Correcting biased observation model error in data assimilation

Tyrus Berry
Dept. of Mathematical Sciences, GMU

PSU-UMD DA Workshop
June 27, 2017

Joint work with John Harlim, PSU

Bias in Observation Models

- Consider the standard filtering problem,

$$x_i = f(x_{i-1}) + \omega_{i-1}$$
$$y_i = h(x_i) + \eta_i$$

- We assume the true observation function $h(x)$ is unknown
- An approximate model is available $\tilde{h}(x)$ so that

$$y_i = h(x_i) + \eta_i = \tilde{h}(x_i) + b_i + \eta_i$$

- Where $b_i \equiv h(x_i) - \tilde{h}(x_i)$ is called the bias

Example 1: Lorenz-96

- Consider the standard 40-dimensional Lorenz-96,

$$\dot{x}_j = x_{j-1}(x_{j+1} - x_{j-2}) - x_j + 8$$

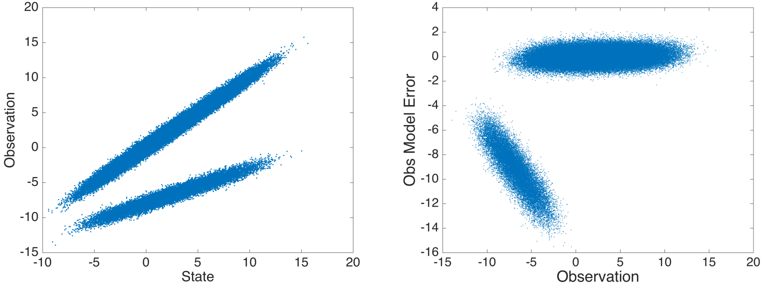
- We observe 20 of the 40 variables
- We draw $\xi_i \sim \mathcal{U}(0,1)$ and let the observations be,

$$h(x_k) = \begin{cases} x_k & \xi_i > 0.8 \\ \beta_k x_k - 8 & \text{else} \end{cases}$$
$$\beta_k \sim \mathcal{N}(0.5, 1/50).$$

- h is applied to 7 randomly chosen variables
- Remaining 13 are directly observed

EXAMPLE 1: LORENZ-96

- The result is a bimodal distribution, “cloudy/clear”
- Obs Model Error = True Obs - $\tilde{h}$(True State)



Correcting the bias

- Our goal is to find $p(b_i \mid y_i)$
- We can then adjust the filter by defining a new innovation

$$\hat{\epsilon}_i = \epsilon_i + \hat{b}_i = y_i - \tilde{h}(x_i^f) + \hat{b}_i$$

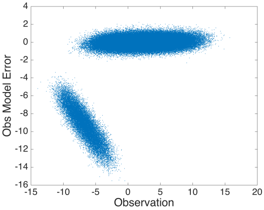
- Where $\hat{b}_i = \mathbb{E}_{p(b_i \mid y_i)}[b_i]$
- We also inflate the obs covariance by $R_i = R^o + \hat{R}_{b_i}$
- Where $\hat{R}_{b_i} = \mathbb{E}_{p(b_i \mid y_i)}[(b_i - \hat{b}_i)(b_i - \hat{b}_i)^\top]$

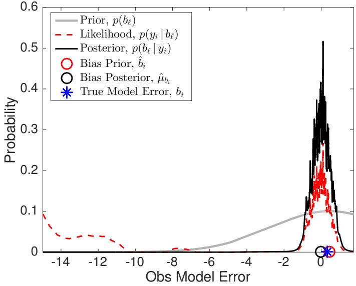
Correcting the bias

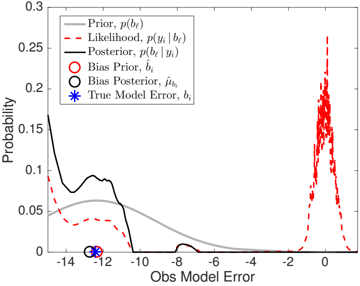
- If we can estimate $p(b_i \mid y_i)$ we can 'fix' the obs
- We will use Bayes' to find $p(b_i \mid y_i) = p(b_i)p(y_i \mid b_i)$
- We will use a simple prior $p(b_i) = \mathcal{N}(\epsilon_i, P_i^y)$
  - $\epsilon_i = y_i - h(x_i^f)$ is the innovation
  - P_i^y is the innovation covariance estimate
- The real challenge is to estimate $p(y_i \mid b_i)$
- We will learn $p(y_i \mid b_i)$ from training data using the kernel estimation of conditional distributions

