Bridging Physics and Learning: application to ocean dynamics

R. Fablet et al.
ronan.fablet@imt-atlantique.fr
web: rfablet.github.io

Webinar IMT Data & AI, July 2020

Lab-STICC
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

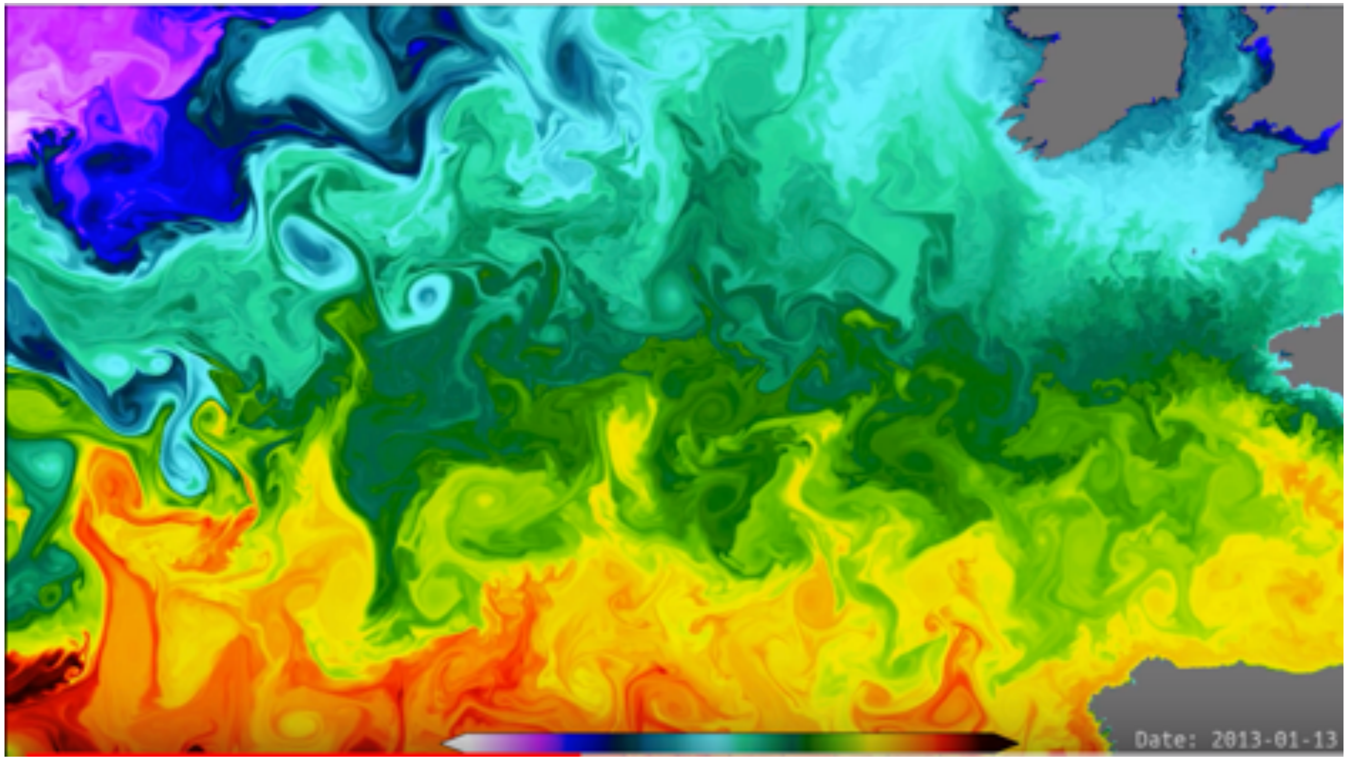
Why the ocean(s) as sandbox ?

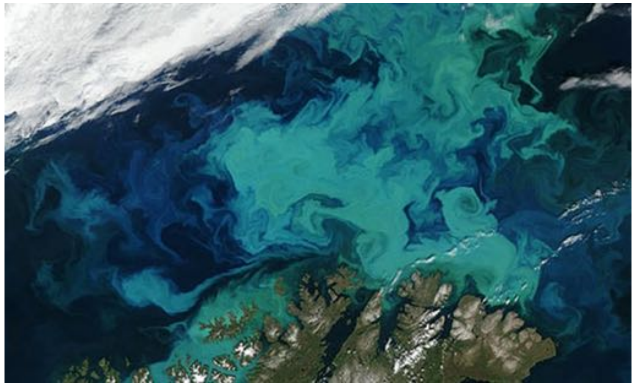
ML/DL & Physics

Focus on Learning variational models for inverse problems

Beyond Ocean Dynamics

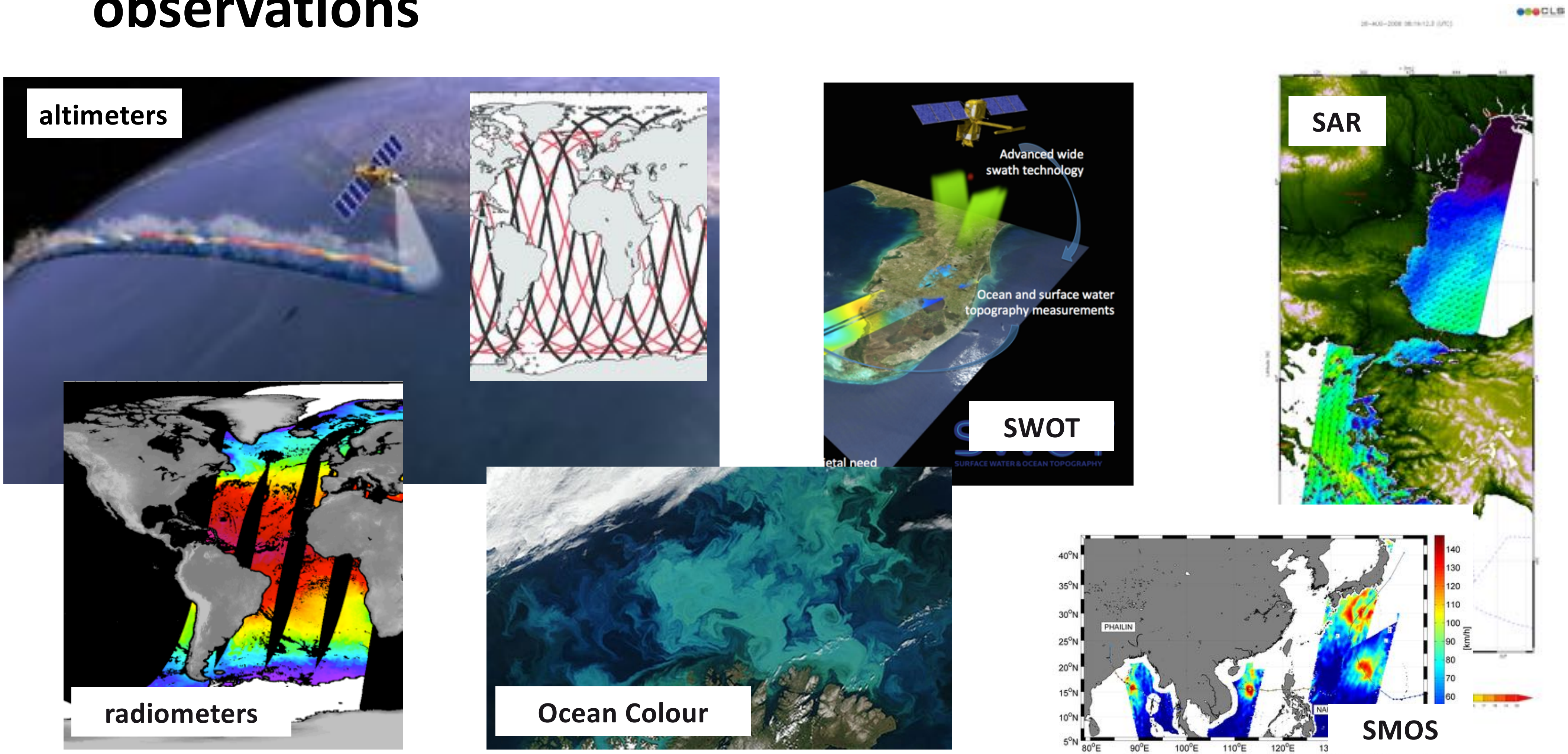
Context: No observation / simulation system to resolve all scales and processes simultaneously



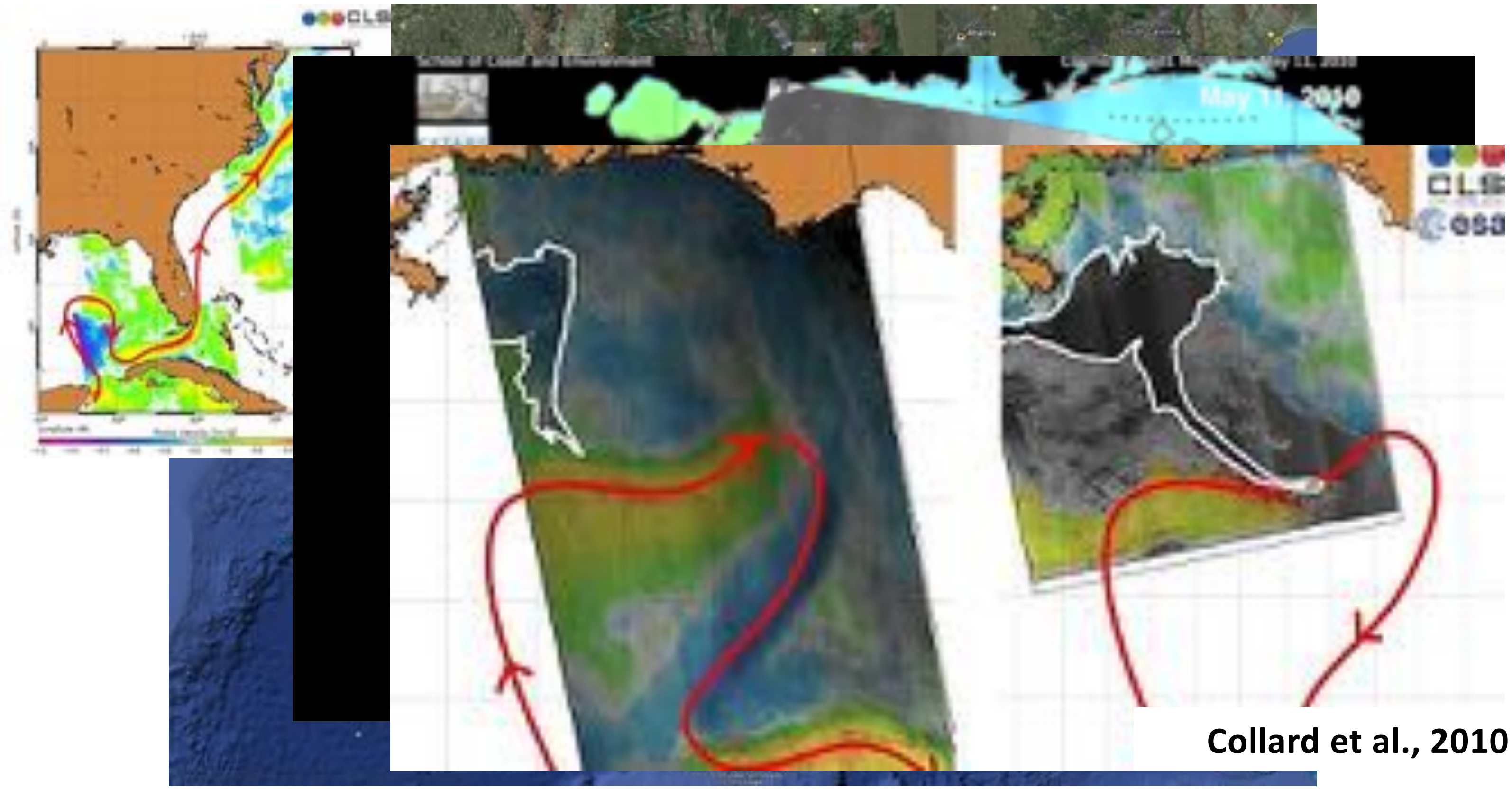


General question: how could data-driven/learning-based tools contribute to solving sampling gaps and higher-level information ?

