Data assimilation using ensemble methods for hurricane forecasting

Pushkar Kumar Jain, Yuxiang Lin, Kyle Mandli, Clint Dawson

The University of Texas at Austin
Aerospace Engineering and Engineering Mechanics
Cockrell School of Engineering

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES

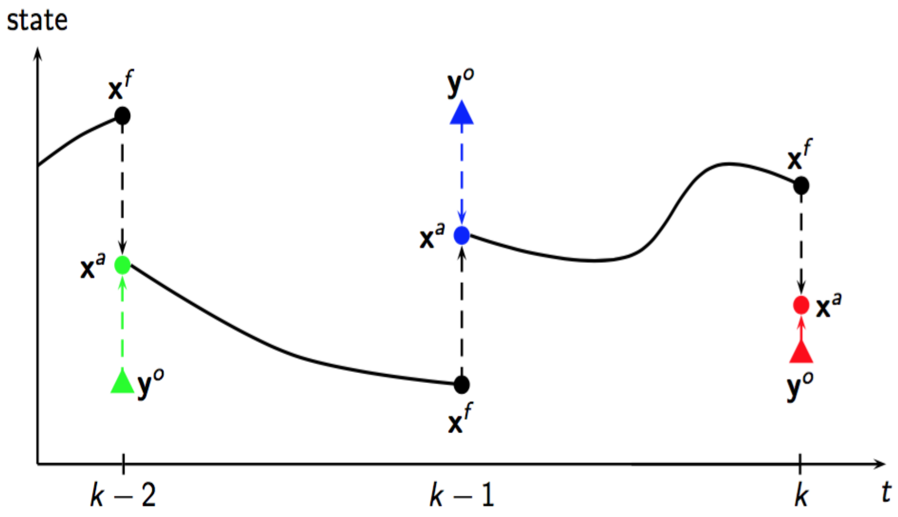
COMPUTATIONAL HYDRAULICS GROUP
THE UNIVERSITY OF TEXAS AT AUSTIN

Outline

- Introduction
- Ensemble Kalman filter algorithm
- Schematic - EnKF in action
- Implementation
- Conclusion

Introduction

- Data assimilation addresses the problem of incorporating observations into a model of the system in an optimal way
- Given a noisy discrete model of dynamics of the system and noisy observations of the system, find estimates of the state of the system
- The objective is to investigate computational methods to advance the science of real time prediction of coastal and hydrological hazards and improve the designs of observational systems
- Techniques – Optimal interpolation, 3D Var, 4D Var, EKF, EnKF



"Data driven simulation"

Algorithm

- $\boldsymbol{X}_b$: Prior guess of initial state
- $\boldsymbol{x}_{new} = \mathcal{M}(\boldsymbol{x}_{old})$: State evolution in time
- $\boldsymbol{y}$: Observation vector
- $\mathcal{H}(\cdot)$: Observations map to state
- $\boldsymbol{B}$: Prior guess error covariance matrix
- $\boldsymbol{R}$: Observation error covariance matrix
- $\boldsymbol{Q}$: Model error covariance matrix

1. Generate ensemble of size N:

$$\boldsymbol{x}_{\boldsymbol{b}}^{(\boldsymbol{i})} = \boldsymbol{x}_{\boldsymbol{b}} + \boldsymbol{\eta}\ for\ i = 1, N \text{ and } \eta \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{B})$$

1. Prediction step:
   a) Propagate each ensemble member forward in time:

$$\boldsymbol{x}_{\boldsymbol{b}}^{(\boldsymbol{i})}(t+1) = \mathcal{M}\left(\boldsymbol{x}_{\boldsymbol{b}}^{(\boldsymbol{i})}(t)\right)$$

