Dynamic Likelihood Approach to Filtering

Juan Restrepo[1]

[1]Affiliation not available

November 21, 2022

Abstract

A Bayesian data assimilation scheme is formulated for advection-dominated or hyperbolic evolutionary problems, and observations. It uses the physics to dynamically update the likelihood in order to extend the impact of the likelihood on the posterior, a strategy that would be particularly useful when the the observation network is sparse in space and time and the associated measurement uncertainties are low. The filter is applied to a problem with linear dynamics and Gaussian statistics, and compared to the exact estimate, a model outcome, and the Kalman filter estimate. By comparing to the exact estimate the dynamic likelihood filter is shown to be superior to model outcomes and to the Kalman estimate, when the observation system is sparse. The added computational expense of the method is linear in the number of observations and thus computationally efficient, suggesting that the method is practical even if the space dimensions of the physical problem are large.

Data Assimilation using Projected Observations

The Dynamic Likelihood Filter (DLF)

Juan M. Restrepo

Department of Mathematics, Department of Statistics,
Physics of Oceans and Atmospheres (CEOAS)
Oregon State University, Corvallis OR 97331 USA

Contact Information:
Department of Mathematics
Oregon State University
Corvallis, OR 97330, USA
Email: `restrepo@math.oregonstate.edu`
URL: `www.math.oregonstate.edu/~restrepo`

Abstract

A Bayesian data assimilation scheme is formulated for advection-dominated or hyperbolic evolutionary problems, and observations. It uses the physics to dynamically update the likelihood in order to extend the impact of the likelihood on the posterior, a strategy that would be particularly useful when the the observation network is sparse in space and time and the associated measurement uncertainties are low. The filter is applied to a problem with linear dynamics and Gaussian statistics, and compared to the exact estimate, a model outcome, and the Kalman filter estimate. By comparing to the exact estimate the dynamic likelihood filter is shown to be superior to model outcomes and to the Kalman estimate, when the observation system is sparse. The added computational expense of the method is linear in the number of observations and thus computationally efficient, suggesting that the method is practical even if the space dimensions of the physical problem are large.

Problem Addressed

Sparse observations can often produce *no* improvements in a data assimilation setting on hyperbolic (wave-like) or advection-dominated problems. The *Dynamics Likelihood Approach* (DLF) to filtering **[1]** exploits the dynamics of hyperbolic systems to extend the range over which observations inform a likelihood. Moreover, the methodology can extend observations into the future, thus allowing Bayesian assimilation of future data assimilation estimates.

Background

Time dependent Bayesian data assimilation combines model outcomes $\mathbf{x}(t) \in \mathbb{R}^N$, $0 \le t < t_f$ and observations $\mathbf{y}(t_m) \in \mathbb{R}^K$, $m = 1, 2, ...,$ $t_m \le t_f$, with the aim at improving estimates of

- **Retrodictions:** $\mathbf{X}(t)$, for $t < t_0$, t_0 is the **present**.
- **Nudictions:** $\mathbf{X}(t)$, for $t = t_0$.
- **Forecasts:** $\mathbf{X}(t)$, for $t \ge t_0$.

The errors inherent in the model outcomes and the observations are taken into account. We obtain estimates $X(t)$ [1, 2] by computing the mean (and variance) of

$$P(\mathbf{x}|\mathbf{y})(t) \propto \Pi_{n=1}^{N_f} P(\mathbf{y}(t_n)\mathbf{x}(t_n))P(\mathbf{x}(t_n))$$

to estimate $\mathbf{X}(t)$, where the likelihood

$$P(\mathbf{y}(t_n)\mathbf{x}(t_n)) = \begin{cases} P(\mathbf{y}(t_m)\mathbf{x}(t_m)), & \text{if} \quad t_m = t_n,\ (t_m \le t_n) \\ 1, & \text{otherwise.} \end{cases}$$

is informed by observations from the past/present.

The Kalman Filter

A sequential model of $\mathbf{V}(t) \approx \mathbf{x}(t)$, and observations $\mathbf{Y}$ are used to produce estimates the **mean** $\langle \mathbf{V} \rangle_n$ and **variance** $\mathbf{P}_n$ via

- **Forecast:**

$$\tilde{\mathbf{V}} = \mathbf{L}_{n-1}\langle \mathbf{V} \rangle_{n-1} + \Delta t \mathbf{f}_{n-1}, \quad n = 1, 2, \ldots, N_f,$$

$\langle \mathbf{V} \rangle_0$, and $\mathbf{P}_0$, known.

