

Bridging Physics and Learning: application to ocean dynamics

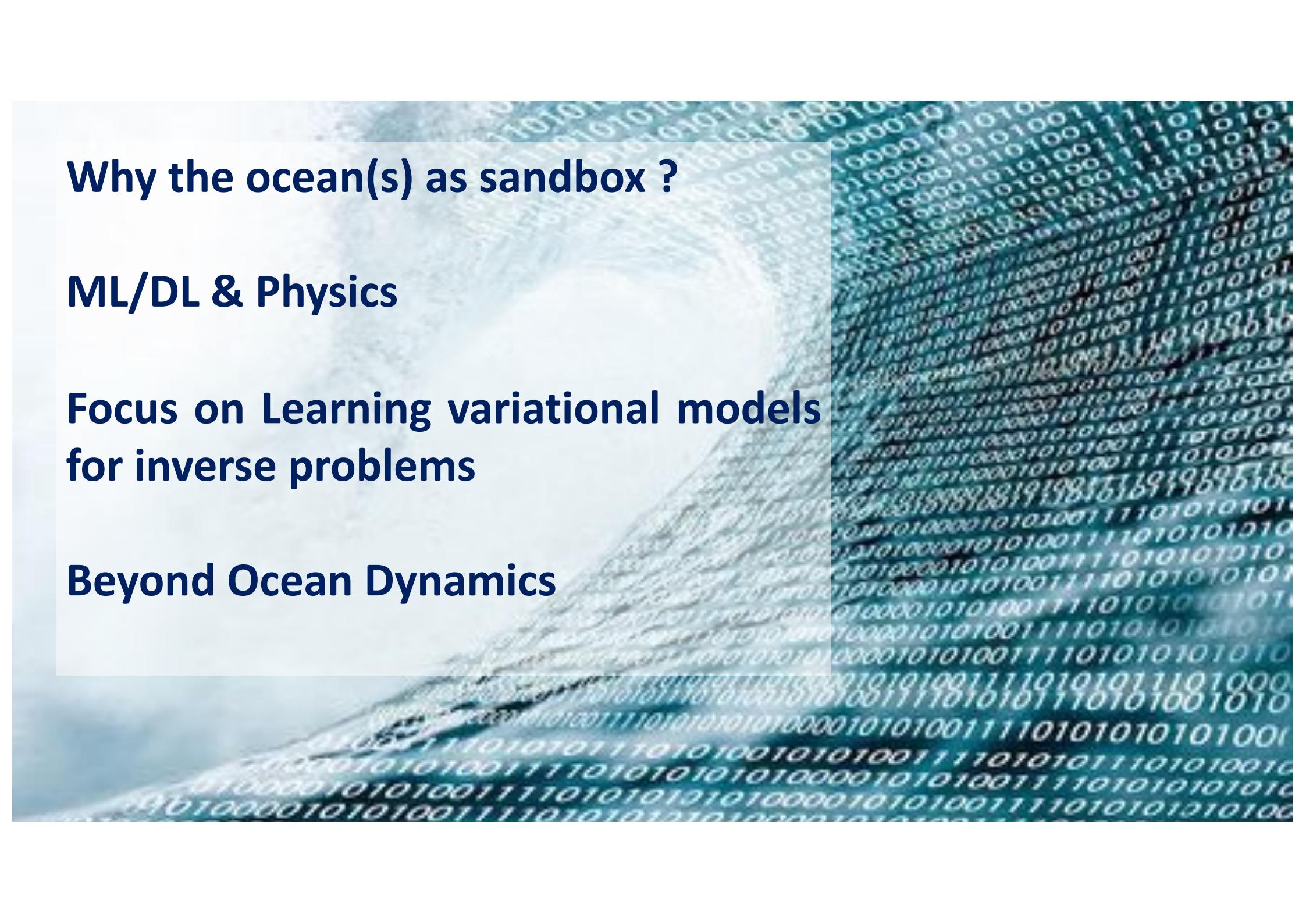
R. Fablet et al.

ronan.fablet@imt-atlantique.fr

web: rfablet.github.io

Webinar IMT Data & AI, July 2020



The background of the slide features a photograph of ocean waves crashing, with a dense grid of binary digits (0s and 1s) overlaid across the entire image.

