Data Assimilation 2: Monte Carlo Methods

"An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem." John W. Tukey
Forecast Cycle

Forecast Step

1: Initial State

Analysis Step

2: Forecast
3: New Observation
4: Updated State

Time
## Uncertainty Propagation

### Applied in the Forecast Step

<table>
<thead>
<tr>
<th>Approach</th>
<th>Distribution</th>
<th>Output Moments</th>
<th>Output Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic</td>
<td>Variable Transform</td>
<td>Analytical Moments</td>
<td>KF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Taylor Series</td>
<td>EKF</td>
</tr>
<tr>
<td>Numeric</td>
<td>Monte Carlo</td>
<td>Ensemble</td>
<td>EnKF</td>
</tr>
<tr>
<td></td>
<td>PF</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Kalman Analysis

- **Forecast:**
  Assume $P(X_{t+1}) \sim N(\mu_f, p_f)$

- **Observation error:**
  Assume $P(Y_{t+1} | X_{t+1}) \sim N(X_{t+1}, r)$

- **Likelihood = Data model**

- Assume $Y$, $\mu_f$, $p_f$ and $r$ are known

- $P(X_{t+1} | Y_{t+1}) \sim N(\mu_a, p_a)$
  \[ X | Y \sim N\left(\frac{\rho}{n\rho + \phi} n\bar{Y} + \frac{\phi}{n\rho + \phi} \mu_f, n\rho + \phi\right) \]  \[ \rho = 1/r \quad \phi = 1/p_f \]
\[ X_a | Y \sim N \left( Y | HX_a, R \right) N \left( X_a | \mu_f, P_f \right) \]

- Solves to be

\[ X_a | Y \sim N \left( \left( H^T R^{-1} H + P_f^{-1} \right)^{-1} \left( H^T R^{-1} Y + P_f^{-1} \mu_f \right), \left( H^T R^{-1} H + P_f^{-1} \right)^{-1} \right) \]

- Mean and variance simplify to

\[ E \left[ X_a | Y \right] = \mu_a = \mu_f + K \left( Y - H \mu_f \right) \]

\[ Var \left[ X_a | Y \right] = P_a = \left( I - KH \right) P_f \]

\[ K = P_{fH}^T \left( R + H P_f H^T \right)^{-1} \quad \text{Kalman Gain} \]
Forecast Step \[ X_{t+1} = MX_t + \epsilon \]

The posterior distribution of \( X_{t+1} \) given \( X_t \) is multivariate normal with

\[ \mu_{f,t+1} = E[X_{f,t+1} | X_{a,t}] = M_t \mu_{a,t} \]

\[ P_{f,t+1} = Var[X_{f,t+1} | X_{a,t}] = Q_t + M_t P_{a,t-1} M_t^T \]
Extended Kalman Filter (EKF)

- Addresses **linear** assumption of the Forecast
  - \( \mu_f = f(\mu_a) \)

- Update variance using a Taylor Series expansion
  - \( F = \text{Jacobian} \left( \frac{df_i}{dx_j} \right) \)
  - \( P_f \approx Q + F P_a F^T \) (was \( Q + M P_a M^T \))

- Can be extended to higher orders

- Jensen’s Inequality: Biased, Normality assumption FALSE
Ensemble Kalman Filter (EnKF)

- Analysis identical to KF
- Uses Monte Carlo samples to approximate Forecast distribution
- Draw m samples from the Analysis posterior
- Run process model + process error for sample

\[
\mu_{f,t+1} = \frac{1}{m} \sum_{i} X_{f,i}
\]

\[
P_{f,t+1} = COV[X_{f,i}]
\]
Ensemble adjustment (Kalman) filter

- 5 member ensemble of prior estimates of an observed variable
- Obtained by applying $h$ to ensemble state vector
Ensemble adjustment (Kalman) filter

![Graph showing probability distribution](image)
Ensemble adjustment (Kalman) filter

observation likelihood
posterior distribution
Ensemble adjustment (Kalman) filter
Ensemble adjustment (Kalman) filter

So first shift prior ensemble so it has the same mean as the posterior.
Ensemble adjustment (Kalman) filter

Posterior PDF

Variance Adjusted

Mean Shifted

posterior
Ensemble Adjustment

- Alt to resampling analysis posterior, nudge current ensemble
- Useful when other uncertainty & latent states

- **SVD:**
  \[ P = VLV^{-1} \]

- **Normalize:**
  \[ Z_i = \sqrt{L_f^{-1}} V_f^{-1} (X_{i,f} - \mu_f) \]

