Use of the SMART Filter in a Data Assimilation Problem

David Grypma and Sam Pimentel

Faculty of Natural & Applied Sciences, Dept. of Mathematics, Trinity Western University, Canada

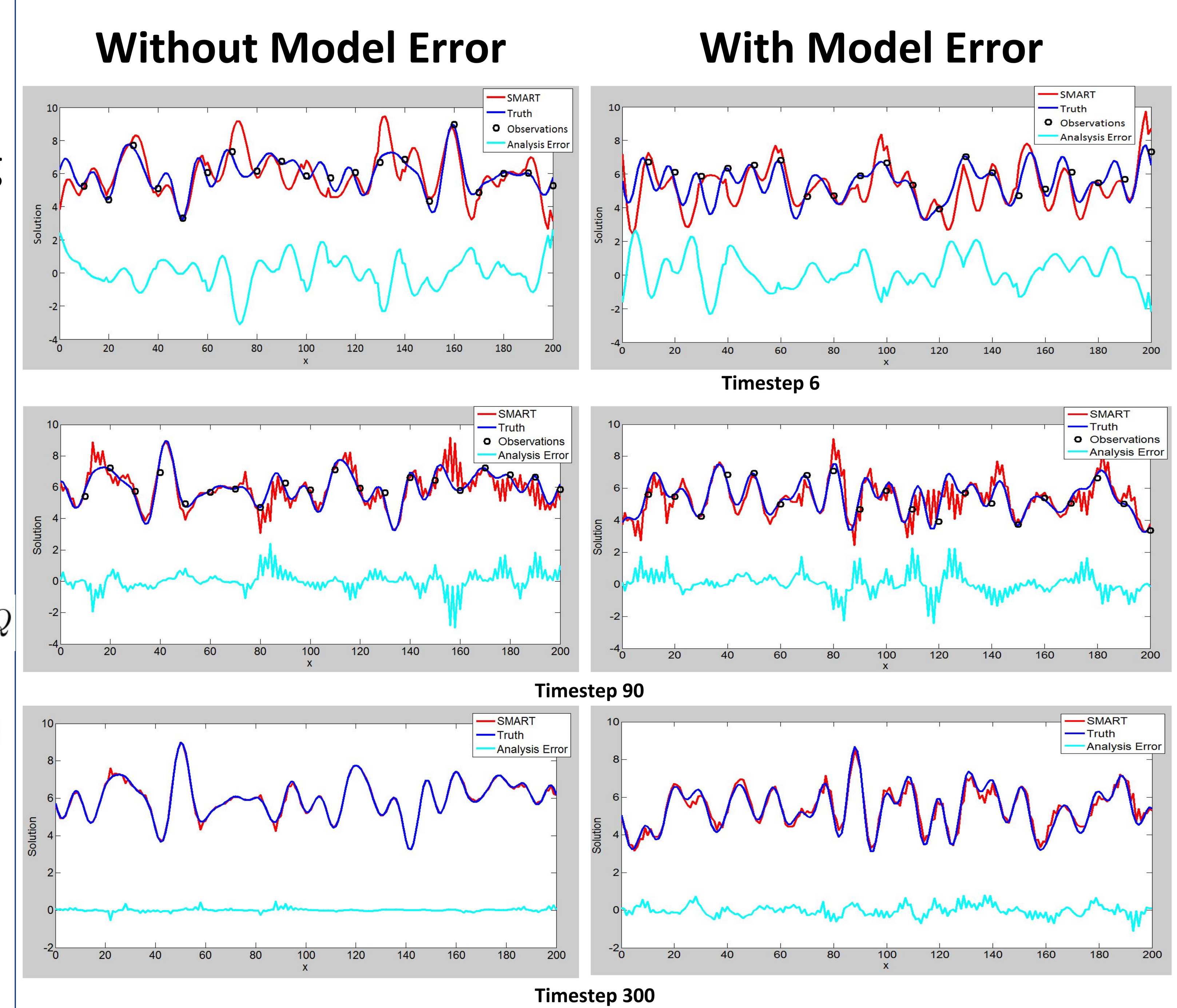
1. Introduction to Problem

- Data assimilation is the process of optimally combining observations and model predictions to produce a best estimate of a system
- Mathematically, the problem of finding the best state estimate, x_k, can be stated as follows

$$min J(x) = \frac{1}{2}||z - Hx||^2_{R^{-1}} + \frac{1}{2}||y - x||^2_{Q^{-1}}$$

Subject to $x_k = y_k + \mu_k$ where the model prediction, $y_k = A_k x_{k-1}$ is subject to model error μ_k $E[\mu] = 0 \quad E[\mu\mu^T] = Q$ and observations in time, z_k, are related to the state, Hx, and contain measurement error $E[\nu] = 0 \quad E[\nu\nu^T] = R$

- Common solution methods include the Kalman Filter and 4D-Var (e.g. see [1])
- The SMART Filter was developed as an inverse method for dynamic medical image reconstruction [2]
- [2] presents the SMART Filter as a computationally efficient alternative to the Kalman Filter
- Here we seek to apply the SMART Filter to a linear advection data assimilation problem



