Research Statement

Shashank Kumar Roy, Postdoctoral Researcher
Lab-STICC, IMT Atlantique, Brest, France

Overview

My research lies at the intersection of data assimilation, uncertainty quantification, and machine learning, with applications focusing on geophysical systems. The scientific direction of my research is to develop useful and physically grounded inference methods for complex dynamical systems, where theory, numerics, and deep learning-based models can be combined to improve both predictive skill and uncertainty representation. I am generally fond of computational problems in prediction and inference in the context of dynamical systems, where well-defined mathematical notions, physical theories, statistics and algorithms can be understood through simulations and tested.

Under the Horizon Europe project called Edito Model Lab, which focuses on the European digital twin ocean, my work involves developing algorithms for deep learning-based end-to-end trainable models for ocean data assimilation and forecasting. Currently, I am investigating how to embed uncertainty quantification in end-to-end neural DA schemes in 4DVarNet [1]. Modelling a physically-consistent random perturbation model, I am developing a novel neural DA scheme, 4DVarNet-LU, toward improving both accuracy and uncertainty quantification at inference. In another work, I developed a variational data assimilation based on neural ODE and automatic differentiation formulations to solve weak-4DVar problems for the quasi-geostrophic dynamics, an intermediate-complexity model for data assimilation. Earlier during my PhD, my thesis research work concentrated on data assimilation for chaotic dynamical systems using EnKF[2], where I demonstrated numerical filter stability, a crucial property of a filter, using optimal transport-based distance between probability distributions and stability of computing Lyapunov vectors from erroneous trajectories.

Following is a brief description of my research projects in more details.

Location uncertainty for ensemble generation: 4DVarNet-LU

The emergence of location uncertainty due to incorrect physics or oversimplifying dynamical/statistical assumptions, errors in model parameters, inadequate model resolution, among others is well known [3]. Localised weather phenomena such as cyclones, thunderstorms, precipitations are examples where the features develop but at incorrect locations in space and time. Using this idea to model coherent spatiotemporal random displacements in physical space, we can sample states around a given state.

We integrate this strategy into end-to-end training to provide an initial guess state that reflects both prior knowledge and positional error. Trained deterministically in output and stochastically in input, we demonstrate that at inference, 4DVarNet-LU can generate an ensemble of reconstructions for a set of observations.

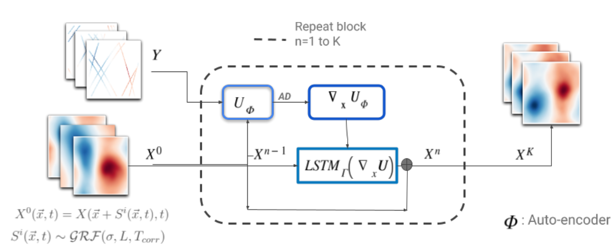

Figure 1: Schematic of 4DVarNet-LU showing the adaptation of original architecture [1] to include an initialisation method based on a random coherent displacement model.

Spatiotemporal coherent random displacements To generate a new perturbed state X^{ptb} from the true state X, we sample a coherent random displacement field $S^i(\vec{x}, t)$ from a GRF, denoted by $\mathcal{GRF}(\sigma_{disp}, L, T)$ and use it to displace the field X in space and time, given by,

$$X^{ptb}(\vec{x}, t) = X(\vec{x} + S^i(\vec{x}, t), t) \tag{1}$$
$$S^i(\vec{x}, t) = (S^i_x(\vec{x}, t), S^i_y(\vec{x}, t)) \sim \mathcal{GRF}(\sigma_{disp}, L, T).$$

The considered random perturbation models location-based uncertainties through spatiotemporal coherent dis-