- **Update:**
  \[ X_{i,a} = V_a \sqrt{L_a} Z_i + \mu_a \]
EnKF pro/con

- Nonlinear
- Existing code: No Jacobian
- Simple to implement, understand
- Sample size chosen based on power analysis
  - Con: larger than Analytical methods
- Simpler to add other sources of uncert. (e.g. driver)
- Moments OK on Jensen’s Inequality
- Normal, but violates Normality
  - Analysis not hard to generalize (Likelihood * Prior) but unlikely to have an analytical sol’n
Localization

☐ All KF flavors involve matrix inversion

☐ Cheaper if correlation matrix is sparse

☐ Often assume correlations beyond some distance are zero

☐ avoids spurious correlations

☐ distance need not be physical
Filter Divergence

- Practitioners of DA in atm sci frequently worry about model variance collapsing to zero
- Model then ignores (diverges from) data
- Process error is TUNED [BAD]
- Ecology is far less chaotic
  - Occasionally, convergence is right answer
  - In others, indicates misspecified process model or partitioning of process error
No KF variant can estimate process and observation errors

Random Walk State Space

\[ P(X_t, \tau_{obs}, \tau_{proc} \mid Y_t) \propto N(Y_t \mid X_t, \tau_{obs}) \times \]
\[ N(X_t \mid X_{t-1}, \tau_{proc}) \Gamma(\tau_{proc}) \Gamma(\tau_{obs}) \]
What if we forecast with a large Monte Carlo sample?

- Can eliminate distributional assumptions!
- Can eliminate Normal x Normal Analysis
- How to do Analysis step when prior is a sample, not an equation?
Particle Filter
Particle Filter

- Weights provided by the likelihood
  - posterior $\propto$ likelihood $\times$ prior
- Estimates based on weighted mean, variance, CI, etc.
- a.k.a. Sequential Monte Carlo
Resampling PF

- Problem: weights can converge on 1 ensemble member
- Solution: resampling & split to maintain a distribution
Particles from Step $t-1$

\[ \{x^i_t, w^i_t\}_{i=1}^{N_s} \]

Particles after Resampling and Propagation

\[ \{x^i_{t+1}, w^i_{t+1}\}_{i=1}^{N_s} \]
When to resample?

- Too often: loose particles through drift
- Not enough: converges (degeneracy), poor distribution
- Typically resample when effective sample size, $1/\sum(W^2)$, drops below some threshold (e.g. $N/2$)
- NOTE: At resample, weights reset to 1!!
Particle Filter pro/con

- **Con:**
  - Computation!

- **Pros:**
  - Simple to implement
  - General, Flexible
  - Can evaluate all params
  - Parallelizable
Kernel Smoothing

- Parameters lack process error, subject to degeneracy
- Can be resampled from kernel smoother = continuous approx of joint PDF
- Req choice of smoothing/bandwidth
- Even better if M-H accept/reject proposed Moves
- Global, Gaussian smoothing

\[ \theta_i^* = \bar{\theta} + h(\theta_i - \bar{\theta}) + e_i \sqrt{1 - h^2} \]

\[ e_i \sim \text{MVN}(0, \Sigma) \]

h=1 no smoothing
h=0 redraw iid
### UNCERTAINTY PROPAGATION APPLIED IN THE FORECAST STEP

<table>
<thead>
<tr>
<th>Approach</th>
<th>Distribution</th>
<th>Output Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytic</strong></td>
<td>Variable Transform</td>
<td>Analytical Moments, Taylor Series</td>
</tr>
<tr>
<td><strong>Numeric</strong></td>
<td>Monte Carlo</td>
<td>Ensemble</td>
</tr>
<tr>
<td></td>
<td>PF</td>
<td></td>
</tr>
</tbody>
</table>

**KF**  
**EKF**  
**EnKF**

**WHAT ABOUT THE ANALYSIS STEP?**
What about MCMC?

- Option 1: Refit full State-Space Model
- Option 2: Just update forecast from State-Space
  - A: treat priors (forecast & params) as samples -> PF
  - B: approximate priors w/ dist’n
q = 0.25
r = 0.0625
m = 0.9
GENERALIZED ENSEMBLE FILTER

Multivariate Tobit
- Range restrictions
- Zero inflated

Estimated Process Error

Raiho et al in prep
Take Homes

▪ Iterative Forecast-Analysis Cycle (Data Assimilation) allow us to continually confront models with data

▪ All DA variants are forward-only special cases of State Space model

▪ Forecast Step: Standard DA methods map to uncertainty propagation axes

▪ Analysis step: Do not feel constrained by Kalman, Take Assumptions into your own hands, MCMC often viable