Correcting the Bias

- Below plots have $y_i \approx -4$
- Left is clear, right is cloudy
- Notice bimodal likelihood







Learning the conditional distribution

- Given training data (y_i, b_i) our goal is to learn $p(y_i \mid b_i)$
- For a kernel $K(\alpha, \beta) = e^{-\frac{||\alpha-\beta||^2}{\delta^2}}$ we define Hilbert spaces

$$\mathcal{H}_y = \left\{ \sum_{i=1}^{N} a_i K(y_i, \cdot) \,:\, \vec{a} \in \mathbb{R}^N \right\}, \mathcal{H}_b = \left\{ \sum_{i=1}^{N} a_i K(b_i, \cdot) \,:\, \vec{a} \in \mathbb{R}^N \right\}$$

- For example the kernel density estimate (KDE) $\hat{q}$ is in $\mathcal{H}_y$

$$\hat{q}(y) = \frac{1}{m_0 N} \sum_{i=1}^{N} K(y_i, y)$$

- Eigenvectors ϕ_ℓ of $K_{ij} = K(y_i, y_j)$ form an orthonormal basis for $\mathcal{H}_y$. Similarly φ_k are a basis for $\mathcal{H}_b$.

Learning the conditional distribution

- We assume that $p(y \mid b)$ can be approximated in $\mathcal{H}_y \otimes \mathcal{H}_b$
- Let $C_{ij}^{yb} = \langle \phi_i, \varphi_j \rangle$ and $C_{ij}^{bb} = \langle \varphi_i, \varphi_j \rangle$ then define

$$C^{y|b} = C^{yb}\left(C^{bb} + \lambda I\right)^{-1}$$

- We can then define a consistent estimator of $p(y \mid b)$ by

$$\hat{p}(y \mid b) = \sum_{i,j=1}^{N} C_{i,j}^{y|b} \phi_i(y) \varphi_j(b) \hat{q}(y)$$

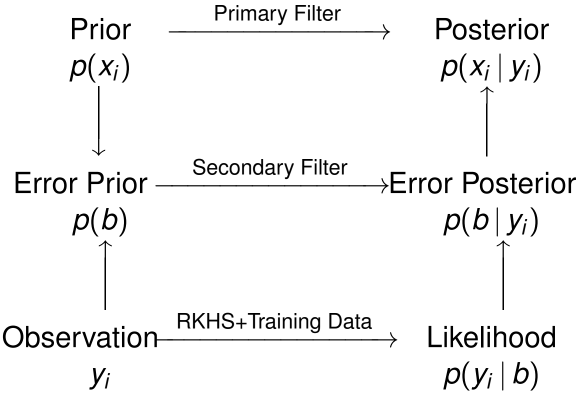
- We define eigenfunctions with Nystöm extension