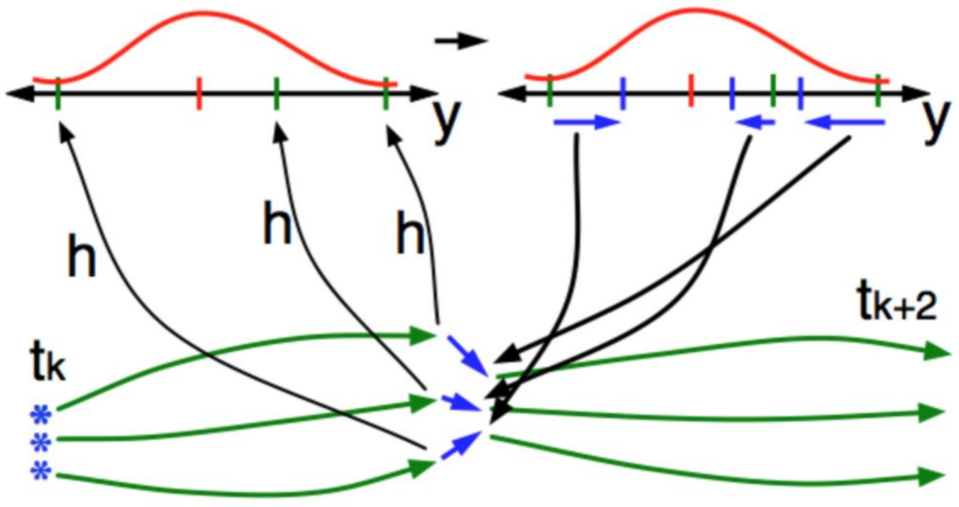
2. Update step:
   a) $\mathbf{P_b}$ is evolved B matrix: $\mathbf{P_b} = \frac{1}{\mathrm{N}-1}\sum\left(\mathbf{x_b^{(i)}} - \bar{\mathbf{x}}\right)\left(\mathbf{x_b^{(i)}} - \bar{\mathbf{x}}\right)^{\mathrm{T}}$
   b) Calculate Kalman gain: $\boldsymbol{K} = \boldsymbol{P_b}\boldsymbol{H}^T(\boldsymbol{R} + \boldsymbol{H}\boldsymbol{P_b}\boldsymbol{H}^{\boldsymbol{T}})^{-1}$
   c) Update each ensemble member:

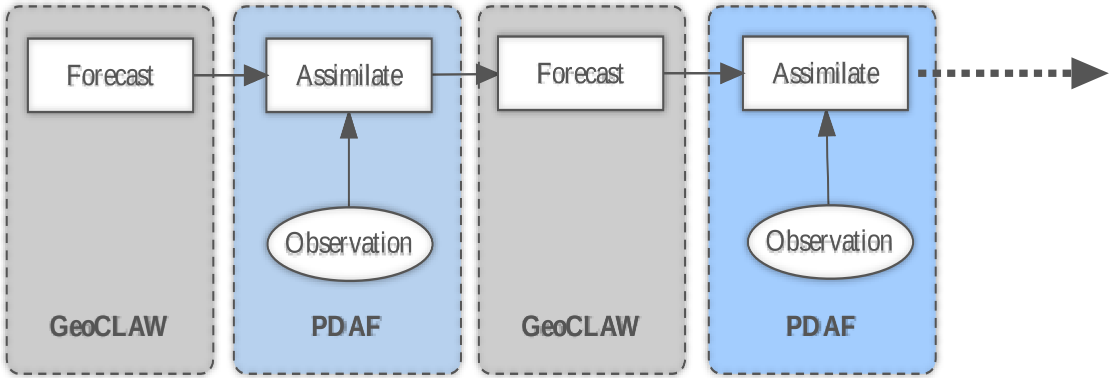
$$\boldsymbol{x}_{\boldsymbol{b}}^{(i)} \longleftarrow \boldsymbol{x}_{\boldsymbol{b}}^{(i)} + \boldsymbol{K}\left(\boldsymbol{y} + \boldsymbol{\epsilon} - \mathcal{H}\left(\boldsymbol{x}_{\boldsymbol{b}}^{(i)}\right)\right), \text{ where } \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$$