- **Analysis:**

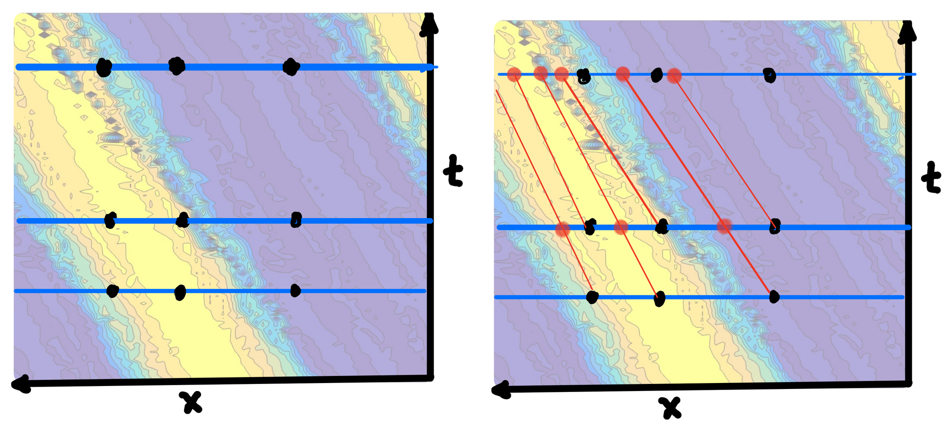
$$\begin{aligned} \langle \mathbf{V} \rangle_n &= \tilde{\mathbf{V}} + \mathbf{K}_m \left(\mathbf{Y}_m - \mathbf{H}_m \tilde{\mathbf{V}} \right), \\ \mathbf{P}_n &= (\mathbf{I} - \mathbf{K}_m \mathbf{H}_m)\tilde{\mathbf{P}}. \end{aligned}$$

The Kalman Gain is defined as

$$\mathbf{K}_m = \tilde{\mathbf{P}}\mathbf{H}_m^\top \left[\mathbf{H}_m \tilde{\mathbf{P}} \mathbf{H}_m^\top + \mathbf{R}_m \right]^{-1} \delta_{n,m}^t,$$

The observation matrices, $\mathbf{H}(t_m) : \mathbb{R}^N \to \mathbb{R}^K$.

The Dynamic Likelihood Filter



Kalman Likelihood uses $\mathbf{Y}(\mathbf{x}, t_i)$. The DLF Likelihood uses $\mathbf{Y}(\boldsymbol{\zeta}, t_i, t)$.

- **Propagate observations and their uncertainties** (using $\mathbf{Y}_t - c(x,t)\mathbf{Y}_x = 0$):

$$\begin{aligned} \mathbf{Y}(\Delta t c(\boldsymbol{\zeta}_n, t_n) + \boldsymbol{\zeta}_n, , t_{n+1}) &= \mathbf{Y}(\boldsymbol{\zeta}_n, t_n), \\ \mathbf{R}_m^{n+1} &= \mathbf{A}_n(t)[\mathbf{A}_n(t)]^\top \Delta t + \mathbf{R}^n, \quad t_n \ge t_m, \end{aligned}$$

with $\boldsymbol{\zeta}_0 = \mathbf{H}(t_m)\mathbf{X}$, $\mathbf{Y}(\boldsymbol{\zeta}_0, t_m) = \mathbf{Y}_m$. and $\mathbf{R}_m^m = \mathbf{R}_m$.

- **Project onto** ***model space*****:**

$$\boldsymbol{\mathcal{H}}_m^n \mathbf{Y}_m^n = \mathbf{V}_n + \boldsymbol{\mathcal{H}}_m^n \boldsymbol{\epsilon}_m^n,$$

at time $t_n \ge t_m$. Here $\boldsymbol{\epsilon}_m^m$ is equal to $\boldsymbol{\epsilon}_m$.