$$\varphi_j(b) = \lambda_j^{-1} \sum_{i=1}^{N} \varphi_j(b_i) K(b_i, b)$$

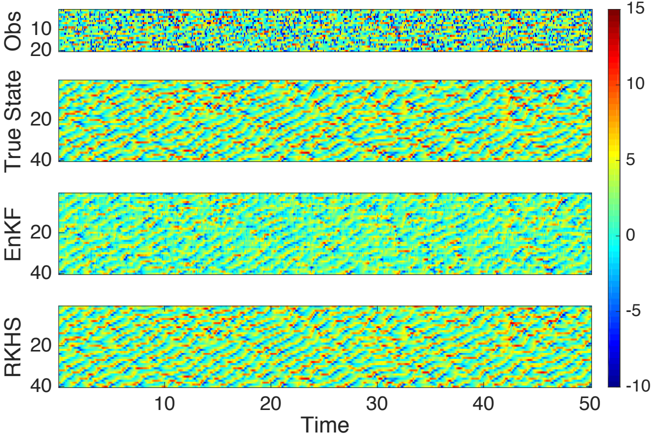
Overview

- **Learning Phase:** Given training data set (x_i, y_i)
  - Compute the biases $b_i = y_i - \tilde{h}(x_i)$
  - Learn the conditional distribution $p(y \mid b)$
- **Filtering:** Forecast $x_i^f \Rightarrow$ innovation $\epsilon_i = y_i - \tilde{h}(x_i^f)$
- Use prior $p(b) = \mathcal{N}(\epsilon_i, P_i^y)$
- Combine with conditional to find $p(b \mid y_i) = p(b)p(y_i \mid b)$
- Estimate conditional mean $\hat{b}_i$ and covariance $\hat{R}_{b_i}$
- Adjust innovation $\hat{\epsilon}_i = \epsilon_i + \hat{b}_i$ and $R_i = R^o + \hat{R}_{b_i}$
- Apply Kalman update, continue to the next filter step

OVERVIEW

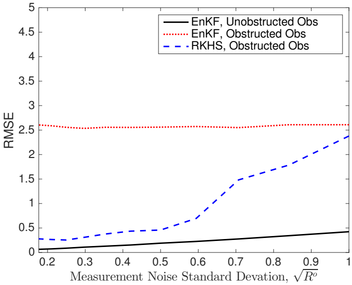

Prior
$p(x_i)$
Primary Filter
Posterior
$p(x_i \mid y_i)$
Error Prior
$p(b)$
Secondary Filter
Error Posterior
$p(b \mid y_i)$
Observation
y_i
RKHS+Training Data
Likelihood
$p(y_i \mid b)$

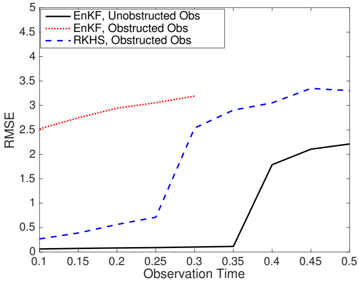
Lorenz-96 Results



LORENZ-96 RESULTS

- Works well with small measurement noise
- Observations need to be precise, but not accurate





Example 2: Multi-cloud "Satellite-like" obs

- Consider a 7-dim'l model for a column of atmosphere
  - Baroclinic anomaly potential temperatures, θ_1 and θ_2
  - Boundary layer anomaly potential temperature, θ_{eb}
  - Vertically averaged water vapor content, q
  - Cloud fractions: congestus f_c, deep f_d, and stratiform f_s
- Extrapolate anomaly potential temperature at height z

$$T(z) = \theta_1 \sin(\frac{z\pi}{Z_T}) + 2\theta_2 \sin(\frac{2z\pi}{Z_T}), \quad z \in [0, 16]$$

Khouider, B., J. Biello, and A. J. Majda, 2010: A stochastic multicloud model for tropical convection.

Example 2: Multi-cloud "Satellite-like" obs

- Extrapolate anomaly potential temperature at height z

$$T(z) = \theta_1 \sin(\frac{z\pi}{Z_T}) + 2\theta_2 \sin(\frac{2z\pi}{Z_T}), \quad z \in [0, 16]$$