Illustration of satellite-derived sea surface observations

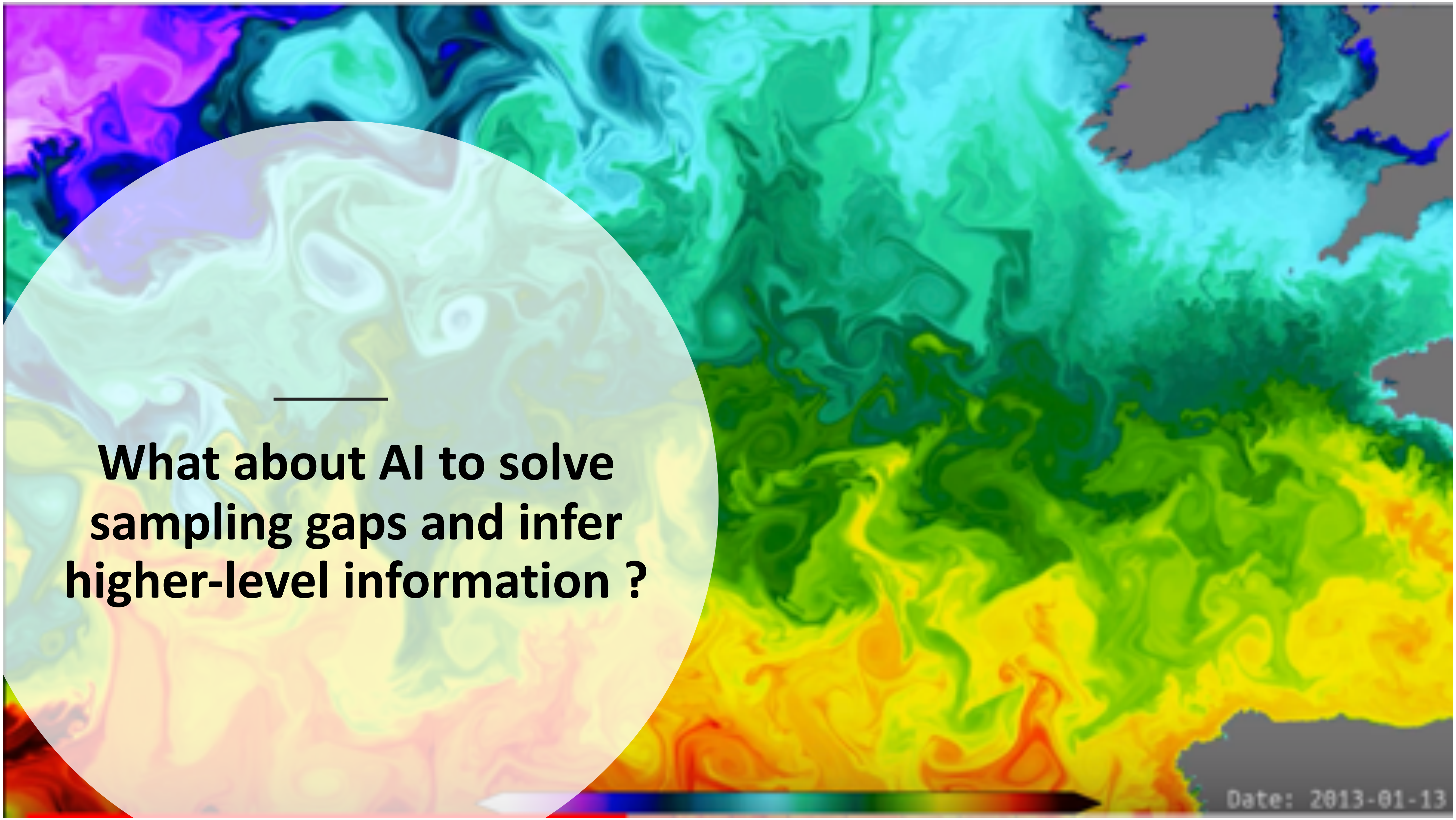


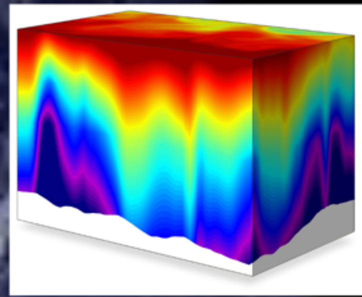
Deepwater horizon [2010]



Collard et al., 2010

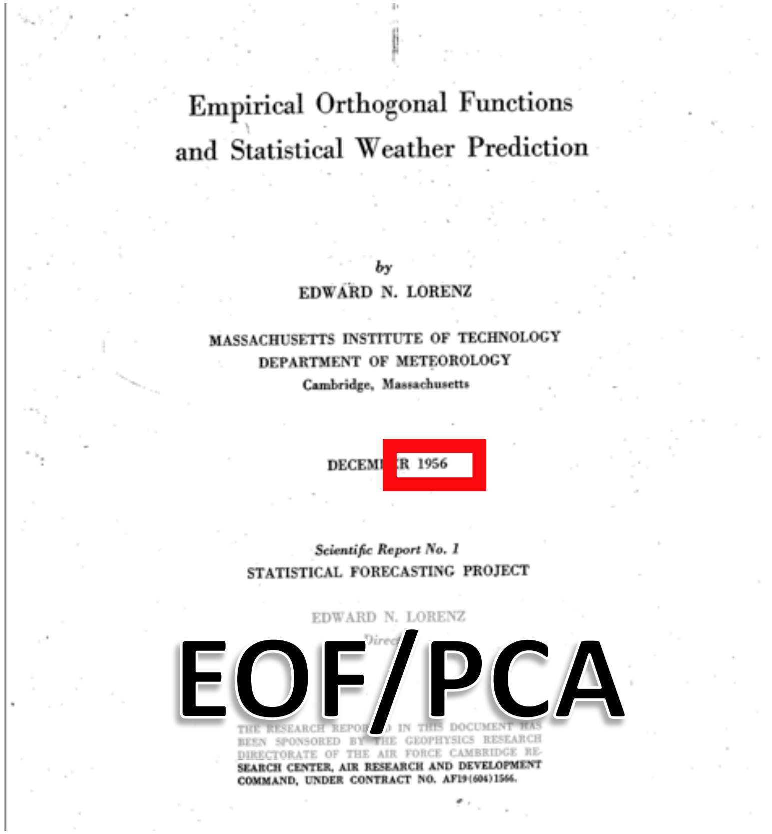
What about AI to solve sampling gaps and infer higher-level information ?





Context: Data-driven and learning-based approaches for ocean monitoring & surveillance

Learning & Geoscience: nothing new ?



Empirical Orthogonal Functions and Statistical Weather Prediction

by

EDWARD N. LORENZ

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF METEOROLOGY
Cambridge, Massachusetts

DECEMBER 1956

Scientific Report No. 1
STATISTICAL FORECASTING PROJECT

EDWARD N. LORENZ
Director

THE RESEARCH REPORTED IN THIS DOCUMENT HAS BEEN SPONSORED BY THE GEOPHYSICS RESEARCH DIRECTORATE OF THE AIR FORCE CAMBRIDGE RESEARCH CENTER, AIR RESEARCH AND DEVELOPMENT COMMAND, UNDER CONTRACT NO. AF19(604)1566.

EOF/PCA

Deterministic Nonperiodic Flow[1]

EDWARD N. LORENZ

Massachusetts Institute of Technology

(Manuscript received 18 November 1962, in revised form January 1963)

ABSTRACT

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions.

A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.

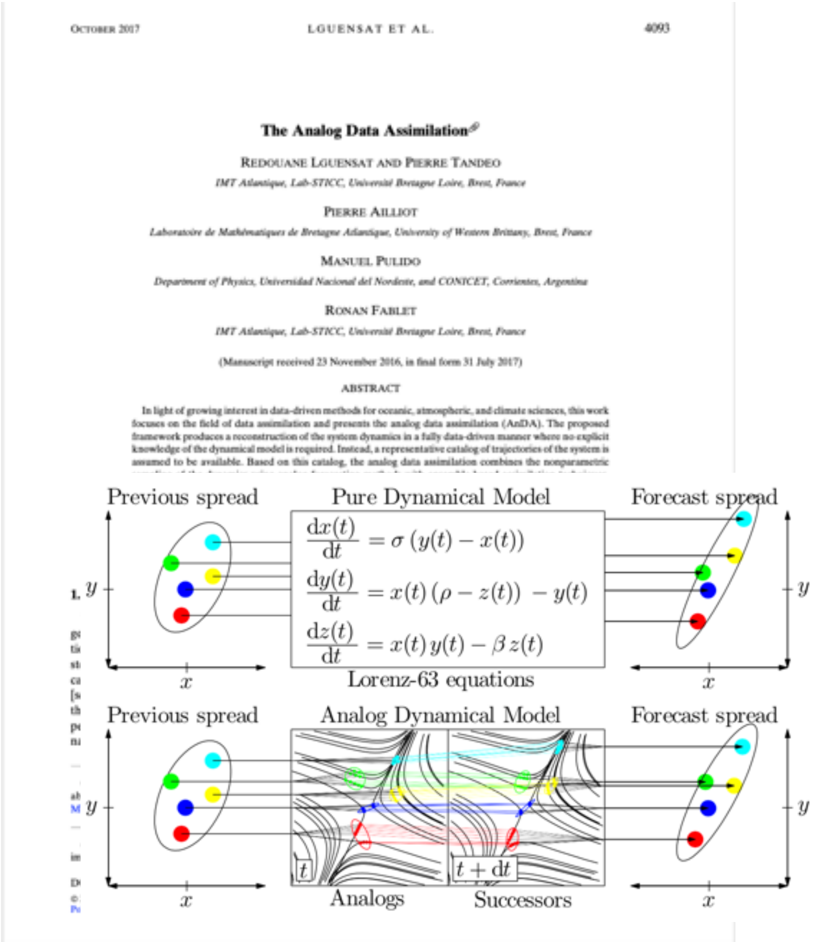
1. Introduction

Certain hydrodynamical systems exhibit steady-state flow patterns, while others oscillate in a regular periodic fashion. Still others vary in an irregular, seemingly haphazard manner, and, even when observed for long periods of time, do not appear to repeat their previous history.

Thus there are occasions when more than the statistics of irregular flow are of very real concern.