- Under the exact same conditions, for both with and without model error, the SMART was ~ 2 times as fast as a Kalman filter (0.57s vs 1.01s and 0.55s vs 1.05s)

3. Experiment and Results

- We illustrate SMART using a data assimilation problem given in [4], whereby a smooth periodic pseudorandom wave is advected on a periodic domain of length 200 m
- The model has a constant advection speed of 1 ms^{-1}, the grid spacing is equal to 1 m, and a time step of 1 s is used
- The initial condition is offset from the truth by adding a normally distributed error with 0 mean, variance 1, and a de-correlation length of 20m
- We take 20 equally spaced observations of the truth every 6 seconds
- In both experiments we applied normally distributed observation error of mean 0 and variance 0.01
- The solution of the model is known, thus we are able to run simulations without model error (graphs on left panel) and with model error (graphs on right panel)
- The model error is normally distributed of mean 0, variance 0.001, and de-correlation length 20m.
- The length of the total integration is 300 seconds

2. SMART Filter Derivation

- The SMART algorithm [3] solves the regularization problem

$$\min_{x \geq 0} \quad F(x) = \alpha KL(Px, d) + (1 - \alpha) KL(x, y)$$

for a linear system d = Px, and uses cross entropy (or Kullback-Leibler) distance

$$KL(a, b) = a \log \frac{a}{b} + b - a, \quad a, b \geq 0, \quad and, \quad parameter \quad 0 \leq \alpha \leq 1$$

- [2] develop this into a filtering algorithm to solve a temporal regularization problem

$$\min_{x_k \geq 0} \quad F(x_k) = KL(H_k x_k, z_k) + KL(x_k, y_k)$$

- This function can be weighted according to model error covariance, Q, and observation error covariance, R

$$F(x) = KL_{R^{-1}}(Hx, z) + KL_{Q^{-1}}(x, y)$$

- A Cholesky decomposition, where U_1 and U_2 are non-negative diagonal matrices,

$$R^{-1} = U_1^T U_1 \qquad\qquad Q^{-1} = U_2^T U_2$$

allows us to rewrite the cost function as:

$$F(x) = KL(U_1 Hx, U_1 z) + KL(U_2 x, U_2 y)$$

- For our example we take Q and R to be of the form

$$R_k = \hat{\sigma}_k^2 I \qquad\qquad U_1 = \frac{1}{\hat{\sigma}_k} I$$

$$Q_k = \sigma_k^2 I \qquad\qquad U_2 = \frac{1}{\sigma_k} I$$

- Hence

$$\begin{aligned} F(x) &= KL_{R^{-1}}(Hx, z) + KL_{Q^{-1}}(x, y) \\ &= KL(U_1 Hx, U_1 z) + KL(U_2 x, U_2 y) \\ &= KL\left(\frac{1}{\hat{\sigma}} Hx, \frac{1}{\hat{\sigma}} z\right) + KL\left(\frac{1}{\sigma} x, \frac{1}{\sigma} y\right) \end{aligned}$$

- Which can easily be manipulated into the required form:

$$F(x) = \underbrace{\frac{\sigma - 1}{\sigma}}_{\alpha} KL\left(\underbrace{\frac{\sigma}{\hat{\sigma}(\sigma - 1)}}_{P} Hx, \underbrace{\frac{\sigma}{\hat{\sigma}(\sigma - 1)} z}_{d}\right) + \underbrace{\frac{1}{\sigma}}_{1-\alpha} KL(x, y)$$

4. Future Steps

- Currently the SMART filter as implemented here assumes that the model error applied at each update

$$x_k = A_k x_{k-1} + \mu_k$$

is white noise therefore to improve the effectiveness of the SMART filter in our example we should first apply a pre-whitening process, i.e.

$$v = \Lambda^{\frac{-1}{2}} E^T \mu$$

where E is the orthonormal matrix of eigenvectors and Λ is the diagonal matrix of positive eigenvalues of the spectral decomposition of the definite positive covariance matrix Q (see [2] for details)

- We would also like to apply the SMART filter to a nonlinear system

Sources

[1] E. Kalnay, Atmospheric Modelling, Data Assimilation and Predictability. Cambridge: Cambridge University Press, 2003.

[2] J. Qranfal and C. Byrne, SMART filter for dynamic SPECT image reconstruction, International Journal of Pure and Applied Mathematics, 73, 405–434, 2011.

[3] C. L. Byrne, Iterative image reconstruction algorithms based on cross-entropy Minimization, IEEE Trans. On Image Processing, 2, 96-103, 1993.

[4] G. Evensen, Data Assimilation, The Ensemble Kalman Filter, Springer, 2009.

Acknowledgements

This summer research project was funded by a NSERC USRA. We thank the TWU math department for support, Peter Grypma for his MATLAB code and Joel Schwartz for expedient poster printing.