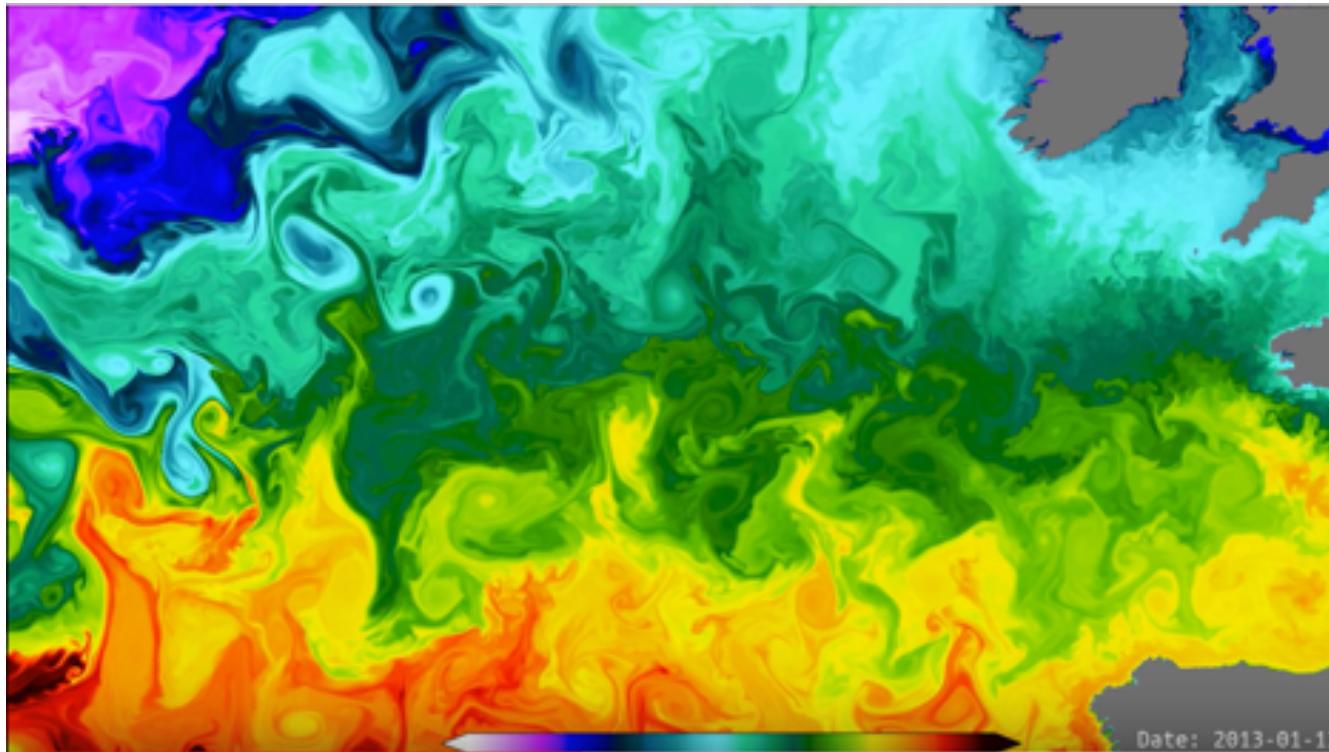
Why the ocean(s) as sandbox ?