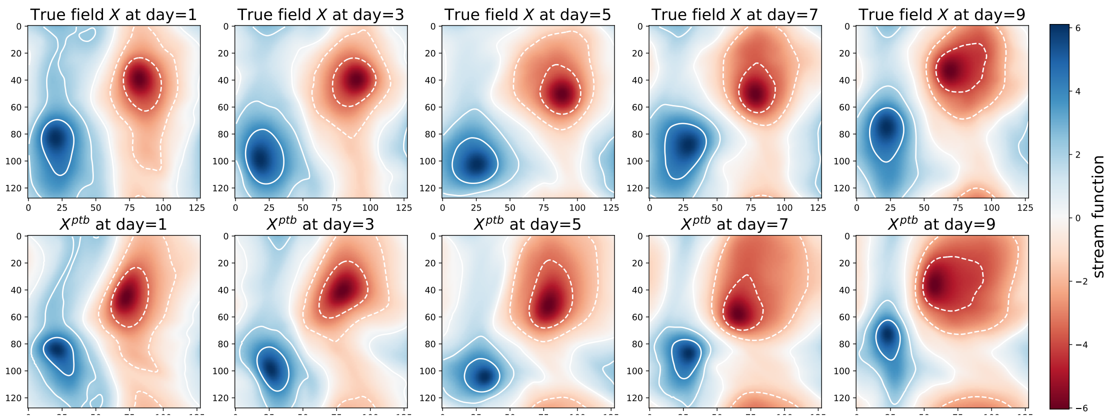

Figure 2: The top row illustrates the true stream function field. The second row shows the perturbed stream function field obtained using the coherent random perturbations with $L = 5.0$, $T = 20.0$ and $\sigma_{disp} = 7.0$. The contour lines show iso-curves at 5 levels.

placements via Gaussian random fields can be further developed together with learnable covariance kernels, particularly adapted for multi-scale systems via the parameters.

Future research directions In the context of deep learning for climate and weather prediction systems, the following are possible directions for developing the current coherent displacement scheme for both the error model and ensemble generations:

1. Developing spatiotemporal displacement with error structure which are based on physical vectors of the system.

2. Understanding the forecasting and uncertainty quantification due to such location uncertainty in the input field for deep learning models.

3. Understanding how methods could be used to capture forward uncertainty growth in forecasts and for ensemble generation alternatives to denoising diffusion probabilistic models.

Numerical filter stability

Bayesian filtering is defined as the sequential estimation of the conditional distribution in phase space of the state of a physical system, taking into account the likelihood of new information arriving from the observations using Bayes' theorem [4]. The true state is unknown, and the choice of $\rho(\mathbf{x_0})$ is arbitrary; it becomes crucial that the conditional distribution of the state becomes independent of the initial choice so that quantities estimated using the posterior are consistent.

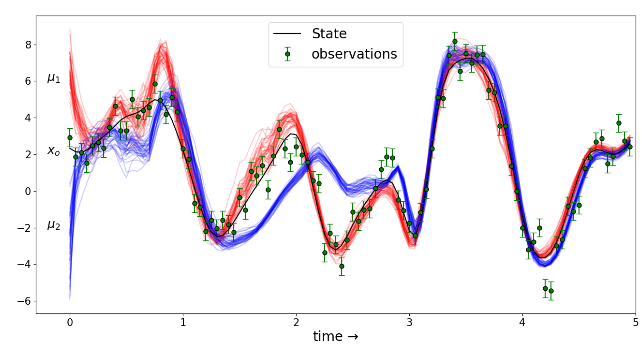

Figure 3: Schematic of filter stability as seen for one component of the state.

Stability of a filtering algorithm.

Filter stability is the property that the conditional posterior distribution computed sequentially over long time is independent of the choice of the distribution at $k = 0$ used to initialise the filtering algorithm. The question we ask is how to numerically check if a filter is stable using a distance D on the space of probability distributions $P(\mathcal{R}^d)$. For two different initial distributions ν_1, ν_2 and their corresponding posterior at time n, denoted by $\hat{\pi}_n(\nu_1)$ and $\hat{\pi}_n(\nu_2)$ should converge over time. In the setting of a deterministic dynamical system, with observation operators h and measurement noise ϵ, we have

$$\mathbf{x}_k = \mathbf{M}(\mathbf{x}_{k-1}), \qquad \mathbf{y}_k = h(\mathbf{x}_k) + \epsilon_k, \qquad \epsilon_k \sim \mathcal{N}(0, \Sigma), \text{ where } \mathbf{x}_k \in R^d, \mathbf{y}_k \in R^p \text{and } p < d. \tag{2}$$

Mathematically,

$$\lim_{n\to\infty} \mathbf{E}[D(\hat{\pi}_n(\nu_1), \hat{\pi}_n(\nu_2))] = 0 \tag{3}$$

where, D is a distance on $P(\mathcal{R}^d)$, the space of probability measures on $\mathcal{R}^d$ and the expectation is over observation noise accounting for different noise realisations [5]. Using Sinkhorn distance [6], on the space of probability distributions, this distance utilises samples from their respective distributions in order to demonstrate and computationally show filter stability for EnKF and Particle filters. I studied the exponential rate of convergence for the distance over time by numerically computing the Sinkhorn distance [7] between the conditional distribution over time . Using different observation gap and observation covariance, I showed that stability for enkf is quite robust to the above two parameters. I also studied numerically the relationship between the rmse, a measure of filter accuracy with the filter stability.

Future research directions Below are some of problems where application of the above ideas of filter stability may be practically useful.