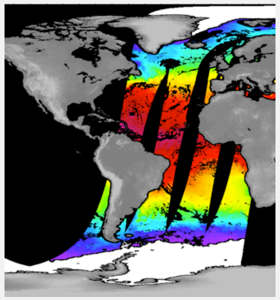
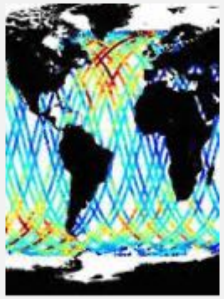
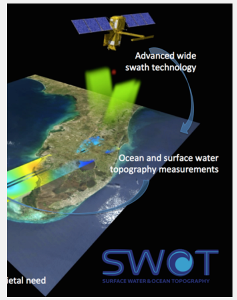
Analogs / Nearest-neighbors

Learning & Geoscience: Data-driven approaches for data assimilation



The analog data assimilation [Lguensat et al., 2017]

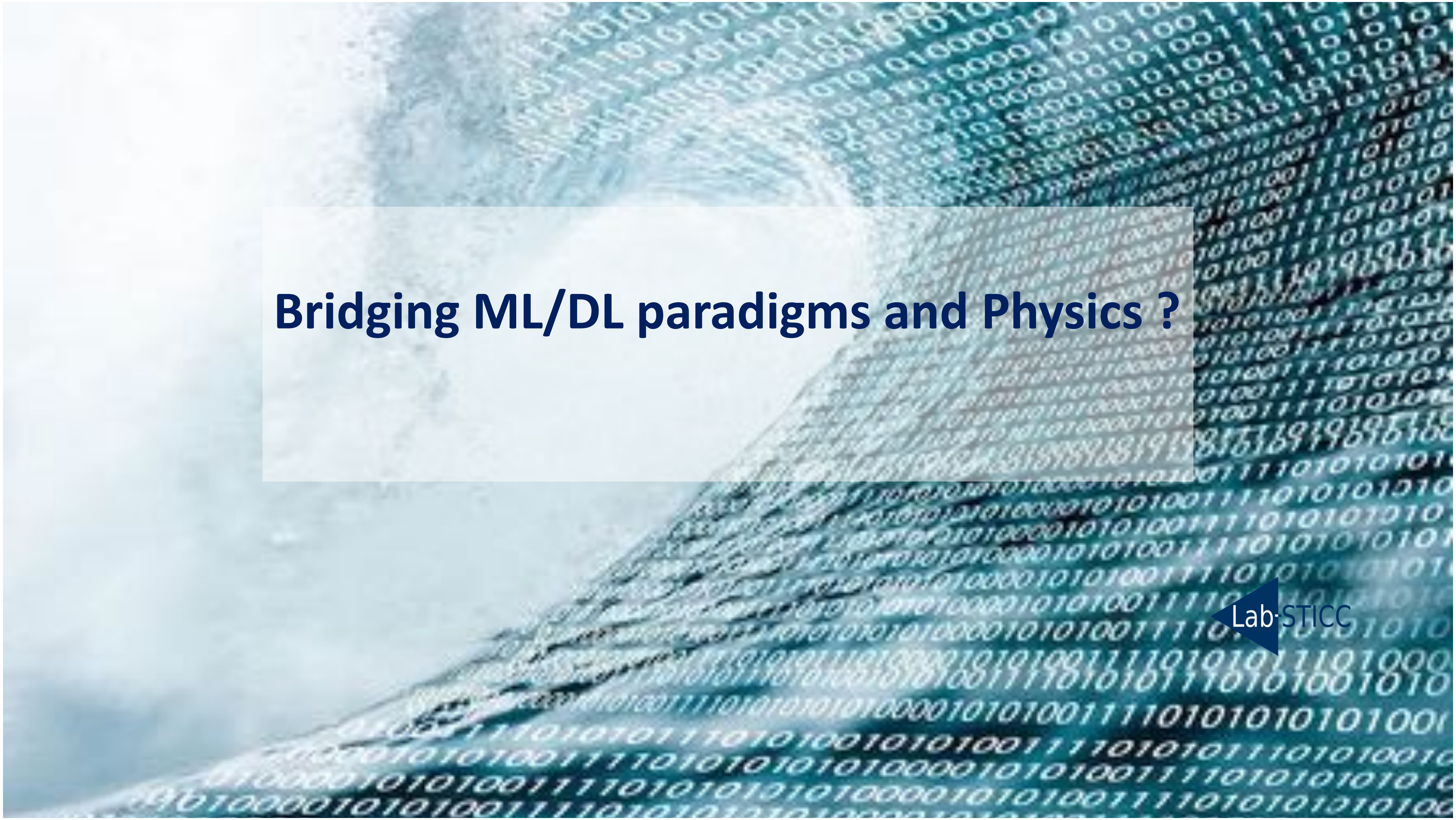
- Combination of analog forecasting strategies and EnKF assimilation schemes
- Extension to 2D+t geophysical dynamics

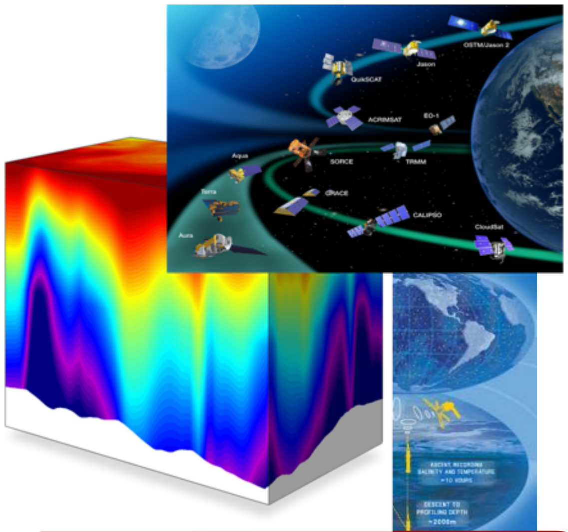
Open questions

- Bridging model-driven and data-driven paradigms
- Learning data-driven representations from real observation data

Bridging ML/DL paradigms and Physics ?



Bridging Physics & AI: Expected breakthroughs



Model-data synergies

AI toolbox

Machine (Deep) learning

Reinforcement learning

Automatic differentiation

GPU

Smart observing systems

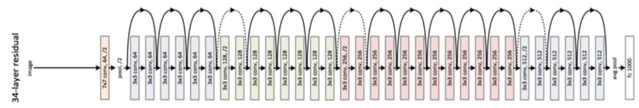
Bridging Physics & AI: Expected breakthroughs

Representation learning

Physical model

$$\frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u$$

Data-driven representation



AI toolbox

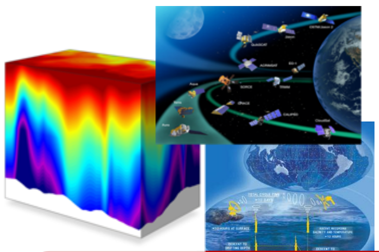
Machine (Deep) learning

Computational acceleration

Big data

Automatic differentiation

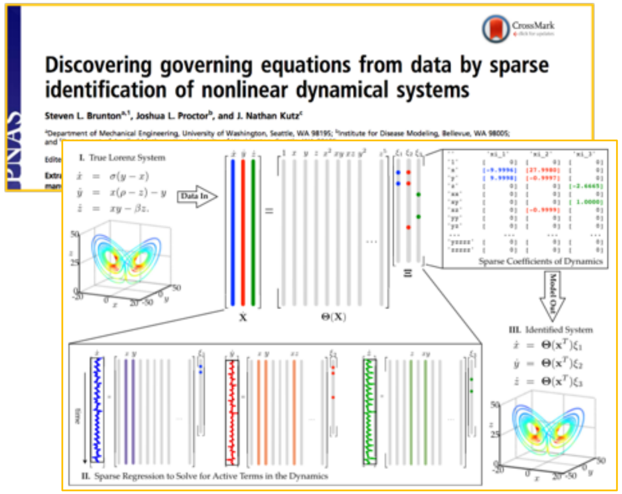
GPU

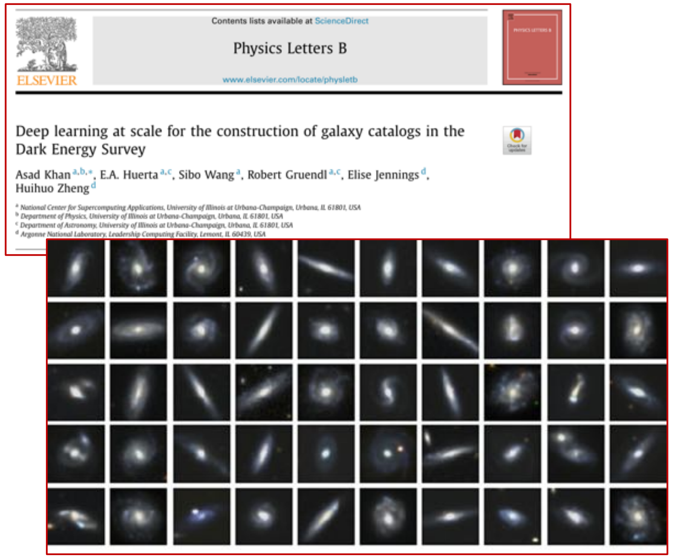


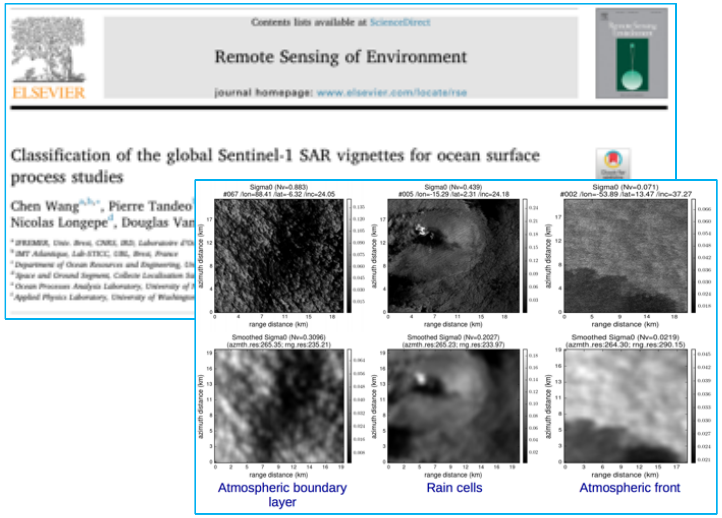
Model-data synergies

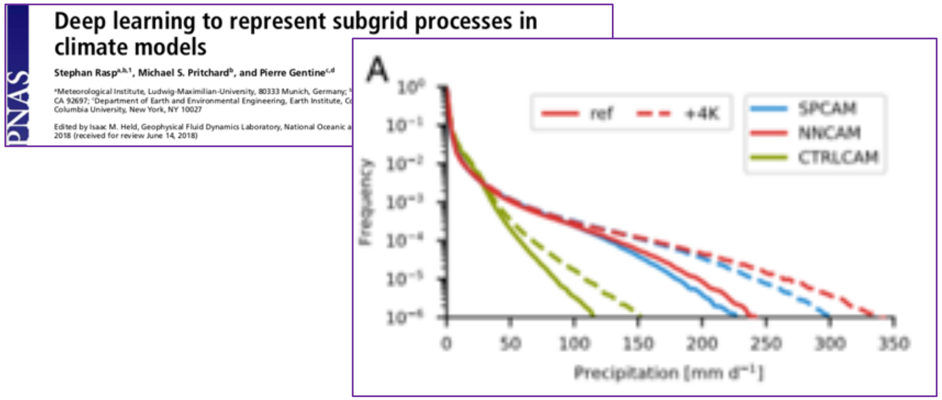
Smart observing systems

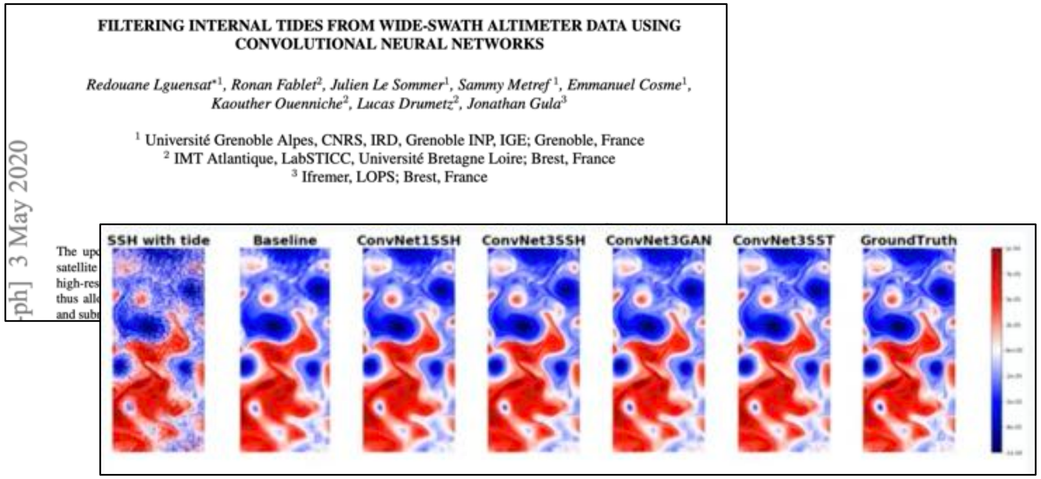
Direct applications of DL schemes to physics-related issues