ML/DL & Physics

**Focus on Learning variational models
for inverse problems**

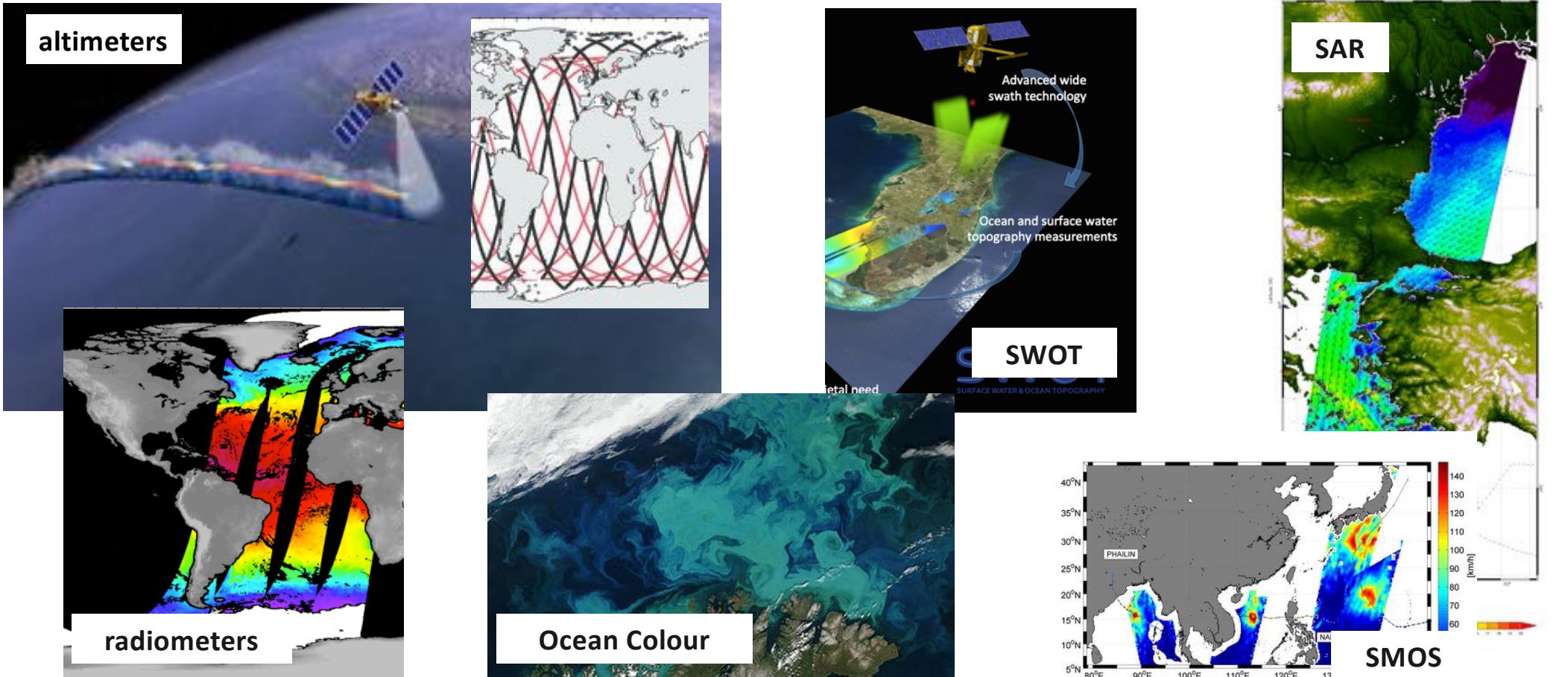
Beyond Ocean Dynamics