- **Forecast:**

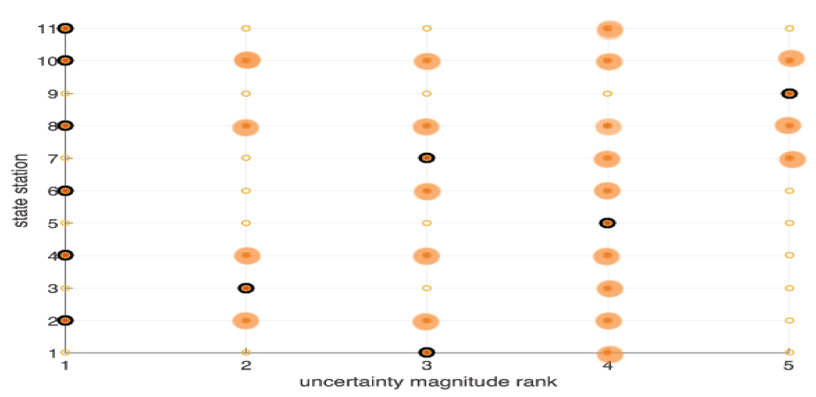
$$\tilde{\mathbf{V}} = \mathbf{L}_{n-1}\langle \mathbf{V} \rangle_{n-1} + \Delta t \mathbf{f}_{n-1}, \quad n = 1, 2, \ldots, N_f,$$

- **Multi-Analysis:**

$$\langle \mathbf{V} \rangle_n = \tilde{\mathbf{V}} + \boldsymbol{\mathcal{K}}_m(\boldsymbol{\mathcal{H}}_m \mathbf{Y}_m - \tilde{\mathbf{V}})\delta_{m,n}^t.$$

$$\mathbf{P}_n = (\mathbf{I} - \delta_{m,n}^t \boldsymbol{\mathcal{K}}_m)\tilde{\mathbf{P}}.$$

$$\boldsymbol{\mathcal{K}}_m = \tilde{\mathbf{P}}(\tilde{\mathbf{P}} + \boldsymbol{\mathcal{H}}_m \mathbf{R}_m \boldsymbol{\mathcal{H}}_m^\top)^{-1}\delta_{m,n}^t.$$



Rank Ordering of Projected and Actual Observations and Uncertainties.

Computational Complexity of the DLF

The *additional* computational load of the method is *linear* in the number of observations $K \ll N$, and the number of time steps N_f:

$$\mathcal{O}(K \times N_f).$$

Computational Example

Aim: Comparison of the Kalman, DLF estimates, a finite-difference model outcome and the exact answer (truth) of

$$\begin{aligned} u_t - C(x,t)u_x &= F(x,t), \quad t > 0, x \in [0, L], \\ u(x, 0) &= \mathcal{U}(x), \quad x \in [0, L], \end{aligned}$$

with random initial conditions, forcing, and phase speed:

$$\begin{aligned} F_\ell \, dt &= f_\ell(x,t)dt + A_\ell(t)dW_\ell^{(f)}(t), \\ C_\ell(x,t) \, dt &= c_\ell(x,t)dt + B_\ell(t)dW_\ell^{(c)}(t), \\ \mathcal{U} &\sim \mathcal{N}(0,1), \quad dW(t)^{(\cdot)} \sim \mathcal{N}(0,1), \text{normal variates,} \end{aligned}$$

given observations (noisy samples of the exact solution).

Observations:

$$\mathbf{Y}(t_m) = \mathbf{H}(t_m)\mathbf{V}(t_m) + \boldsymbol{\epsilon}(t_m), \quad m = 1, ..., M.$$

The observation matrices, $\mathbf{H}(t_m) : \mathbb{R}^N \to \mathbb{R}^K$. The observation errors are normally distributed, with variance

$$\mathbf{R}_m := \langle \boldsymbol{\epsilon}_m \boldsymbol{\epsilon}_{m'}^\top \rangle \delta_{m,m'}.$$

Outcomes

Let the vector $\boldsymbol{\Phi}(t)$ be such that $\Phi_\ell(0) = \mathcal{U}(x_\ell)$. For $\ell = 1, 2, ..., N$,

$$\begin{aligned} \frac{d\Phi_\ell}{dt} &= F_\ell(x,t), \quad t > 0, \\ \Phi_\ell(0) &= \mathcal{U}(x_\ell). \end{aligned}$$

$$\begin{aligned} \frac{dx_\ell(t)}{dt} &= C_\ell(x,t), \quad t > 0, \\ x_\ell(0) &= X_\ell, \quad \ell = 1, 2, ..., N, \end{aligned}$$