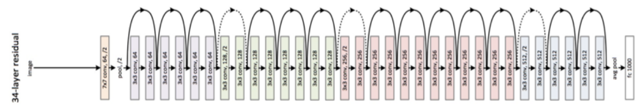
Bridging physics & AI: Expected breakthroughs

Representation learning

Physical model

$$\frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u$$

Data-driven representation



Making the most of AI and Physics Theory

- Model-Driven/Theory-Guided & Data-Constrained schemes
- Data-Driven & Physically-Sound schemes (eg, Ouala et al., 2019)

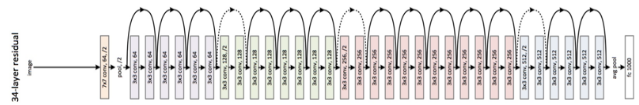
Bridging physics & AI: Expected breakthroughs

Physical model

$$\frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u$$

Representation learning

Data-driven representation

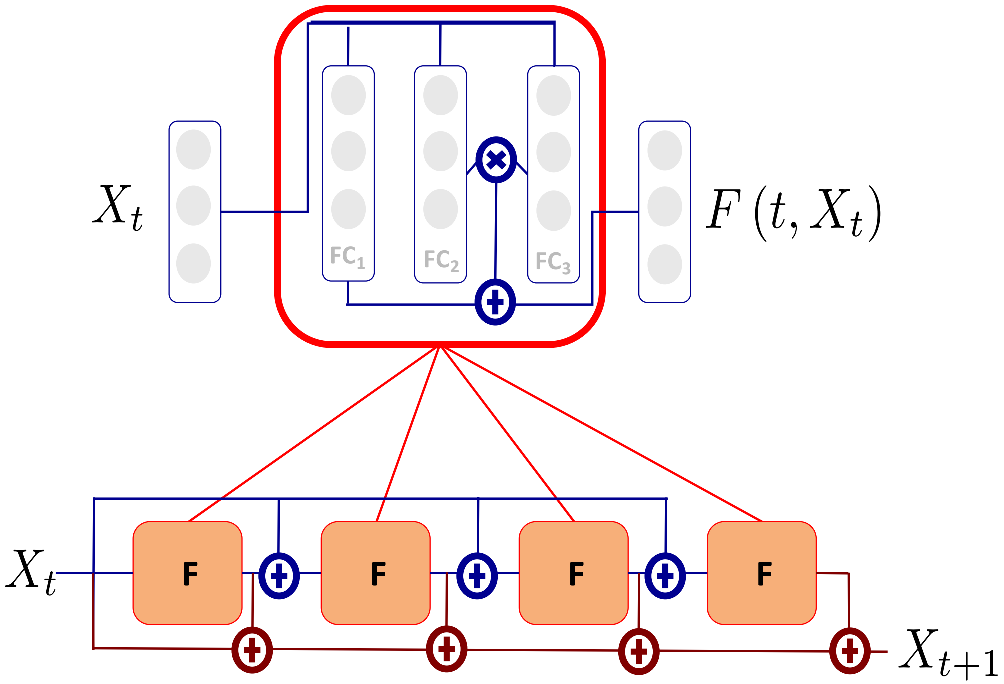


Making the most of AI and Physics Theory

- **Model-Driven/Theory-Guided & Data-Constrained schemes**
- Data-Driven & Physically-Sound schemes (eg, Ouala et al., 2019)

DL representations for ODEs/PDEs (Neural ODE)

An example: Residual RK4 Bilinear Network [Fablet et al., 2018]



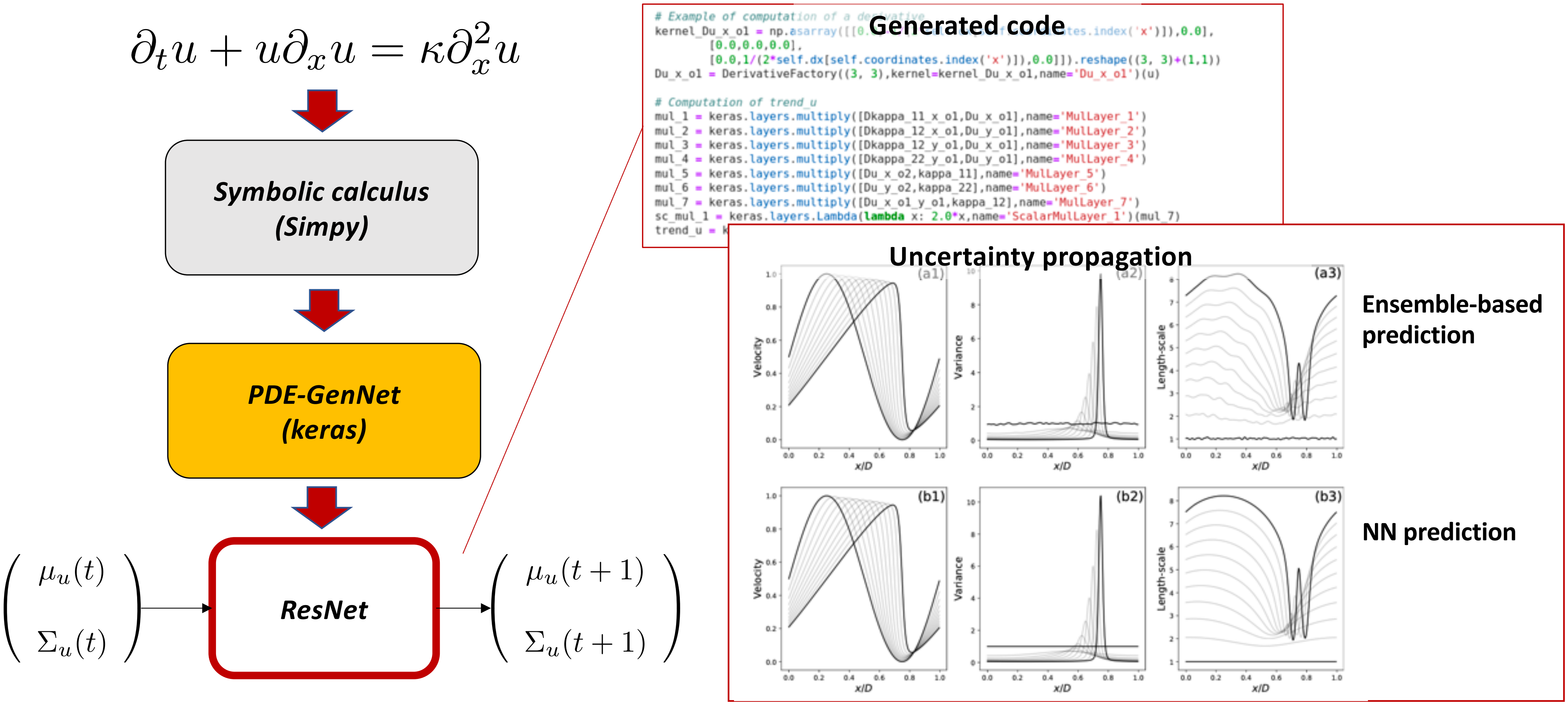
$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \sigma\left(y(t) - x(t)\right)$$
$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = x(t)\left(\rho - z(t)\right) - y(t)$$
$$\frac{\mathrm{d}z(t)}{\mathrm{d}t} = x(t)\,y(t) - \beta\,z(t)$$

Lorenz-63 equations

Noise-free training data

Forecasting time step	**t_0+h**	**t_0+4h**	**t_0+8h**
Analog forecasting	$<10^{-6}$	0.002	0.005
Sparse regression	$<10^{-6}$	0.002	0.006
MLP	$<10^{-6}$	0.018	0.044
Bi-NN(4)	***$<10^{-6}$***	***$<10^{-6}$***	***$<10^{-6}$***

NN Generator from Symbolic PDEs (Pannekoucke et al., 2020)



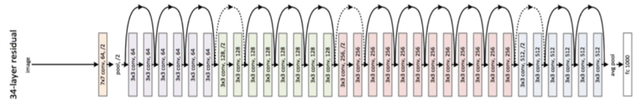
Bridging physics & AI: Expected breakthroughs

Physical model

$$\frac{\partial u}{\partial t} + \langle \nabla u, v \rangle = \kappa \Delta u$$

Representation learning

Data-driven representation



Making the most of AI and Physics Theory

- **Model-Driven/Theory-Guided & Data-Constrained schemes**
- **Data-Driven & Physically-Sound schemes (eg, Ouala et al., 2019)**

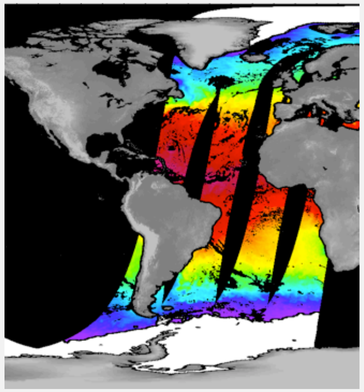
Dealing with real systems, including Irregularly-sampled, noisy and/or partially-observed systems ?

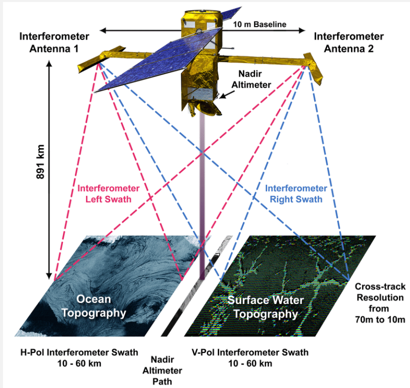
Lab-STICC

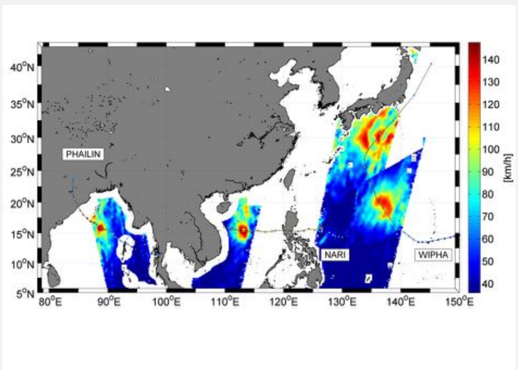
End-to-end learning from irregularly-sampled data

[Nguyen et al., 2019; Fablet et al., 2019]

Can we learn directly from observation data ?