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

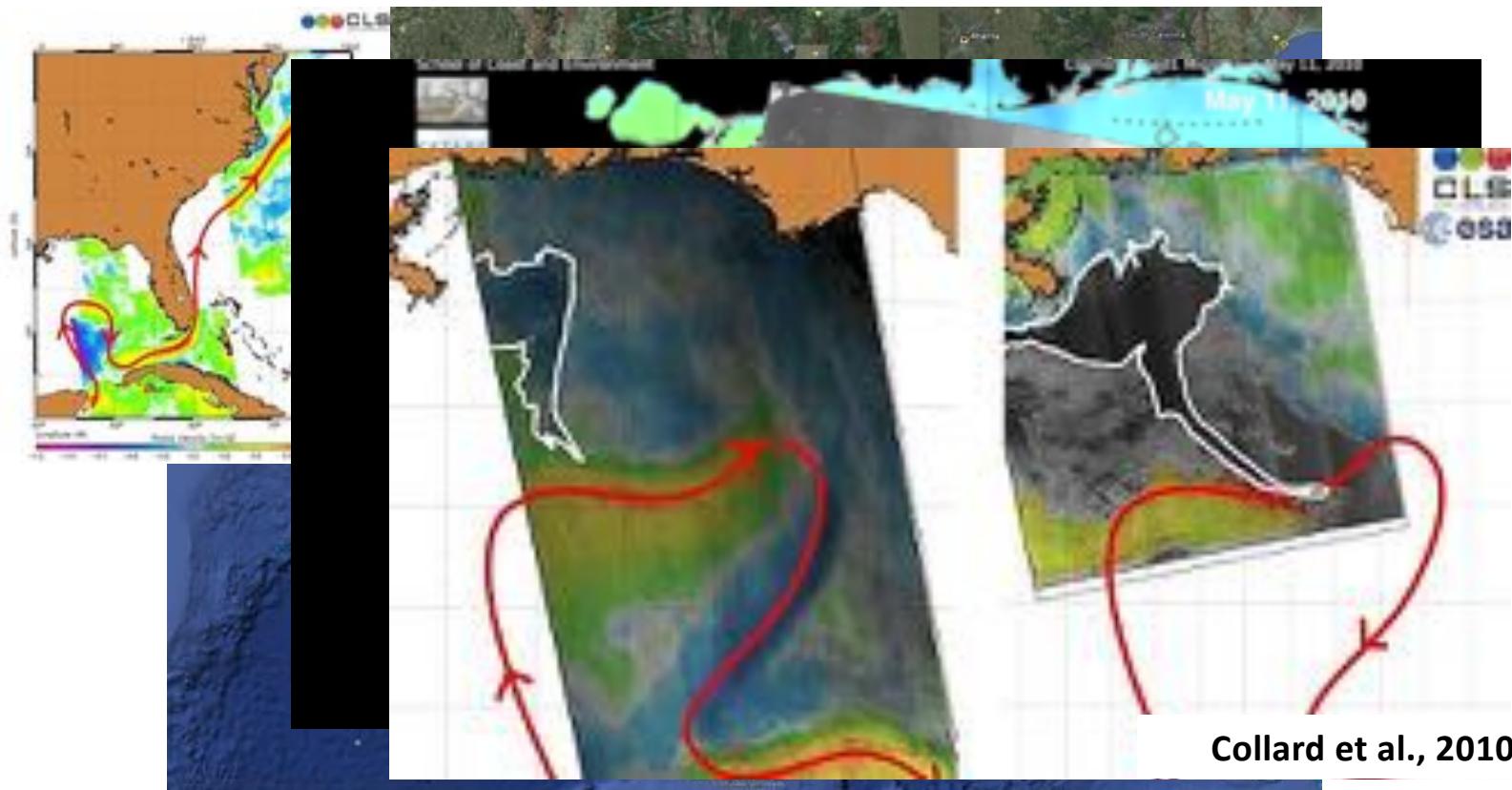


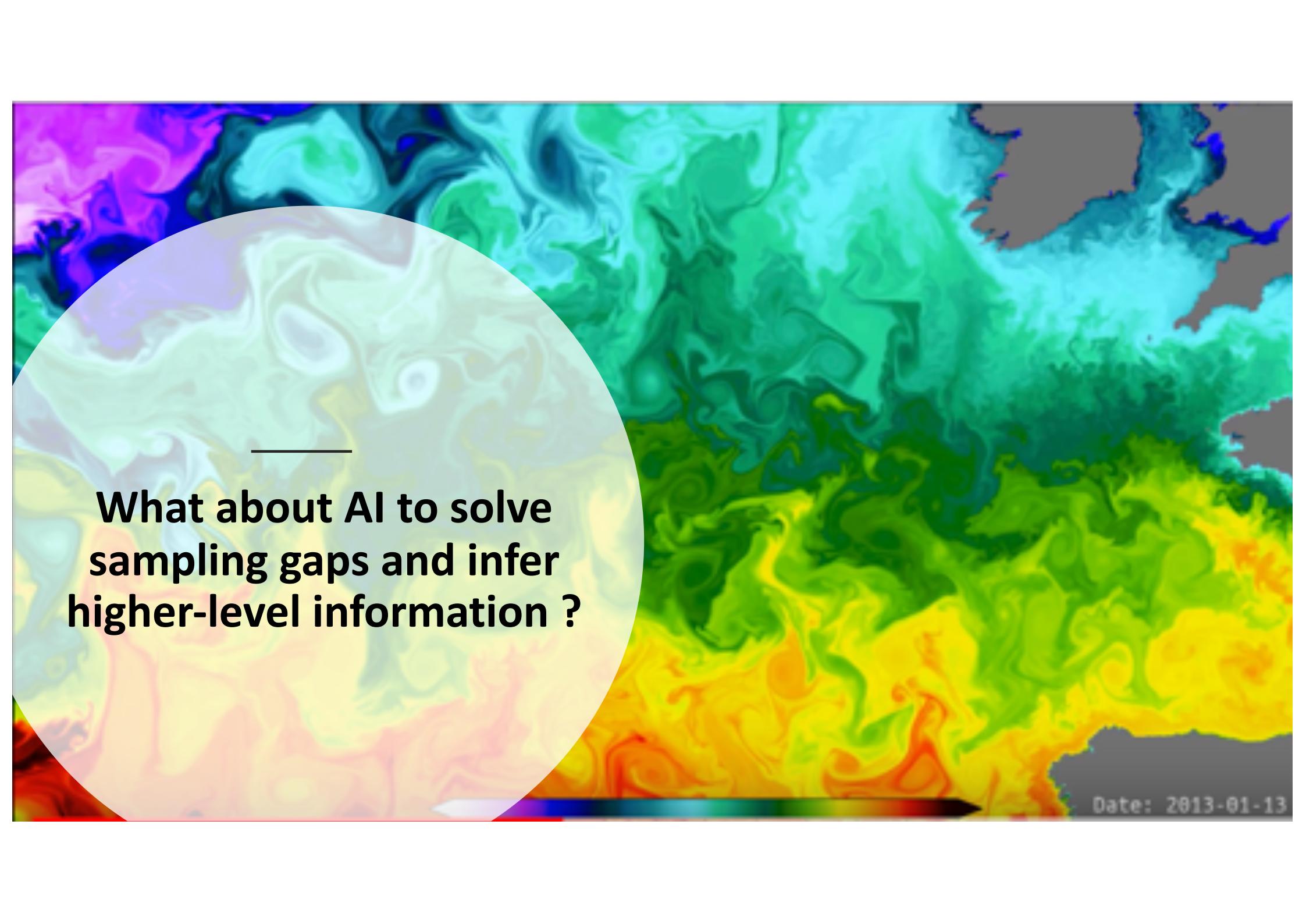
General question: how could data-driven/learning-based tools contribute to solving sampling gaps and higher-level information ?