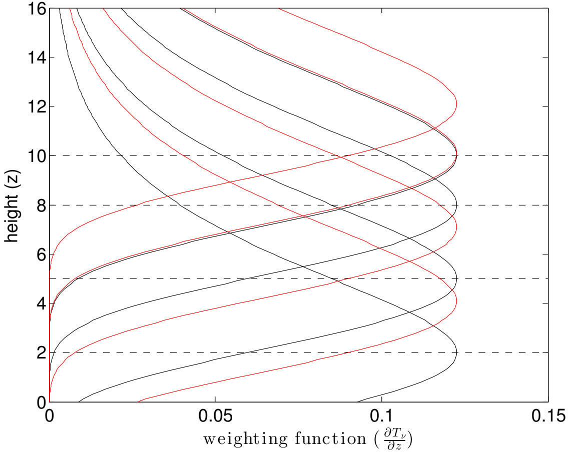
- Brightness temperature-like quantity at wavenumber-ν

$$\begin{aligned} h_\nu(x, f) = (1 - f_d - f_s)\Big[(1 - f_c)\big(\theta_{eb} T_\nu(0) + \int_0^{z_c} T(z)\frac{\partial T_\nu}{\partial z}(z)\,dz\big) \\ + f_c T(z_c) T_\nu(z_c) + \int_{z_c}^{z_d} T(z)\frac{\partial T_\nu}{\partial z}(z)\,dz\Big] \\ + (f_d + f_s) T(z_d) T_\nu(z_d) + \int_{z_d}^{\infty} T(z)\frac{\partial T_\nu}{\partial z}(z)\,dz, \end{aligned} \tag{1}$$

- Setting $f = 0$ is the clear sky model

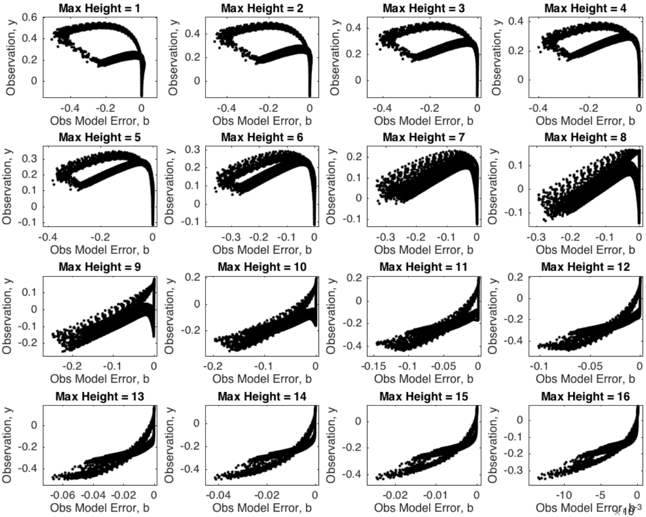
Example 2: Multi-cloud "Satellite-like" obs

- Weighting functions define RTM at different wavenumbers



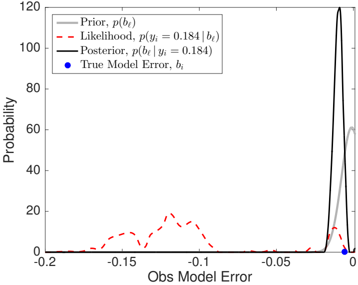
Example 2: Multi-cloud "Satellite-like" obs

