Bridging linear state estimation and machine learning

EGU General Assembly, 25 May 2022

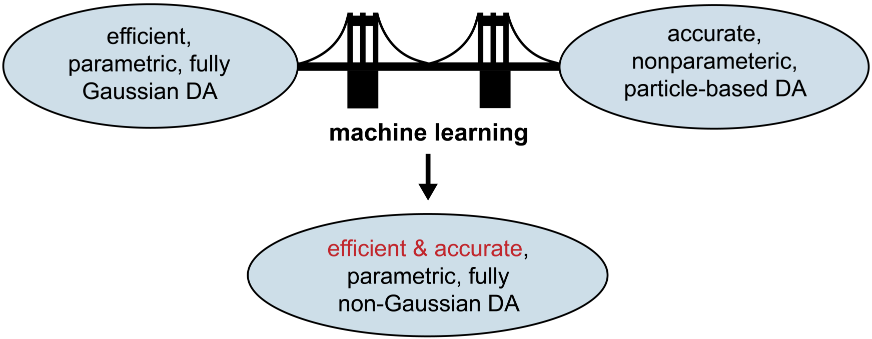
Hristo G. Chipilski
ASP Postdoctoral Fellow

Acknowledgments:
Jeff Anderson, Moha Gharamti, and many others

NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

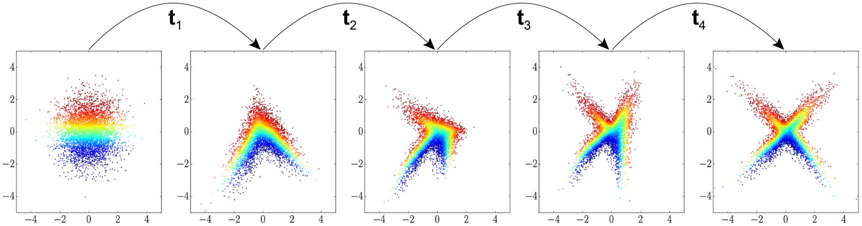
NSF

Objective


efficient, parametric, fully Gaussian DA
accurate, nonparameteric, particle-based DA
machine learning
efficient & accurate, parametric, fully non-Gaussian DA

Introducing non-Gaussianity via ML

- This research was inspired by a subfield of Machine Learning (ML) known as Invertible Neural Networks (INNs).
  - INN functions are diffeomorphisms: invertible and differentiable with differentiable inverses.
- INNs can be used to derive more flexible non-Gaussian distributions:



Adapted from Fig. 1 in Papamakarios et al. (2021)

A new state-space model

- Discrete-time formulation:

$$\begin{aligned} \mathrm{X}_k &= \mathbf{f}_k \left(\mathbf{M}_{k-1}\tilde{\mathrm{x}}_{k-1} + \mathcal{E}^m_{k-1} \right) \\ \mathrm{Y}_k &= \mathbf{g}_k \left(\mathbf{H}_k \tilde{\mathrm{x}}_k + \mathcal{E}^o_k \right) \end{aligned}$$

  - $\{\mathrm{X}_k, \mathrm{Y}_k\}$ and $\{\tilde{\mathrm{X}}_k, \tilde{\mathrm{Y}}_k\}$ are the state and observation processes together with their latent Gaussian representations, respectively.
  - $\mathbf{M}_k$ and $\mathbf{H}_k$ are the linear model and observation operators in the latent Gaussian space.
  - $\{\mathcal{E}^m_k\}$ and $\{\mathcal{E}^o_k\}$ are zero-mean i.i.d Gaussian noises.

- Special structure: The highlighted terms are Gaussian random vectors which represent inputs to some nonlinear diffeomorphisms $\mathbf{f}_k$ and $\mathbf{g}_k$, in analogy to the INN schematic from the previous slide.

Exact nonlinear filtering

- Despite its nonlinear characteristics, the new state-space model (SSM) leads to closed-form solutions:

Theorem

The filtering distribution $\pi_{X_k|Y_{1:k}}$ is given by $\mathbf{f}_{k\,\sharp}\mathcal{N}(\mu_{k|k}, \boldsymbol{\Sigma}_{k|k})$ with

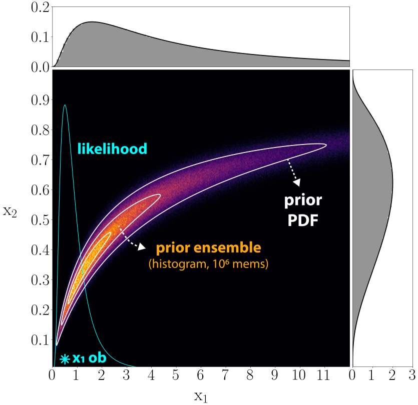
$$\mu_{k|k} = \mu_{k|k-1} + \mathbf{K}_k\left(\tilde{y}_k - \mathbf{H}_k\mu_{k|k-1}\right)$$
$$\boldsymbol{\Sigma}_{k|k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\boldsymbol{\Sigma}_{k|k-1}$$

Properties

1. Generalization of the Kalman filter.
2. The prior is conjugate to the likelihood regardless of how $\mathbf{g}_k$'s are chosen $\Rightarrow$ preservation of dynamical balances after the DA step.

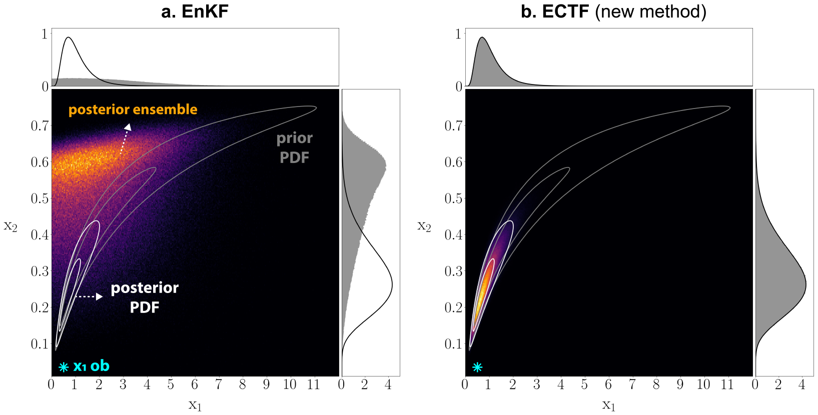
Ensemble Conjugate Transform Filter

- Main idea: Similar to EnKF, construct an ensemble update that converges to the analytically tractable non-Gaussian posterior.
- Numerical example: Prior distribution and observation.



Ensemble Conjugate Transform Filter

- NUMERICAL EXAMPLE: Analysis update.



- RESULT: EnKF produces biased analysis that violates the x_1 bound ($\sim 36\%$ of all members), while ECTF recovers the true posterior.

Summary

- Concepts from machine learning were used to derive an exact non-linear estimation theory which generalizes traditional data assimilation methods to arbitrarily non-Gaussian distributions.
- An ensemble implementation of the new estimation framework (ECTF) was shown to provide significant benefits over the standard EnKF for a challenging geophysical problem.

Future work

- Systematic comparison of ECTF against other nonlinear ensemble filters in idealized statistical experiments.
- Testing different machine learning implementations of ECTF in a hierarchy of dynamical models.

Contact me at hristoc@ucar.edu.