Illustration of satellite-derived sea surface observations



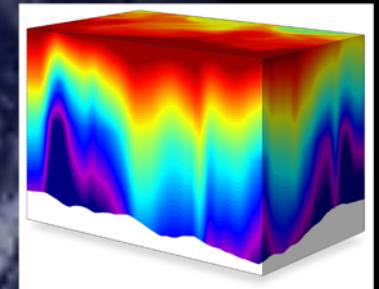
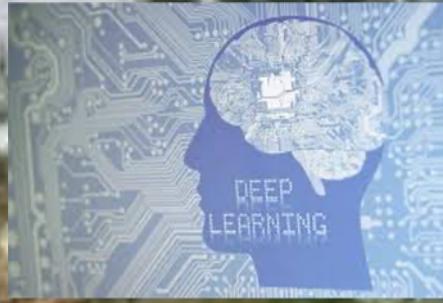
Deepwater horizon [2010]





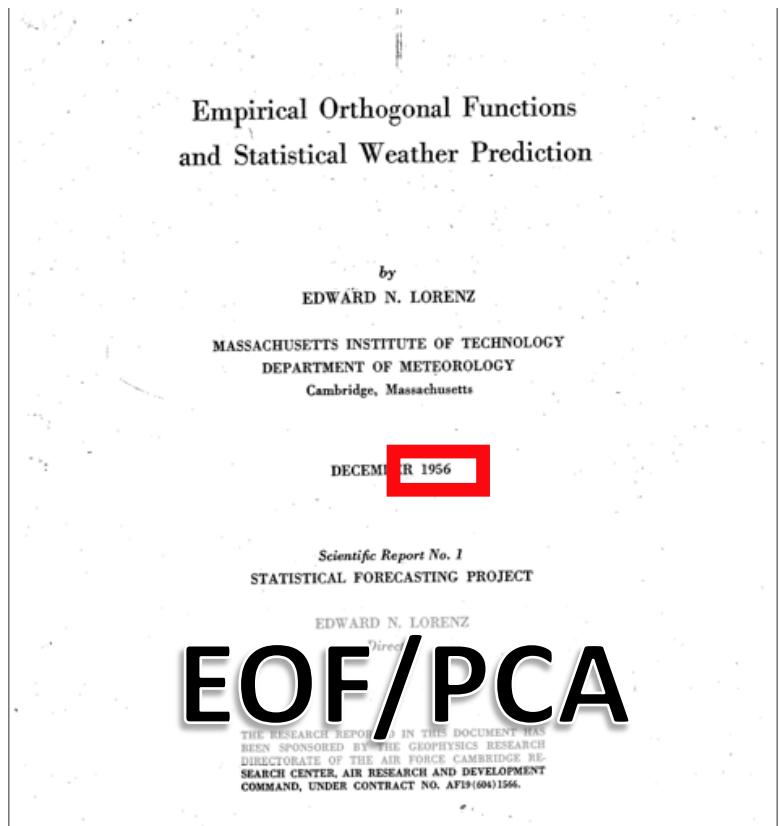
**What about AI to solve
sampling gaps and infer
higher-level information ?**

Date: 2013-01-13



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

Learning & Geoscience: nothing new ?



Deterministic Nonperiodic Flow¹

EDWARD N. LORENZ

Massachusetts Institute of Technology

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

Abstract

Finitesimal 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 essentially 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.

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 histories.

These modes of familiar rotation (cf. 1959) are cylindrical waves about the axis, and is heated near the rim and cooled at its center in a steady symmetrical fashion. Under certain conditions the resulting flow is as symmetric and steady as the heating which gives rise to it. Under different conditions a system of regularly spaced waves develops, and progresses at a uniform rate without changing its shape. Under still different conditions the flow forms, and moves in a nonperiodic manner.

Lack of periodicity is a characteristic of turbulence, and is one of the distinguishing features of turbulent flow. Because instantaneous turbulent flow patterns are so irregular, attention is often confined to the statistics of turbulence, which, in contrast to the details of turbulence, often behave in a regular well-organized manner. The short-range weather forecaster, however, is forced to pay attention to the details of the small-scale turbulent flow patterns which occur in the

¹The research reported in this work has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, under Contract No. AF 33(657)-4950.