Exact (Truth) Outcome:

$$beginequation dx = (\alpha_0 + \alpha_1 t^{1/2})dt + \beta dW,$$

with associated initial conditions. Here α_0 and α_1 are constants. This problem has a solution

$$x_{n+1} = x_n + \alpha_0 \Delta t + \frac{2}{3}\alpha_1 \Delta t^{3/2} + \sqrt{\beta^2 \Delta t}\mathcal{N}(0,1),$$

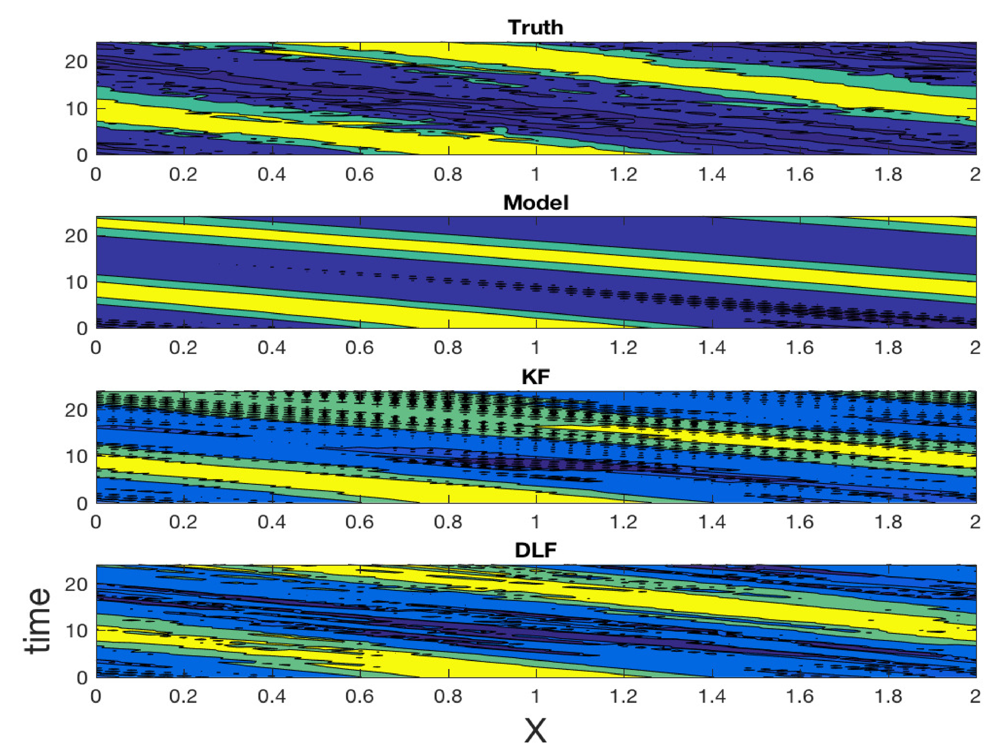
for $n = 0, 1, ...N_f - 1$. The mean and the covariance of the solution are, respectively,

$$\langle x_{n+1} \rangle = \langle x_n \rangle + \alpha_0 \Delta t + \frac{2}{3}\alpha_1 \Delta t^{3/2}, \quad \text{cov}(x_{n+1}) = \text{cov}(x_n) + \beta^2 \Delta t.$$

