoteit	Clint Dawson
Umer Altaf	Troy Butler
Xiaodong Luo	Talea Mayo

Intro: Assimilation

- Data assimilation combines numerical models and data to compute the best possible estimate of the state of a dynamical system

- All assimilation schemes have been derived from the Bayesian filtering theory, determine *pdf* of the state given available data

Uncertainty Quantification **+** *Uncertainty Reduction*

{ Forecast: propagate *pdf* with the model in time

{ Analysis: correct prior *pdf* with new data

HF and Inflation in EnKFs

- Inflation is becoming a standard tool in EnKFs

  Hamill et al. (2011): Since the early implementations of the EnKF, several now-standard modifications are commonly considered to be essential in spatially distributed systems; the first is some form of "localization" of covariances (Houtekamer and Mitchell 2001; Hamill et al. 2001). Another common technique for the stabilization of the EnKF is the enlargement of the prior through "covariance inflation" (Anderson and Anderson 1999)

- No rigorous framework for inflation yet!

  Talagrand on Hoteit's thesis (2001):

  My only critic about this thesis is related to the use of forgetting factor. I do not see any theoretical reason to use it!

Why Using H_∞?

- Better deal with large dimensional geophysical systems with intermittent and fast varying regimes which are subject to
  - ✓ *Important model uncertainties*
  - ✓ *Poor priors*
- Provide a theoretical framework for different inflations

Intro: Robust H_∞ Filtering

- Focus on the robustness of the estimate in the sense that it has better tolerance to possible uncertainties
- Do not assume the complete knowledge of the statistics of the system in assimilation; recognizing that some uncertainties cannot be avoided
- Replace the optimal estimate criterion by a robust criterion, e.g. H_∞ *which is based on a minimax criterion*

HF and Inflation in EnKFs

- Case I-OBS: If $\gamma_i \mathbf{L}_i^T \mathbf{S}_i \mathbf{L}_i = c(\mathcal{H}_i)^T (\mathbf{R}_i)^{-1} \mathcal{H}_i$, which leads to

$$(\hat{\mathbf{\Delta}}_i^a)^{-1} = (\hat{\mathbf{\Delta}}_i^b)^{-1} + (1-c)(\mathcal{H}_i)^T (\mathbf{R}_i)^{-1} \mathcal{H}_i$$

or, in other words, to the inflation of the observation covariance.

- In the EnKF, the observation covariance is generally under-sampled because of the limited ensemble size. This means

$$1 - c > 1 \quad \Longrightarrow \quad \gamma_i < 0$$

implying more confidence in the prior, which could explain some underperformances of the EnKF compared to SR-EnKFs.

- *The EnKF could benefit from the inflation of the observation covariance*

Time is 13 Sept. 08 00:00 UTC

Top: Forecast. Middle: No assimilation Bottom: Difference