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

In this study we shall work with systems of deterministic equations which are idealizations of hydrodynamical systems. We shall be interested principally in nonperiodic solutions, i.e., solutions which never repeat their past history exactly, and where approximate repetitions are of finite duration. The flow shall be as irregular as possible, and the behavior of the solutions, as far as possible, shall reflect the real behavior associated with actual physical systems.

Ogs /

rest-

It is sometimes possible to obtain particular solutions of these equations analytically, especially when the solutions are periodic or invariant with time, and, indeed, much work has been devoted to obtaining such solutions by one scheme or another. Ordinarily, however, aperiodic solutions cannot readily be determined, and it is often necessary to proceed by numerical methods, in which case it may perhaps be the most convenient to start with a regular grid of points, the coordinates of which are the variables in series of orthogonal functions. The governing laws, then become a finite set of ordinary differential

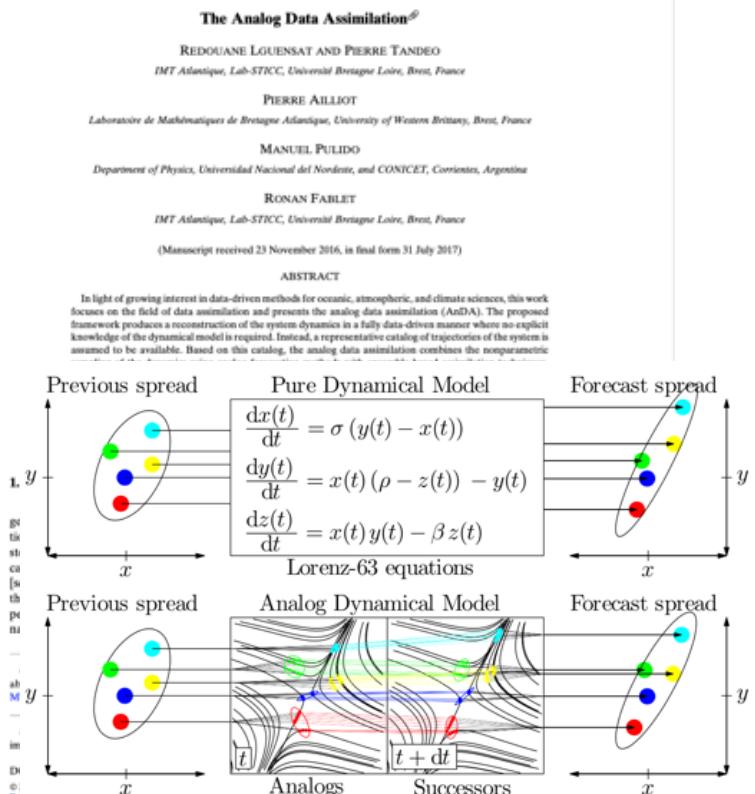
Analogs / Nearest- neighbors

Learning & Geoscience: Data-driven approaches for data assimilation

OCTOBER 2017

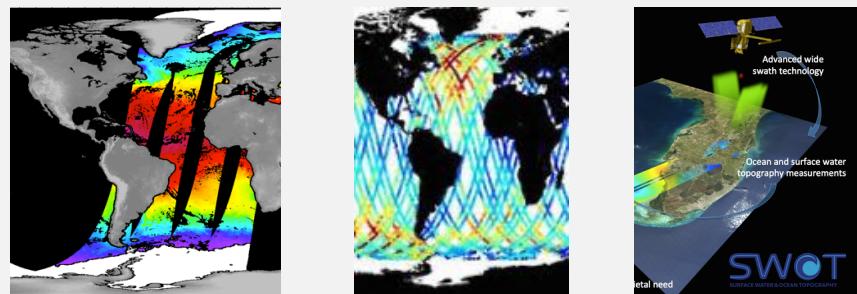
LGUENSAT ET AL.

