ISDA-Online Special Event
Friday, January 12, 2024 from 15-17 UTC

"Machine Learning in Data Assimilation"

During the last few years, we witnessed a fast development on the use of Machine Learning (ML) approaches in combination with Data Assimilation (DA). This special event will feature invited presentations by distinguished ML and DA experts, namely Alan Geer, Stephan Rasp, and Ronan Fablet. The speakers will provide an overview of the state and potential of these developments in their domain, both from the methodological and applied perspectives. Overall, this special event aims to unveil synergies and differences between ML and DA and to highlight that Machine Learning is not just a buzzword but a powerful tool for analysing, predicting, and understanding environmental systems.

Organizers: Tobias Necker, Nora Schenk, Javier Amezcua, James Taylor, Lars Nerger

Invited Speakers: Alan Geer (ECMWF)
Stephan Rasp (Google Research)
Ronan Fablet (IMT Atlantique)

Invited Chairs: Yvonne Ruckstuhl (KU Eichstätt-Ingolstadt)
Walter Acevedo Valencia (DWD)

Program: *(each talk 30 min + 5 min q&a)*

15:00 – 15:05 Welcome

15:05 – 15:40 Combining machine learning and data assimilation at ECMWF
Alan Geer

15:40 – 16:15 The second revolution in numerical weather prediction and its implications for data assimilation
Stephan Rasp

16:15 – 16:50 End-to-end neural data assimilation: from toy examples to uncertainty quantification with real ocean data
Ronan Fablet

(5 min time buffer)

16:55 – 17:00 Closing / Information on upcoming events

Please note:

- The times in UTC are approximate. In case of technical problems, we might have to change the order of the presentations.
- Login to the session is possible from 20 minutes before the event starts. When you login before 15:00 UTC, and everything is quiet, this is most likely because we muted the microphones.

Time zones: 15 – 17 UTC

Europe: 03 – 05 pm GMT (London) | 04 – 06 pm CET (Berlin)
Asia/Australia: 11 – 01 am CST (Shanghai) | 00 – 02 am JST (Tokyo) | 02 – 04 am AEDT (Sydney)
Americas: 07 – 09 am PST (San Fran.) | 08 – 10 am MST (Denver) | 10 – 12 am EST (NewYork)

Abstract Talk 1

Combining machine learning and data assimilation at ECMWF

Alan Geer

ECMWF, UK

Viewed from a Bayesian inverse perspective, machine learning (ML) and data assimilation (DA) are two sides of the same coin. Given observations of the earth system, ML typically learns models that are unconstrained apart from their inputs and outputs. DA typically learns the system state and uses a fixed model based on physical equations. But once ML starts including physical constraints, or DA starts learning model parameters, ML and DA become increasingly hard to tell apart. Perhaps the remaining philosophical difference betwen ML and DA is the extent to which prior knowlege, in the form of known physical equations, is considered useful in constraining the solution of the inverse problem. These ideas are being explored practically at ECMWF. By abandoning most physical constraints, the machine learning equivalent to the forecast model is currently able to generate better forecasts, suggesting that physically-based forecast is being degraded by substantial model errors. To address this without throwing away the physical model, we could use machine learning to dynamically correct the errors in the physical model. But there is also the more granular option of seeing the model as a mix of physical and empirical components. Where the physics is well known, it can be described by a fixed physical model. Where the physics is poorly known, and our models come from earth observations anyway, it can be treated as empirical. Such empirical models can be learned most optimally from observations within the wider DA system. This latter approach is explored through a new sea ice observation modelling framework that mixes empirical and physical state variables along with empirical and physical model components. Looking back to Bayes' theorem, the optimal solution to the inverse problem is likely to require all prior physical knowledge and all information that is available from observations.

Abstract Talk 2

The second revolution in numerical weather prediction and its implications for data assimilation

Stephan Rasp
Google Research, US / Germany
https://sites.research.google/weatherbench/

Machine Learning (ML) weather models can now product forecasts with similar skill to traditional, physics-based methods. In this talk, I will review the state of the art in ML-based weather prediction and speculate what implications these advances could have for data assimilation.

Abstract Talk 3

End-to-end neural data assimilation: from toy examples to uncertainty quantification with real ocean data

Ronan Fablet
Prof. IMT Atlantique, Lab-STICC, INRIA Team Odyssey
https://cia-oceanix.github.io/

The exploration of deep learning schemes for data assimilation has received a growing attention. Recent studies suggest possible breakthroughs to better bridge models and observations. This includes both computational acceleration, performance gain for reanalysis and forecasting problems as well as the ability to better exploit the richness of observation datasets. This talk will focus on end-to-end neural data assimilation. From a methodological point of view, it will address topics at the crossroads of variational data assimilation, variational Bayesian inference, meta-learning and uncertainty quantification. Toy examples as well as ocean case-studies will illustrate the relevance of the proposed trainable DA frameworks.

Related references:

- *R. Fablet et al. Learning Variational Data Assimilation Models and Solvers. JAMES, 2021.*
- *N. Lafon et al. Uncertainty quantification when learning dynamical models and solvers with variational methods. JAMES, 2023*
- *R. Fablet et al. Multimodal 4DVarNets for the reconstruction of sea surface dynamics from SST-SSH synergies. IEEE TGRS, 2023.*
- *M. Beauchamp et al. Learning Neural Optimal Interpolation Models and Solvers. Proc. ICCS'2023. LNCS 14076, 2023.*
- *M. Beauchamp et al. Neural SPDE solver for uncertainty quantification in high-dimensional space-time dynamics. arXiv, 2023.*