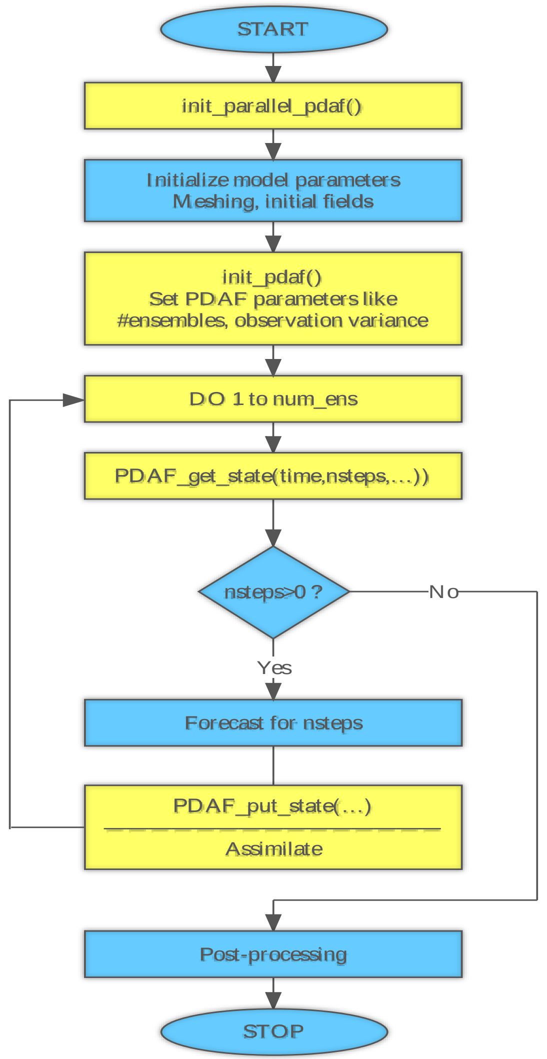
   a) Calculate analysis error covariance: $\boldsymbol{P_a} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{P_b}$ **(if needed)**