Generic issue:

Joint identification and inversion

Dynamical model

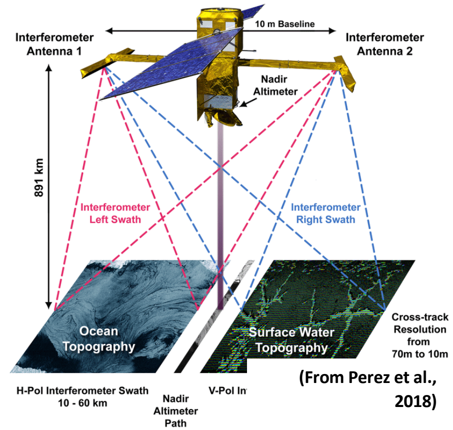
$$X_t \rightarrow \partial_t X = F(X, \xi, t, \theta) \rightarrow X_{t+1}$$

+

Observation model

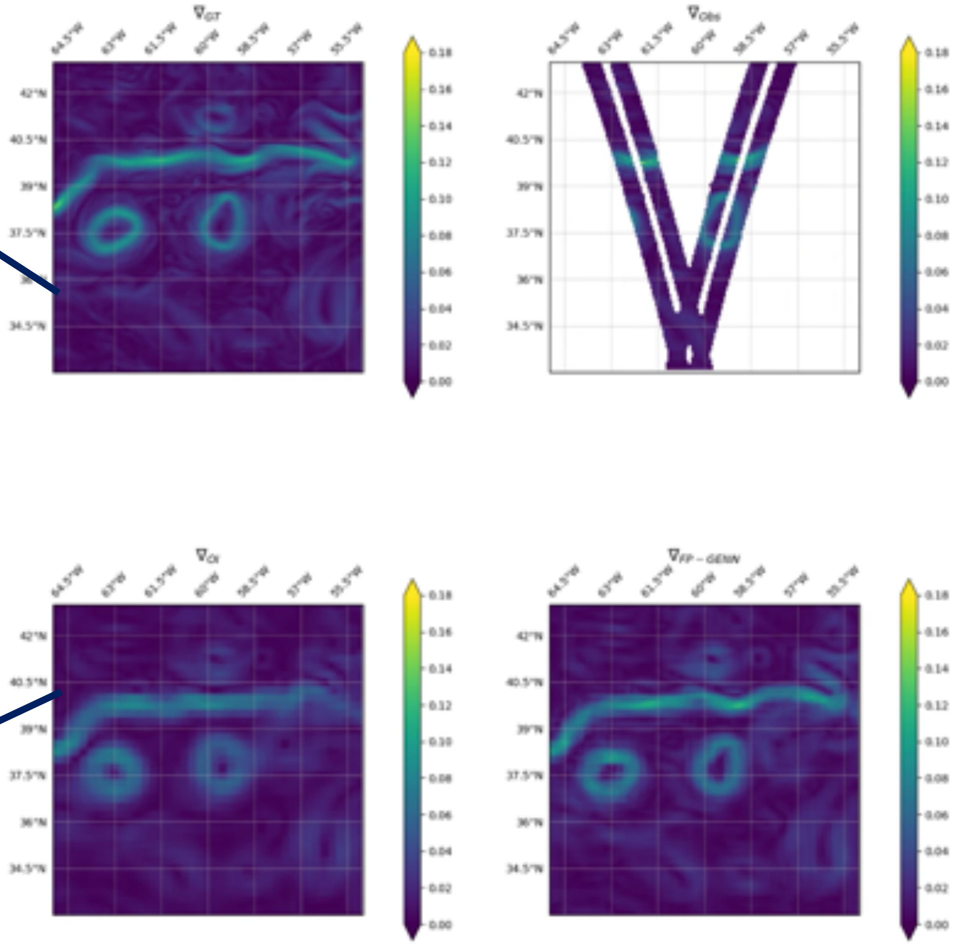
$$Y_t = H(X, \zeta, t, \phi)$$

An example for upcoming SWOT mission



(From Perez et al., 2018)

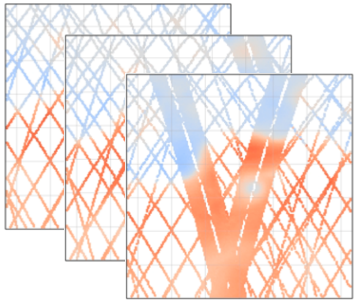
Groundtruth



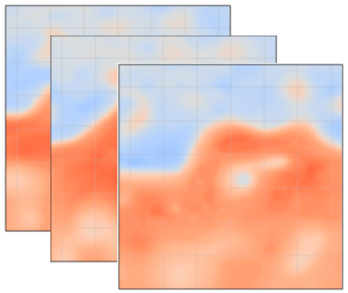
State-of-the-art operational processing

Proposed NN framework (Fablet et al., 2019)

End-to-end learning for inverse problems (Fablet etal., 2020)



Partial observations y



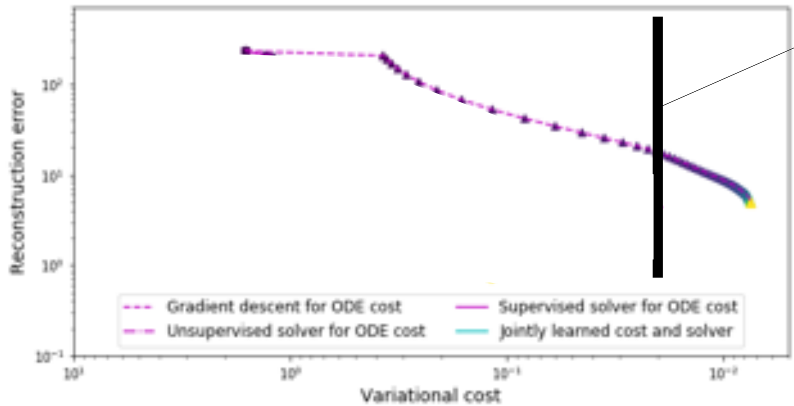
True states x

Model-driven schemes: $$\widehat{x} = \arg\min_x \underbrace{\lambda_1 \left\| x - y \right\|_\Omega^2 + \lambda_2 \Phi(x)}_{U_\Phi\left(x^{(k)}, y, \Omega\right)}$$

Gradient-based solver (adjoint/Euler-Lagrange method):

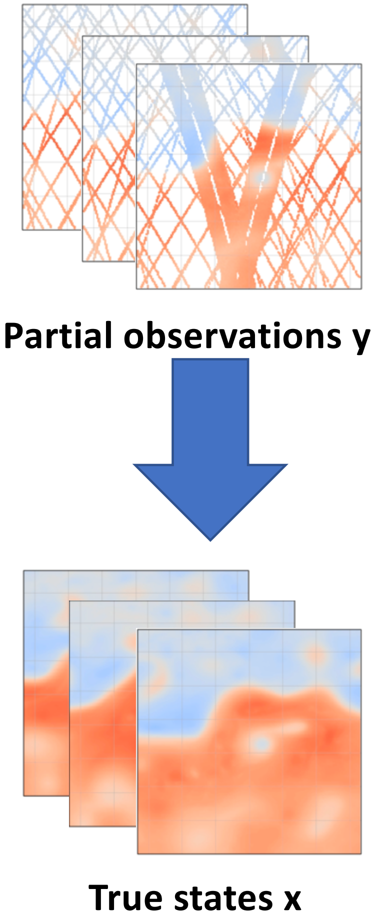
$$x^{(k+1)} = x^{(k)} - \alpha \nabla_x U_\Phi\left(x^{(k)}, y, \Omega\right)$$

No control on the reconstruction error $$x^{true} \neq \arg\min_x U_\Phi\left(x^{(k)}, y, \Omega\right)$$



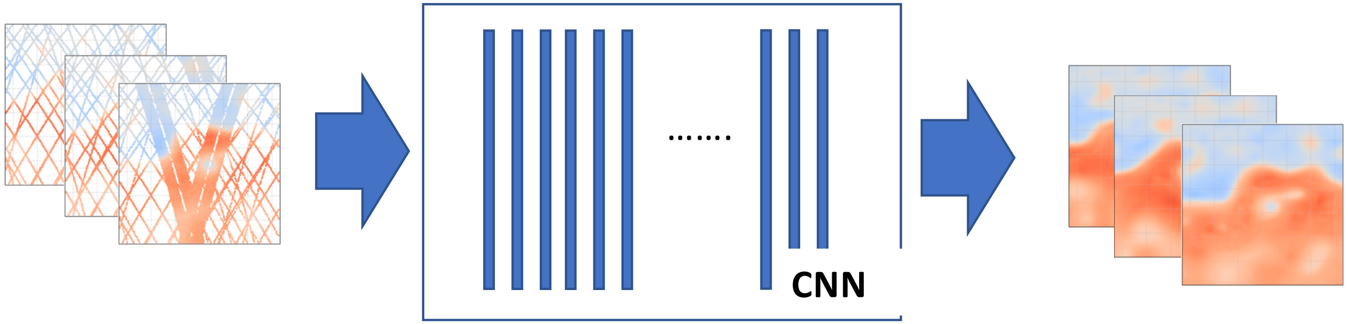
Variational cost for the true state

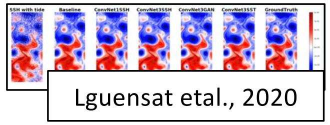
End-to-end learning for inverse problems (Fablet et al., 2020)



Model-driven schemes: $\widehat{x} = \arg\min_{x} \lambda_1 \|x - y\|^2_\Omega + \lambda_2 \Phi(x)$

Direct learning for inverse problems: $\widehat{x} = \Psi(y)$ $\quad y \rightarrow \boxed{\text{CNN}} \rightarrow x$

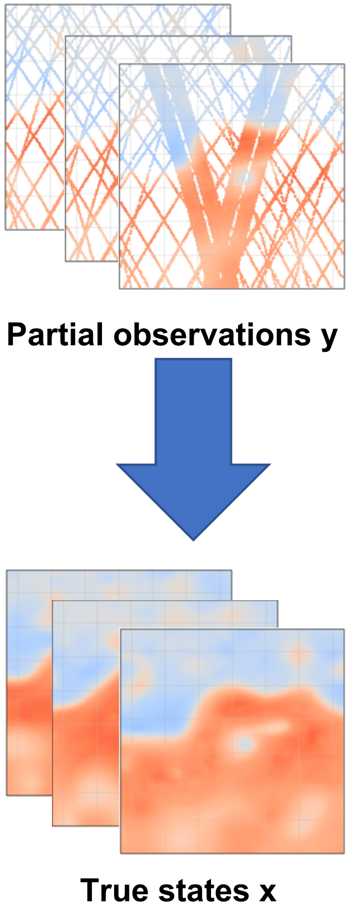




Examples of CNN architectures: Reaction-Diffusion architectures, ADMM-inspired architectures,…

Good performance but possibly weak interpretability/generaiization capacities of the solution byeond the training cases

End-to-end learning for inverse problems (Fablet et al., 2020)



Model-driven schemes: $\widehat{x} = \arg\min_x \lambda_1 \|x-y\|_\Omega^2 + \lambda_2 \Phi(x)$

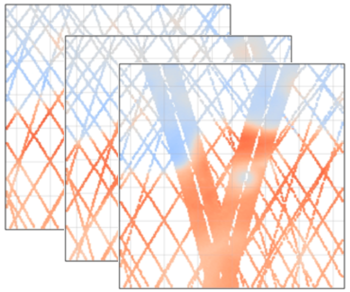
Direct learning for inverse problems: $\widehat{x} = \Psi(y)$ $\quad y \rightarrow$ CNN $\rightarrow x$