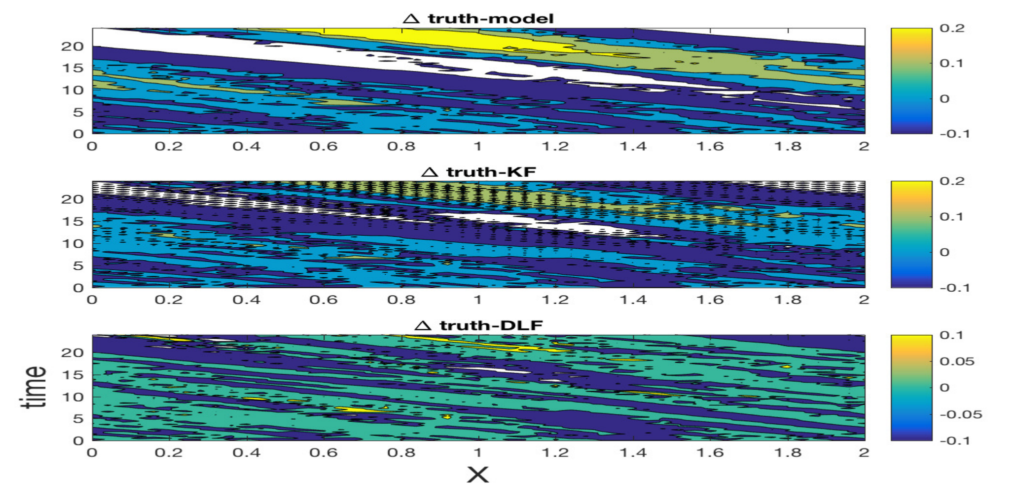
Model Estimate: Lax-Friedrichs,

$$\sqrt{\Delta t}\Delta \mathbf{w}_n = -\mathbf{V}_n + \mathbf{L}_n \mathbf{V}_{n-1} + \Delta t \mathbf{f}_{n-1}, \quad n = 1, 2, \ldots, N_f,$$

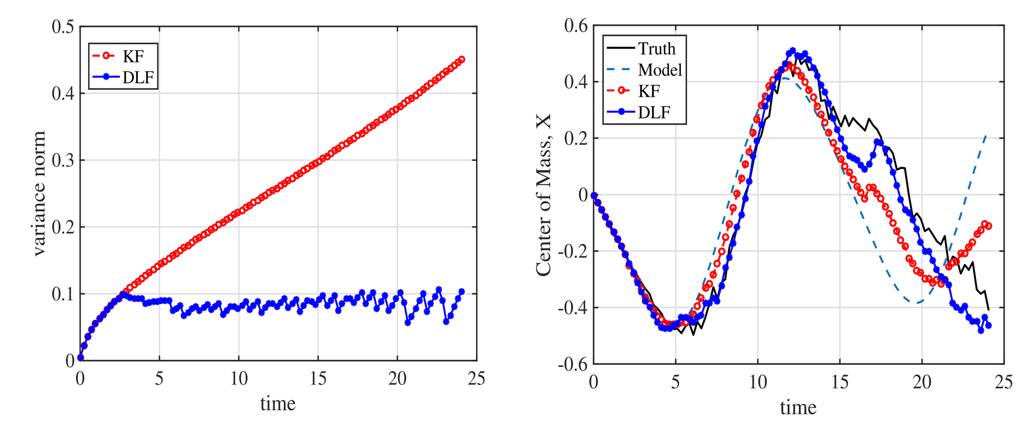
where $\mathbf{L} \in \mathbb{R}^{N \times N}$. Model noise $\Delta \mathbf{w}_n$, normal variates. Model noise variances, $\mathbf{Q}_n = \Delta t \langle \mathbf{w}_n \mathbf{w}_{n'}^\top \rangle \delta_{n,n'}$.



Estimates: Truth, Model, KF and DLF.



Truth-Model, Truth-KF, Truth-DLF.



Uncertainty and **center of mass** of the Truth, Model, KF, DLF.

Data was sampled at every 4 space steps, and every 10 time steps.

Summary

- The DLF is a data assimilation strategy, applicable to hyperbolic, or advection dominated problems, such as linear and nonlinear waves, advected transport.
- The DLF strategy can be applied to linear (Gaussian) as well as nonlinear (non-Gaussian) dynamic problems.
- DLF is computationally-efficient.
- DLF is particularly effective when observations are sparse.
- Unlike other sequential data assimilation schemes, DLF can produce Bayesian forecasts, by projecting observations into the future.

Bibliography

[1] J. M. Restrepo, *A Dynamic Likelihood Approach to Filtering*, Quarterly Journal of the Royal Society of Meteorology, 10.1002/qj.3143 (2017)

[2] G. Eyink, J. M. Restrepo, F. Alexander, *Accelerated Monte-Carlo for Optimal Estimation of Time Series*, Journal of Statistical Physics, **119**, 1331–1345, (2005).

[3] J. M. Restrepo, *A Path Integral Method for Data Assimilation*, Physica D, **237**, pp14–27, (2008).

[4] S. Rosenthal, S. Venkataramani, J. M. Restrepo, A. Mariano, *Displacement Data Assimilation*, Journal of Computational Physics **330**, 594–614, (2017).

.

Acknowledgements

This work was supported by NSF/OCE #1434198 and a PEER research grant, #1123-NCTRYH.