3. Repeat 2 and 3 till all observations are assimilated

EnKF in action



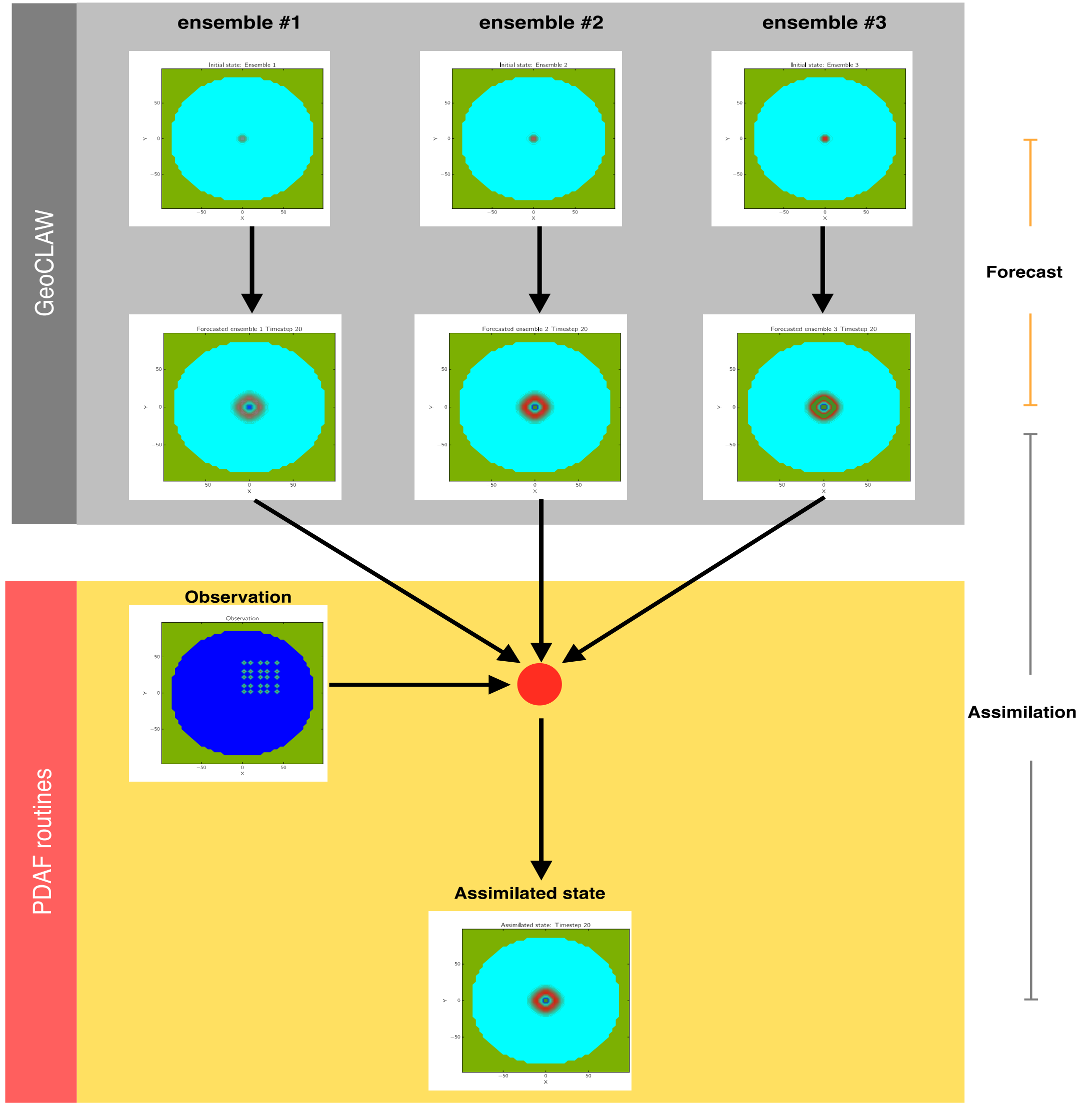


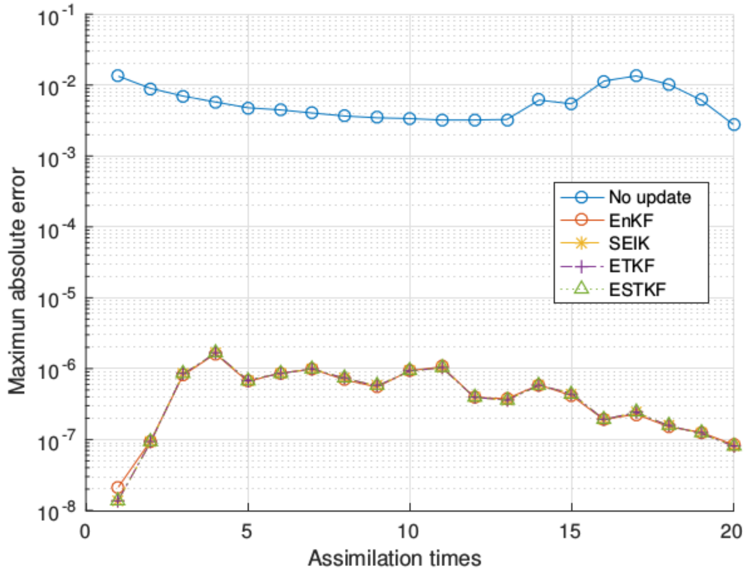


Implementation

- Data assimilation is performed using the library **PDAF** - Parallel Data assimilation Framework
- Objective - Sequential forecasting and analysis step to obtain estimated analysis state
- Goals –
  - Tight two way coupling between AMR and data assimilation
  - Analyze how assimilation affects adpative mesh refinement
- Issues –
  - Varying size of field due to adaptive mesh refinement
  - Code refactoring to accommodate the PDAF routines
  - Forced refinement/fixed reference mesh
  - Multi-level assimilation

Schematic representation





Conclusion

- The technology will be tested using actual data from recent events, and implemented on high-performance computational platforms.
- These advances offer the promise of significantly transforming data-driven, real-time modeling of hydrological hazards, with potentially broader applications in other science domains.

Collaborators

- Dr. Clint Dawson, The University of Texas at Austin
- Dr. Kyle Mandli, Columbia University

- **Acknowledgement:** This work is supported by King Abdullah University of Science & Technology (KAUST)