Proposed scheme: joint learning of the variational model and solver

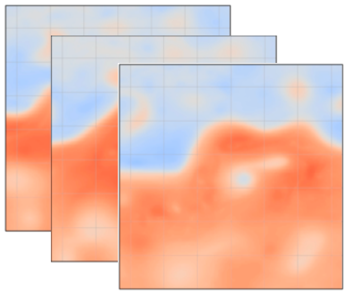
- **Theoretical bi-level optimization**

$$\arg\min_\Phi \sum_n \|x_n - \tilde{x}_n\|^2 \text{ s.t. } \tilde{x}_n = \arg\min_{x_n} U_\Phi(x_n, y_n, \Omega_n)$$

End-to-end learning for inverse problems (Fablet et al., 2020)



Partial observations y



True states x

Model-driven schemes: $\widehat{x} = \arg\min_{x} \lambda_1 \|x - y\|^2_\Omega + \lambda_2 \Phi(x)$

Direct learning for inverse problems: $\widehat{x} = \Psi(y)$ $\quad y \rightarrow \boxed{\text{CNN}} \rightarrow x$

Proposed scheme: joint learning of the variational model and solver

- **Theoretical bi-level optimization**

$$\arg\min_{\Phi} \sum_n \|x_n - \tilde{x}_n\|^2 \text{ s.t. } \tilde{x}_n = \arg\min_{x_n} U_\Phi(x_n, y_n, \Omega_n)$$

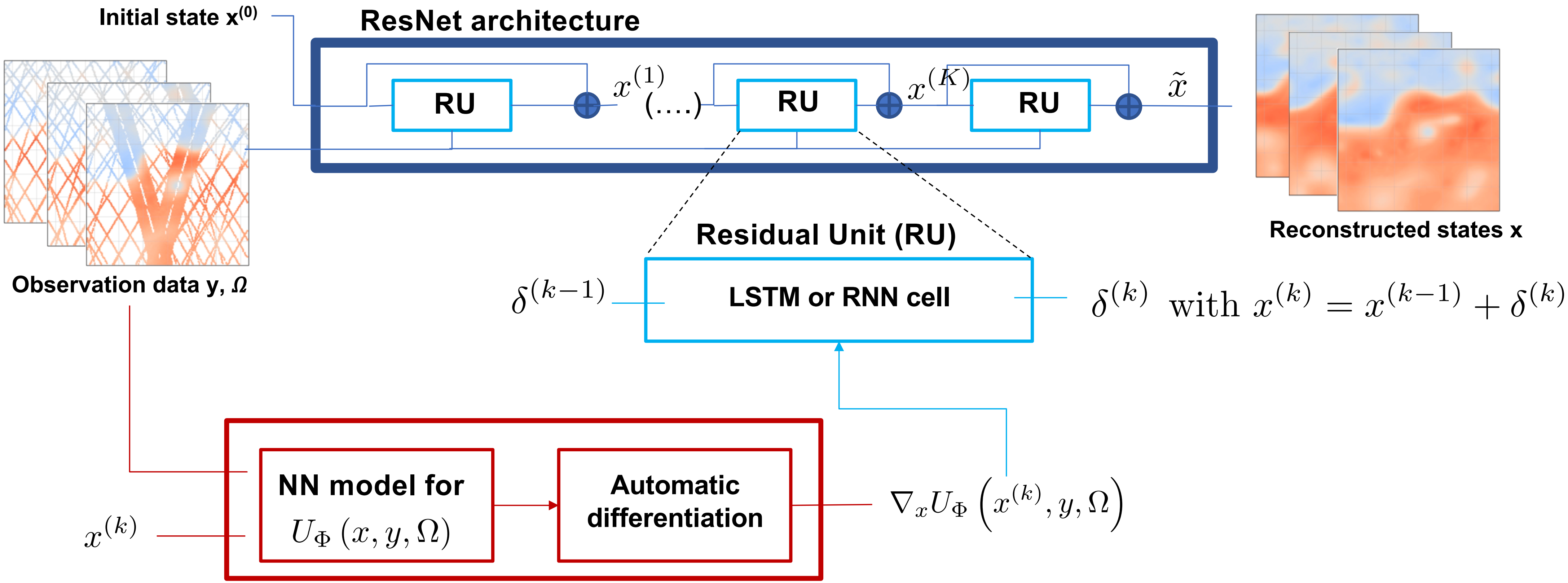
- **Restated with a gradient-based NN solver for inner minimization**

$$\arg\min_{\Phi,\Gamma} \sum_n \|x_n - \tilde{x}_n\|^2 \text{ s.t. } \tilde{x}_n = \Psi_{\Phi,\Gamma}(x_n^{(0)}, y_n, \Omega_n)$$

Iterative NN solver using automatic differentiation to compute gradient $\nabla_x U_\Phi \left(x^{(k)}, y, \Omega\right)$

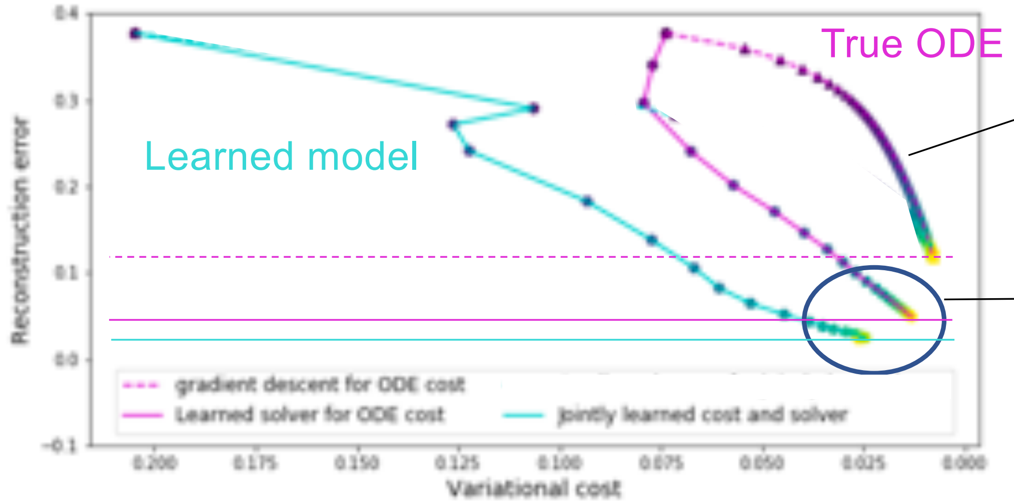
End-to-end learning for inverse problems (Fablet et al., 2020)

Proposed scheme: associated NN architecture



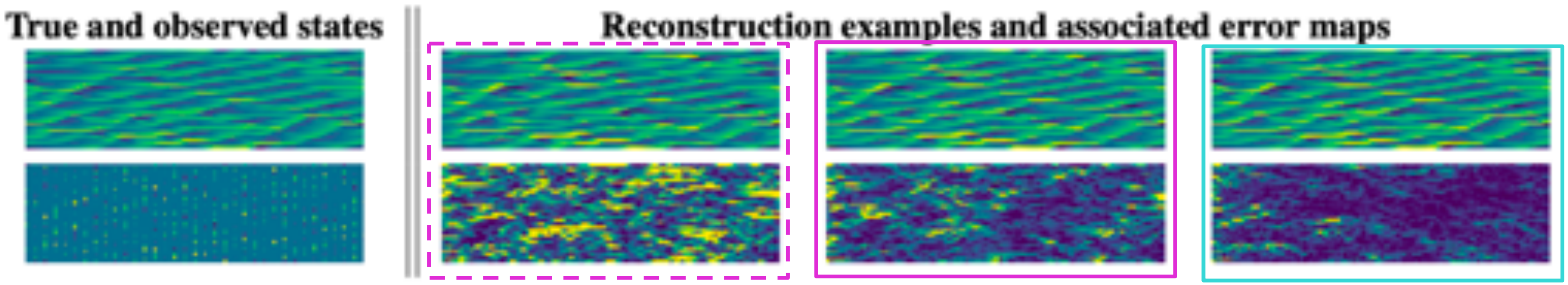
End-to-end learning for inverse problems (Fablet etal., 2020)

Illustration on Lorenz-96 dynamics (Bilinear ODE)



Non-supervised setting
(classic variational minimisation)

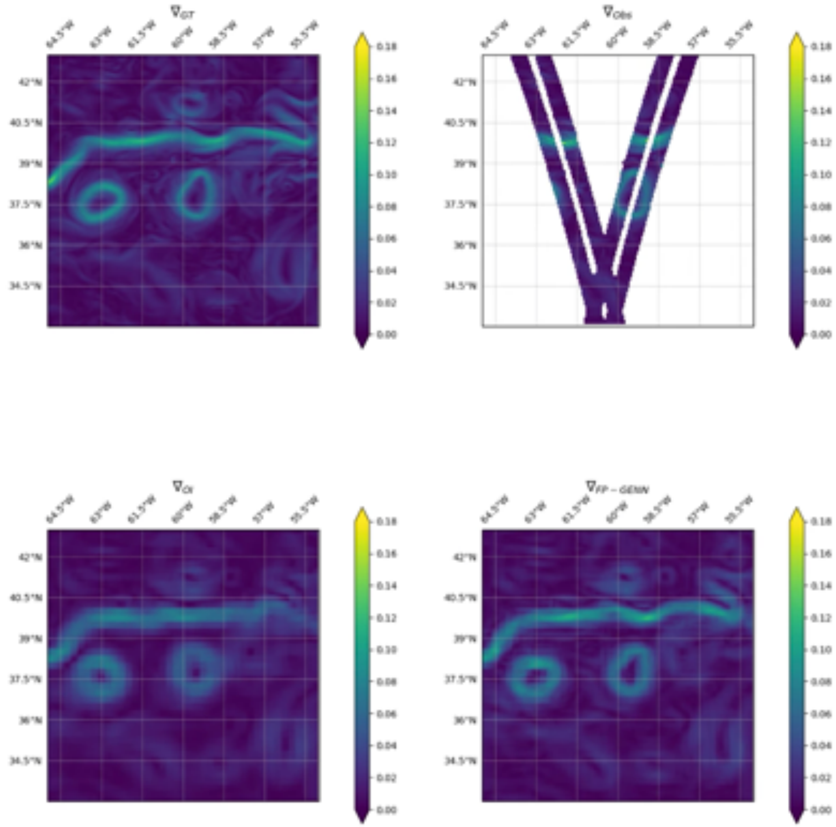
Supervised settings



End-to-end learning for inverse problems (Fablet etal., 2020)

Applications to the reconstruction of sea surface current from SWOT data

NB: preliminary results with a fixed-point Solver rather than a gradient-based solver



End-to-end learning for inverse problems (Fablet etal., 2020)