1. A particularly interesting idea is to see how probabilistic machine learning models proposedly performing filtering, satisfy this criteria improving them.
2. The study can be generalized to different filters in context of model and parameterization errors of unmodelled quantities respectively.
3. Relating filter stability to the chaotic properties of the underlying system, since the interplay between the instability and the informative observations lead to eventually capture the conditional distribution.

Covariant Lyapunov vectors

Unstable vectors [8] and corresponding subspace of a system has been shown to improve forecasts in ocean-atmosphere coupled models when the assimilation takes into account [9]. In the setting of data assimilation with sparse and noisy observations, the best estimate of the true state over time comes with a caveat that the filter estimates or the analysis mean over time is not a dynamical trajectory of the model equations. The filter estimate over time as a proxy of the true trajectory perturbed by the error statistics to recover the Lyapunov vectors(LVs) and exponents. This approach led me to study numerical sensitivity of LVs to perturbations in general for a given dynamical system.

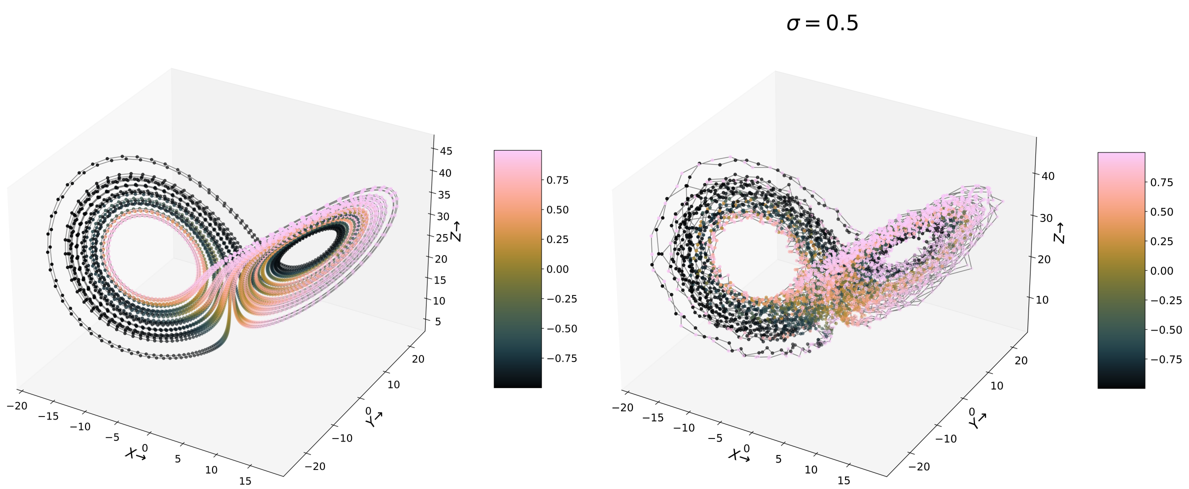

Figure 4: We plot the cosine between first two CLVs for both true trajectory and perturbed trajectory in phase space. The left-most picture corresponds to the true trajectory. σ denotes the standard deviation of noise with zero mean gaussian distribution used to add perturbations.

Sensitivity of Lyapunov vectors The specific questions which I address about lyapunov vectors and the exponents in context of the filtered trajectory are as follows:

- How sensitive are the backward and covariant lyapunov vectors(CLVs)[10] and the corresponding exponents to perturbations in the underlying trajectory?
- Under what conditions can one recover them from a filter estimated trajectory instead of the true trajectory of the dynamical system?
- How robust is the unstable subspace to the perturbation strength σ, and are they more robust than the individual vectors themselves?

In order to start answering the above questions, I studied the effect on the LVs to small perturbations added to the underlying trajectory by systematically adding noise of strength σ following a gaussian distribution $\mathcal{N}(0, \sigma^2 I_d)$, where d is the dimension of the state. Computed using Ginelli's algorithm[11], the angles between the respective LV obtained from the true and the perturbed trajectory help us understand the limitations of such vectors obtained from the numerical state estimates of the filter. In small dimension, I used Lorenz-63, where visualisation and interpretation is straight forward figure 4, where the angle between the first two clvs have been shown to predict regime change in L63 system [12]. Additionally, I studied Lorenz-96 in order to explore dimensional dependence added to the sensitivity problem.
Another interesting directions is using principle angles [13], which summarize the angle between two different subspaces, which I found to be more robust than the individual vectors themselves. Such analysis is useful in the context of problems where subspaces are more important than the individual vectors themselves.
Possible directions for research A set of directions for future work which can be directly extended from the current work are as follows:

1. Using the degree of similarity in the lyapunov vectors between two nearby points on the attractor in phase space for supervised machine learning methods to predict vectors at neaby points in the phase space.