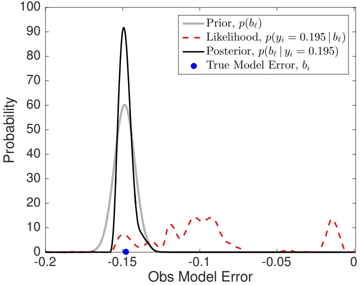
- Biases at the 16 observed wavenumbers



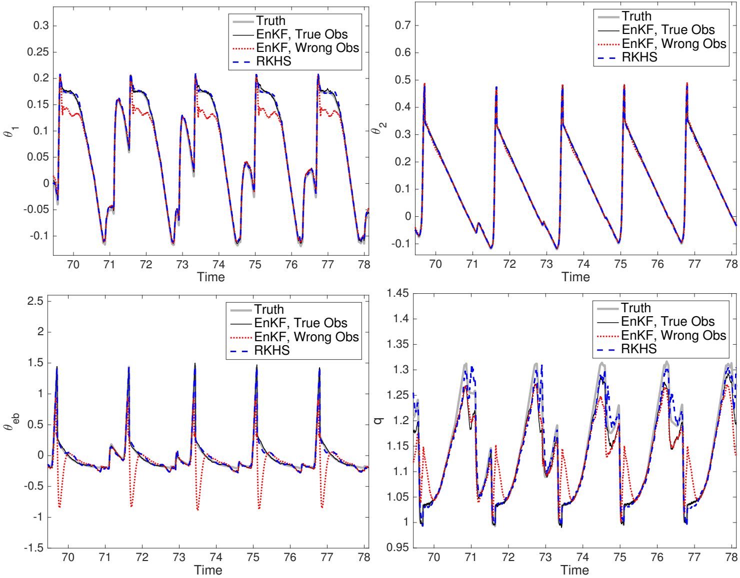
Example 2: Multi-cloud "Satellite-like" obs

- Multimodal likelihood functions

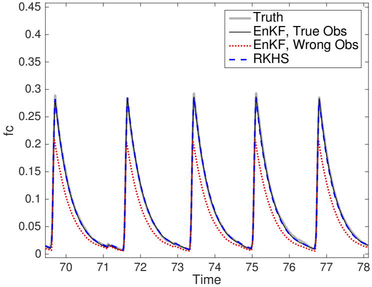


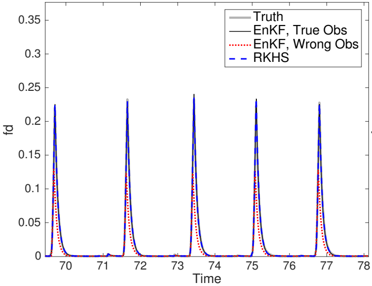


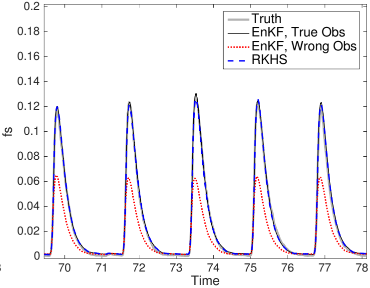
Example 2: Multi-cloud “Satellite-like” obs



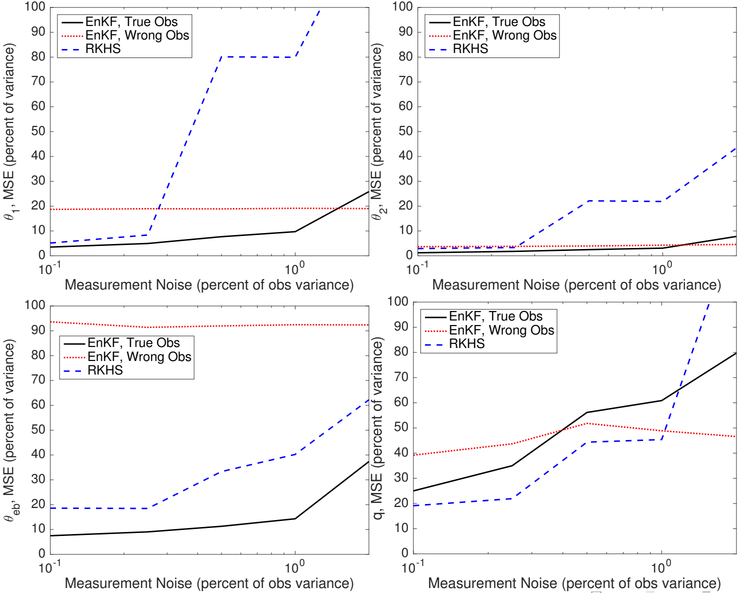
EXAMPLE 2: MULTI-CLOUD "SATELLITE-LIKE" OBS







Example 2: Multi-cloud "Satellite-like" obs



Example 2: Multi-cloud "Satellite-like" obs