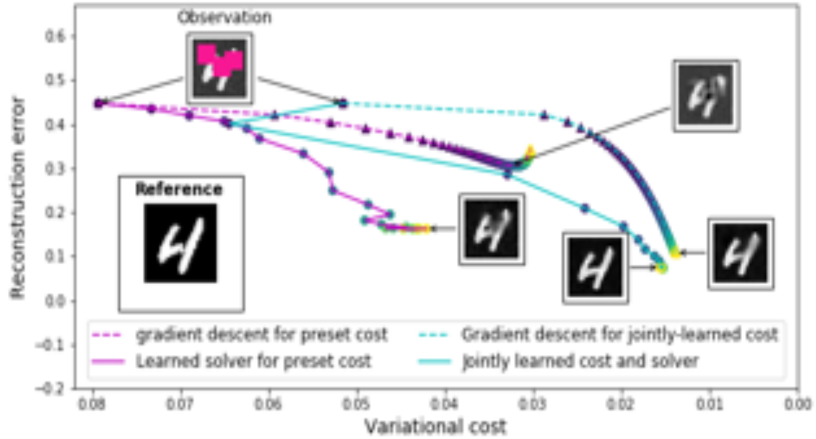
Key messages

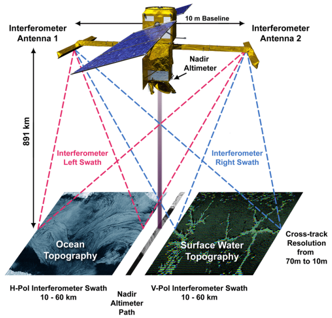
- **We can bridge DNN and variational models to solve inverse problems**

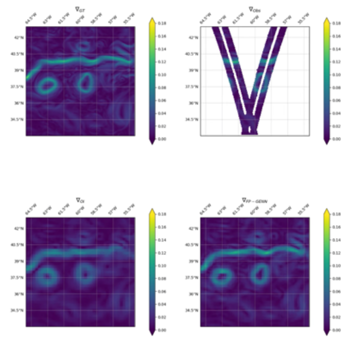
- **Learning both variational priors and solvers using groundtruthed (simulation) or observation-only data**

- **The best model may not be the TRUE one for inverse problems**

- **Generic formulation/architecture beyond space-time dynamics**





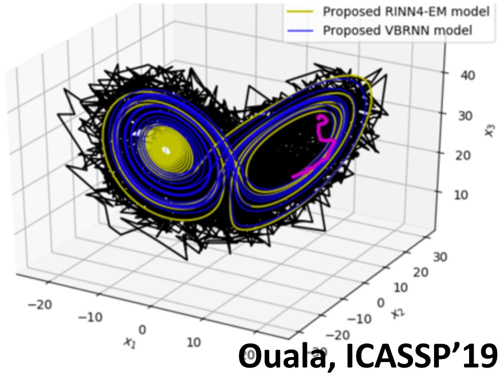


Preprint: https://arxiv.org/abs/2006.03653

Code: https://github.com/CIA-Oceanix

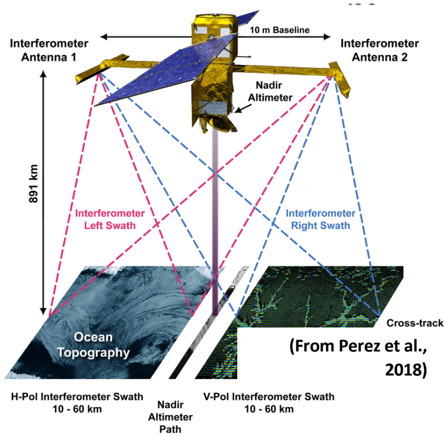
End-to-end learning from real observation data ?

Scarce time sampling

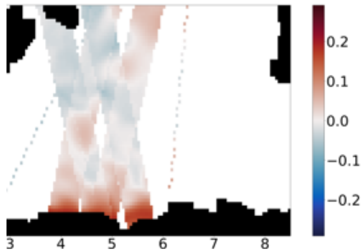


Ouala, ICASSP'19

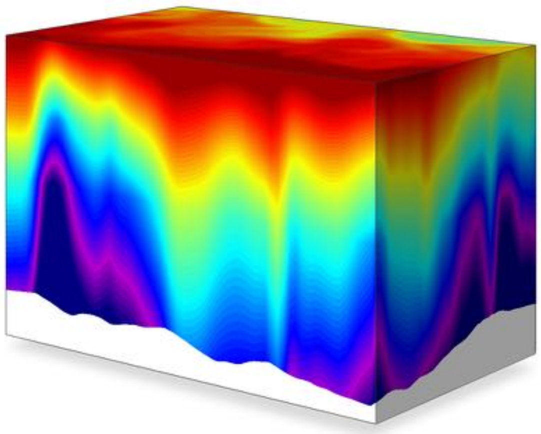
Noisy and irregular sampling



(From Perez et al., 2018)



Nguyen, ICASSP'20



Partially-observed system

Ouala, preprint 2019

Summary

- ***NNs as numerical schemes for ODE/PDE/energy-based representations of geophysical flows***

- ***Embedding geophysical priors in NN representations*** (e.g., Lguensat et al., 2019; Ouala et al., 2019)

- ***End-to-end architecture for jointly learning a representation (**eg, ODE**) and a solver*** (e.g., Fablet et al., 2020)

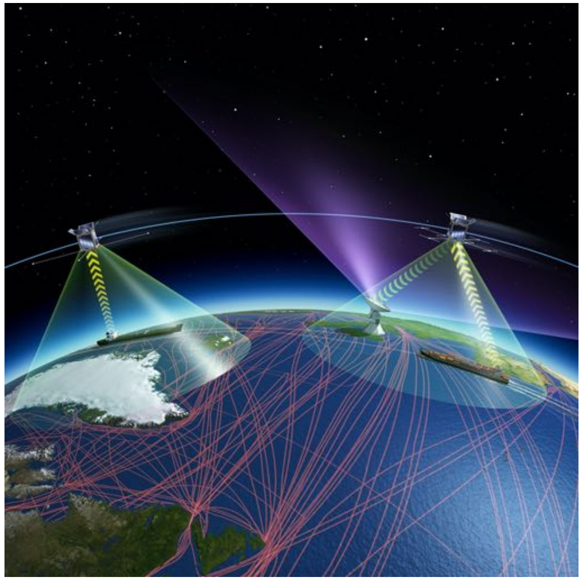
- ***Towards stochastic representations embedded in NN architectures*** (e.g., Pannekoucke et al., 2020, Nguyen etal., 2020)

Beyond Ocean Dynamics

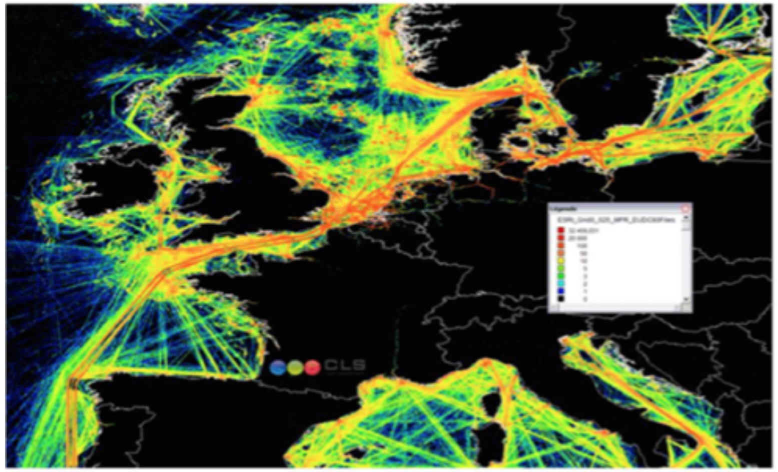
Learning stochastic hidden dynamics

Learning stochastic hidden dynamics [Nguyen et al., 2018]

The example of AIS Vessel trajectory data

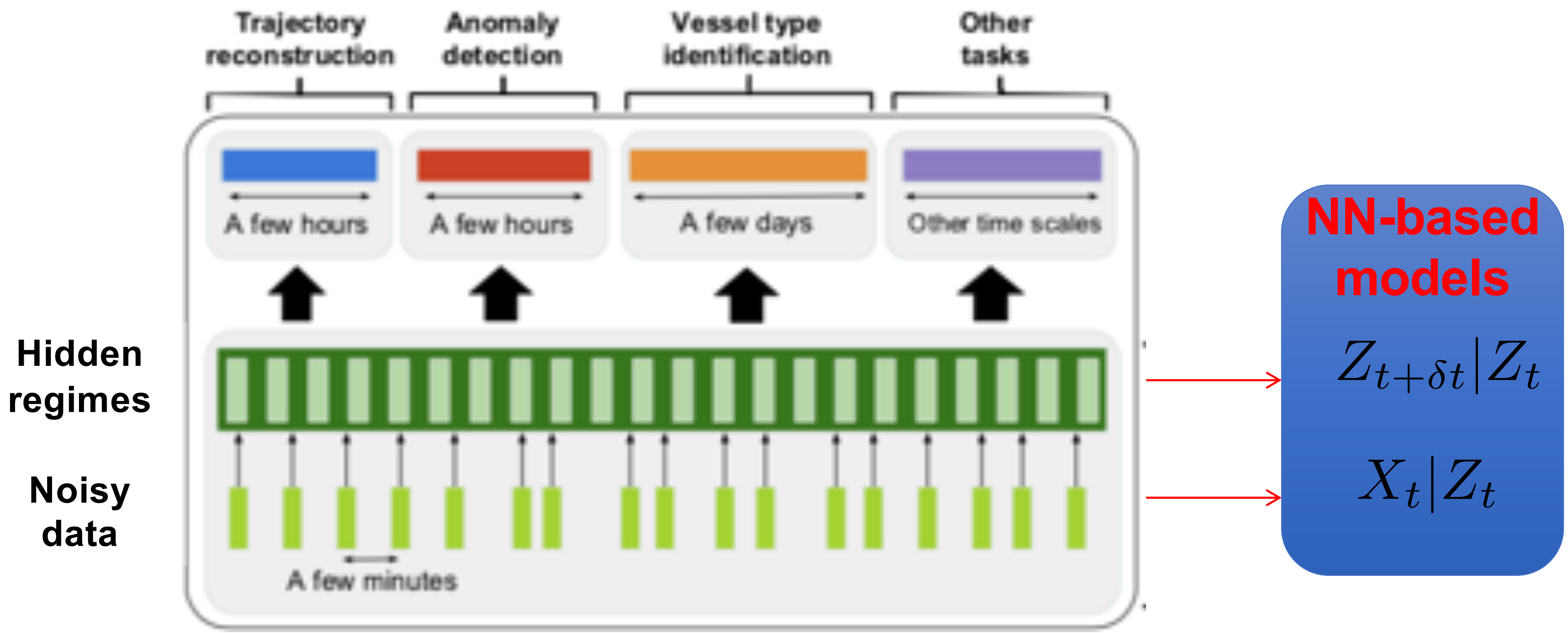


- Millions of AIS positions daily
- Noisy data: irregular sampling, corrupted data



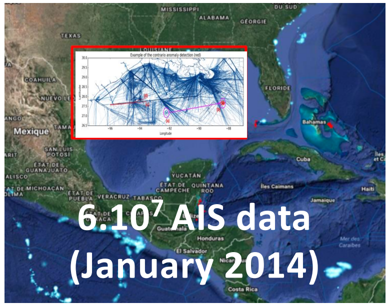
How can we learn from AIS data streams ?

