Data assimilation using ensemble methods for hurricane forecasting

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

• Introduction
• Ensemble Kalman filter algorithm
• Schematic - EnKF in action
• Implementation
• Conclusion
• 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 assimilation using ensemble Kalman filtering

**Algorithm**

1. **Generate ensemble of size** $N$:
   \[ x_b^{(i)} = x_b + \eta \text{ for } i = 1, N \text{ and } \eta \sim \mathcal{N}(0, B) \]

1. **Prediction step**:
   a) Propagate each ensemble member forward in time:
   \[ x_b^{(i)}(t+1) = \mathcal{M}\left(x_b^{(i)}(t)\right) \]

2. **Update step**:
   a) **$P_b$** is evolved B matrix:
   \[ P_b = \frac{1}{N-1} \sum \left(x_b^{(i)} - \bar{x}\right)\left(x_b^{(i)} - \bar{x}\right)^T \]
   b) Calculate Kalman gain:
   \[ K = P_b H^T (R + H P_b H^T)^{-1} \]
   c) Update each ensemble member:
   \[ x_b^{(i)} \leftarrow x_b^{(i)} + K \left(y + \epsilon - \mathcal{H}\left(x_b^{(i)}\right)\right), \text{ where } \epsilon \sim \mathcal{N}(0, R) \]
   a) Calculate analysis error covariance:
   \[ P_a = (I - KH) P_b \text{ (if needed)} \]
3. Repeat 2 and 3 till all observations are assimilated

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

Initialize model parameters
Meshing, initial fields
DO 1 to num_ens
PDAF_get_state(time,nsteps,…))
Forecast for nsteps
PDAF_put_state(…)

Assimilate
Observation
Forecast

PDAF
GeoCLAW
PDAF
GeoCLAW
PDAF_get_state(time,nsteps,…))
set PDAF parameters like ensembles, observation variance

Yes
nsteps>0 ?

Post-processing
STOP
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 adaptive 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
Data assimilation using ensemble Kalman filtering
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)