2. Studying structure of CLVs for discretized pde systems such as kuramoto-sivashinksy equation which has a finite dimensional attractor to shed light on CLV localization problem in data assimilation algorithms such as EnKF.

Future research vision

Looking forward, I envision myself working on problems at the intersection of climate and weather models, both numerical and data-driven deep learning models, to solve complex scientific problems of global importance in the field of geosciences. Problems at the interface of climate model, data, algorithms and society are something that inspires my long-term vision.

- Development and evaluation of future weather prediction and climate models that focus on application to both regional and global scale impacts.

- Understanding the key drivers of physics-based models and how they are helpful in the control/ correction of data-driven models, in terms of accuracy, uncertainty and risk.

- Estimation and prediction algorithms of extreme events with a focus on their impact and early warning systems.

I think that the skills and understanding which I have acquired over my small academic projects, conferences and workshops make me confident to handle complex scientific challenges. Technically, developing methods to handle and model uncertainty, dynamical knowledge, combined with general machine learning techniques for the dynamical system of the Earth, is of practical importance. Here, working with big datasets is ubiquitous, and discovering problem-specific understanding and representations benefits from available deep learning and data science tools. Understanding how new developments in deep learning theory and algorithms, and applying them to geoscientific research, is where I can lend my current skill set for collaborations in new directions of research avenues.

References

[1] R. Fablet, B. Chapron, L. Drumetz, E. Mémin, O. Pannekoucke, and F. Rousseau, "Learning variational data assimilation models and solvers," *Journal of Advances in Modeling Earth Systems*, vol. 13, no. 10, p. e2021MS002572, 2021, e2021MS002572 2021MS002572. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021MS002572

[2] G. Evensen, "The ensemble kalman filter: theoretical formulation and practical implementation," *Ocean Dynamics*, 2003. [Online]. Available: https://doi.org/10.1007/s10236-003-0036-9

[3] S. Ravela, K. Emanuel, and D. McLaughlin, "Data assimilation by field alignment," *Physica D: Nonlinear Phenomena*, vol. 230, no. 1, pp. 127–145, 2007, data Assimilation. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167278906003551

[4] S. E. Cohn, "An introduction to estimation theory (gtspecial issueltdata assimilation in meteology and oceanography: Theory and practice)," *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 1B, pp. 257–288, 1997.

[5] P. Mandal, S. K. Roy, and A. Apte, “Stability of nonlinear filters-numerical explorations of particle and ensemble kalman filters,” in *2021 Seventh Indian Control Conference (ICC)*. IEEE, 2021, pp. 307–312.

[6] J. Feydy, T. Séjourné, F.-X. Vialard, S.-i. Amari, A. Trouvé, and G. Peyré, “Interpolating between optimal transport and mmd using sinkhorn divergences,” in *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR, 2019, pp. 2681–2690.

[7] A. Thibault, L. Chizat, C. Dossal, and N. Papadakis, “Overrelaxed sinkhorn–knopp algorithm for regularized optimal transport,” *Algorithms*, vol. 14, no. 5, p. 143, 2021.

[8] T. N. Palmer and L. Zanna, “Singular vectors, predictability and ensemble forecasting for weather and climate,” *Journal of Physics A: Mathematical and Theoretical*, vol. 46, no. 25, p. 254018, jun 2013. [Online]. Available: https://doi.org/10.1088/1751-8113/46/25/254018

[9] S. Vannitsem and V. Lucarini, “Statistical and dynamical properties of covariant lyapunov vectors in a coupled atmosphere-ocean model-multiscale effects, geometric degeneracy, and error dynamics,” *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 22, p. 224001, may 2016. [Online]. Available: https://doi.org/10.1088/1751-8113/49/22/224001

[10] D. Pazó, I. G. Szendro, J. M. López, and M. A. Rodríguez, “Structure of characteristic lyapunov vectors in spatiotemporal chaos,” *Phys. Rev. E*, vol. 78, p. 016209, Jul 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.78.016209

[11] F. Ginelli, H. Chaté, R. Livi, and A. Politi, “Covariant lyapunov vectors,” *Journal of Physics A: Mathematical and Theoretical*, vol. 46, no. 25, p. 254005, jun 2013. [Online]. Available: https://doi.org/10.1088/1751-8113/46/25/254005

[12] E. L. Brugnago, J. A. C. Gallas, and M. W. Beims, “Predicting regime changes and durations in lorenz's atmospheric convection model,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 10, p. 103109, 2020. [Online]. Available: https://doi.org/10.1063/5.0013253

[13] G. H. Golub and H. Z. ha, “The canonical correlations of matrix pairs and their numerical computation,” 1992.