Learning stochastic hidden dynamics [Nguyen et al., 2018]



Model training from noisy AIS streams using variational Bayesian approximation

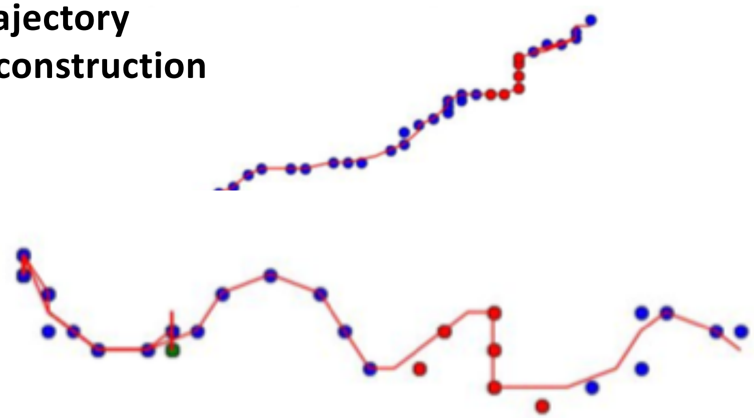
Learning stochastic hidden dynamics [Nguyen et al., 2018]



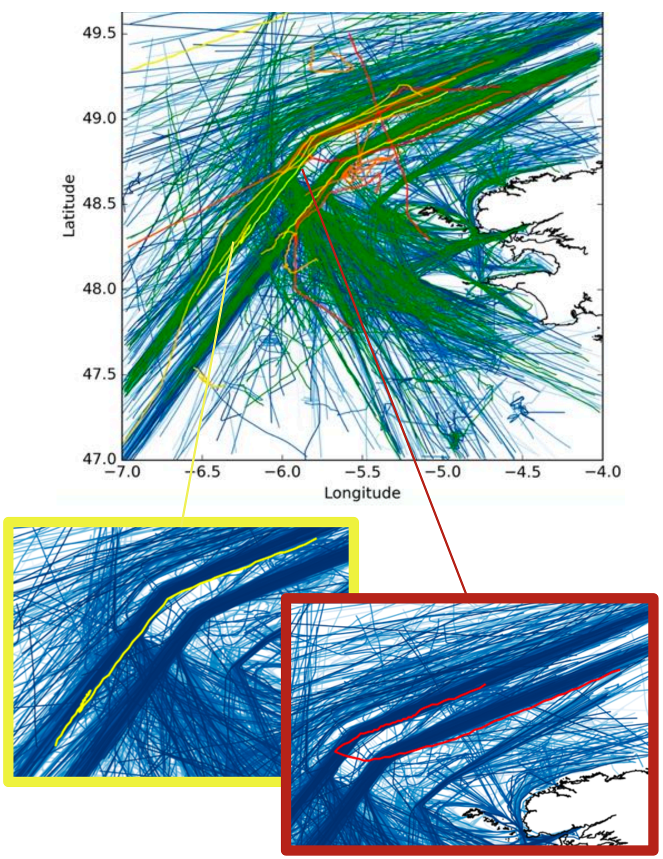
Vessel type recognition

~88% of correct recognition

Trajectory reconstruction



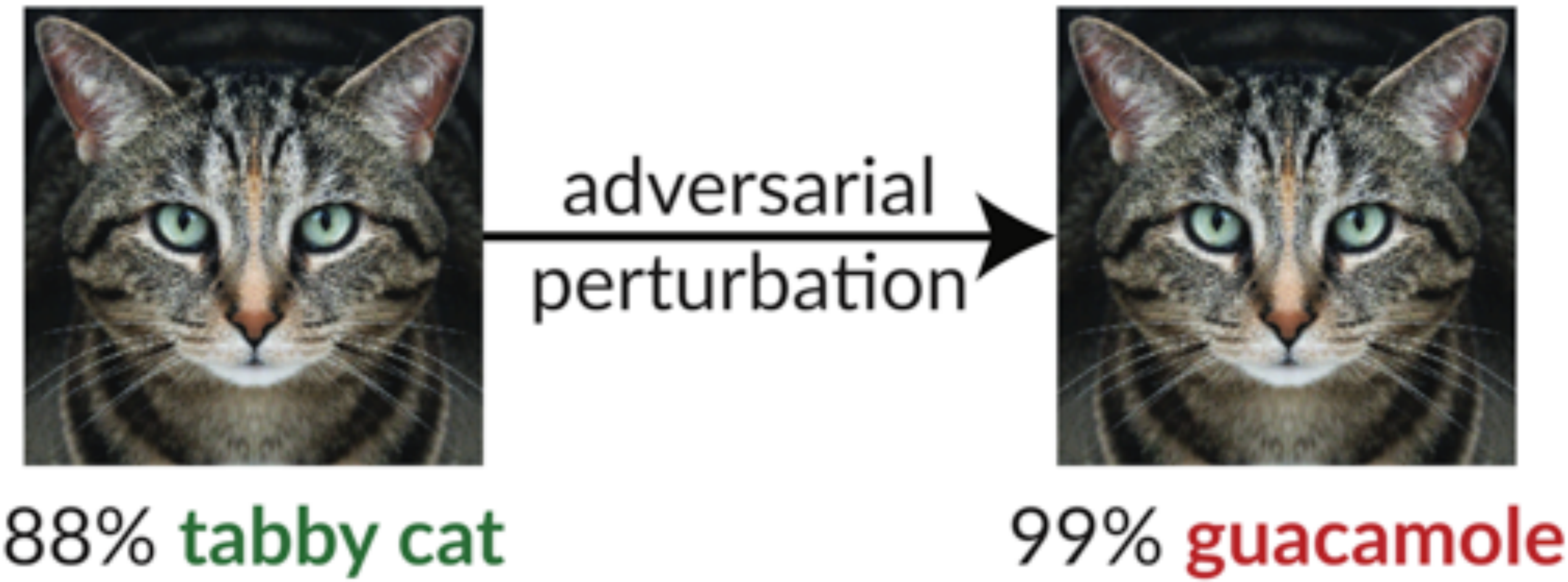
Abnormal behaviour detection



Beyond Ocean Dynamics

Dynamical System Theory for Deep Learning

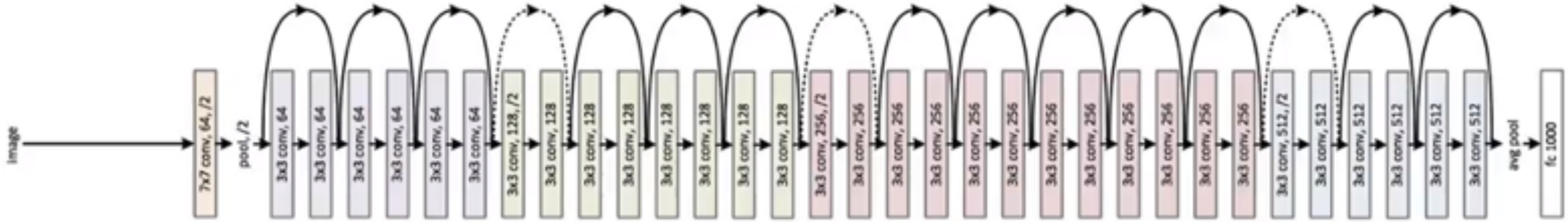
Understanding DL models ?



Szegedy et al., 2015

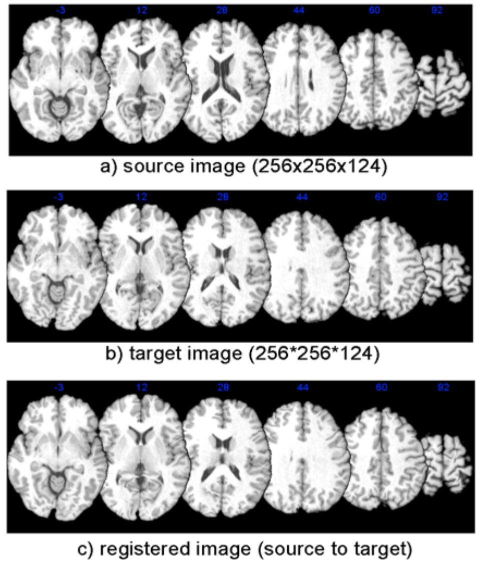
Understanding ResNets [Rousseau et al., 2019]

ResNet [He et al., 2015] regarded as space registration machines



- **Image registration examples**

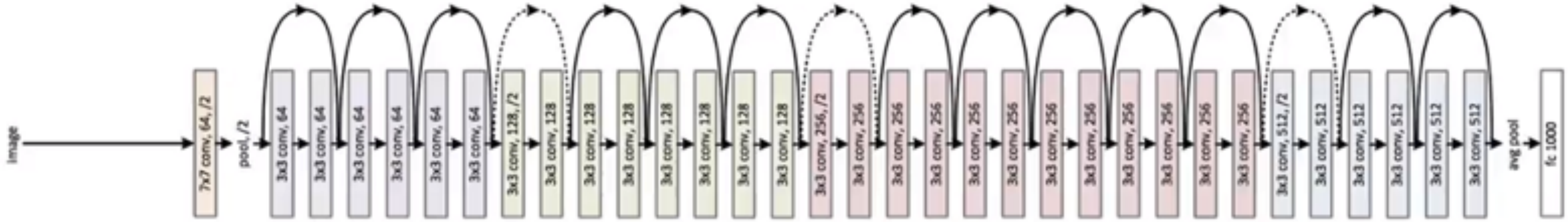
[Matlab tutorial]



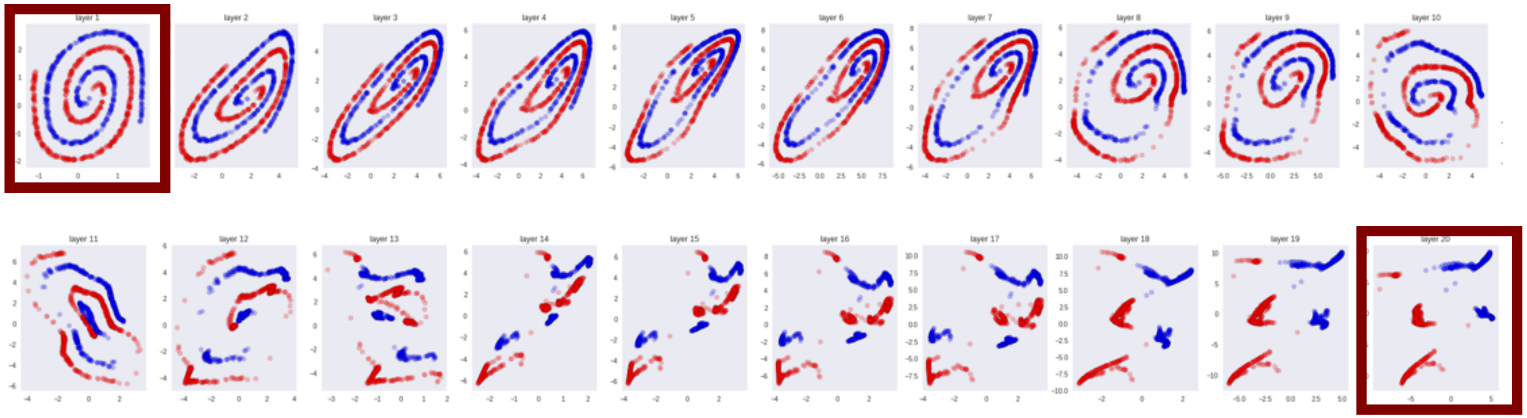
[Dramms tutorial]

Understanding ResNets [Rousseau et al., 2019]

ResNet [He et al., 2015] regarded as space registration machines



Original feature space



Registered space to make feasible a linear separation between classes

AI Chair OceaniX 2020-2024

Physics-informed AI
for Observation-Driven Ocean AnalytiX

PI: R. Fablet, Prof. IMT Atlantique, Brest

Internship, PhD and postdoc opportunities

(https://rfablet.github.io/)

Thank you.

Joint work with B. Chapron, F. Collard, L. Drumetz, J. Le Sommer, R. Lguensat, D. Nguyen, S. Ouala, A. Pascual, F. Rousseau, P. Tandeo, J. Verron, O. Pannekoucke, …

More:

- **Webpage:** https://rfablet.github.io/
- **Preprints**: https://www.researchgate.net/profile/Ronan_Fablet