4093



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

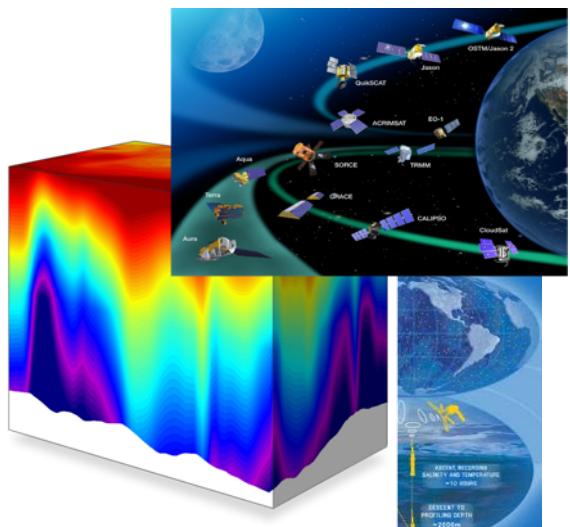
- Bridging model-driven and data-driven paradigms
- Learning data-driven representations from real observation data



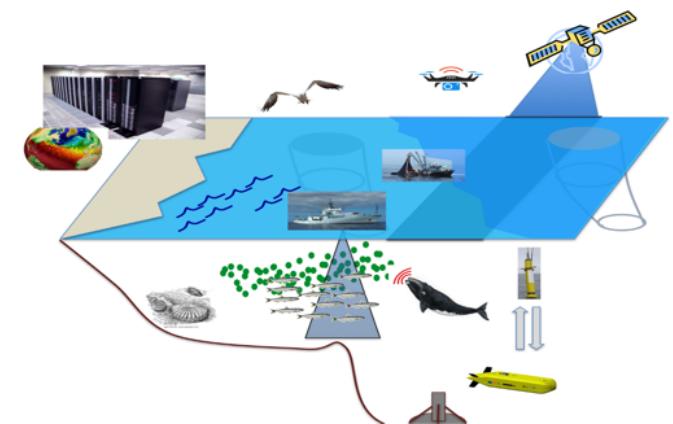
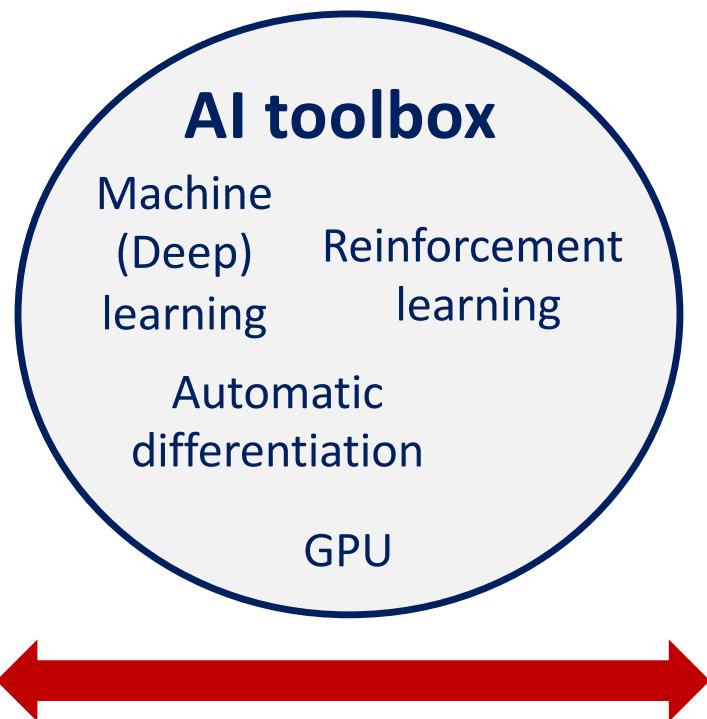
Bridging ML/DL paradigms and Physics ?



Bridging Physics & AI: Expected breakthroughs

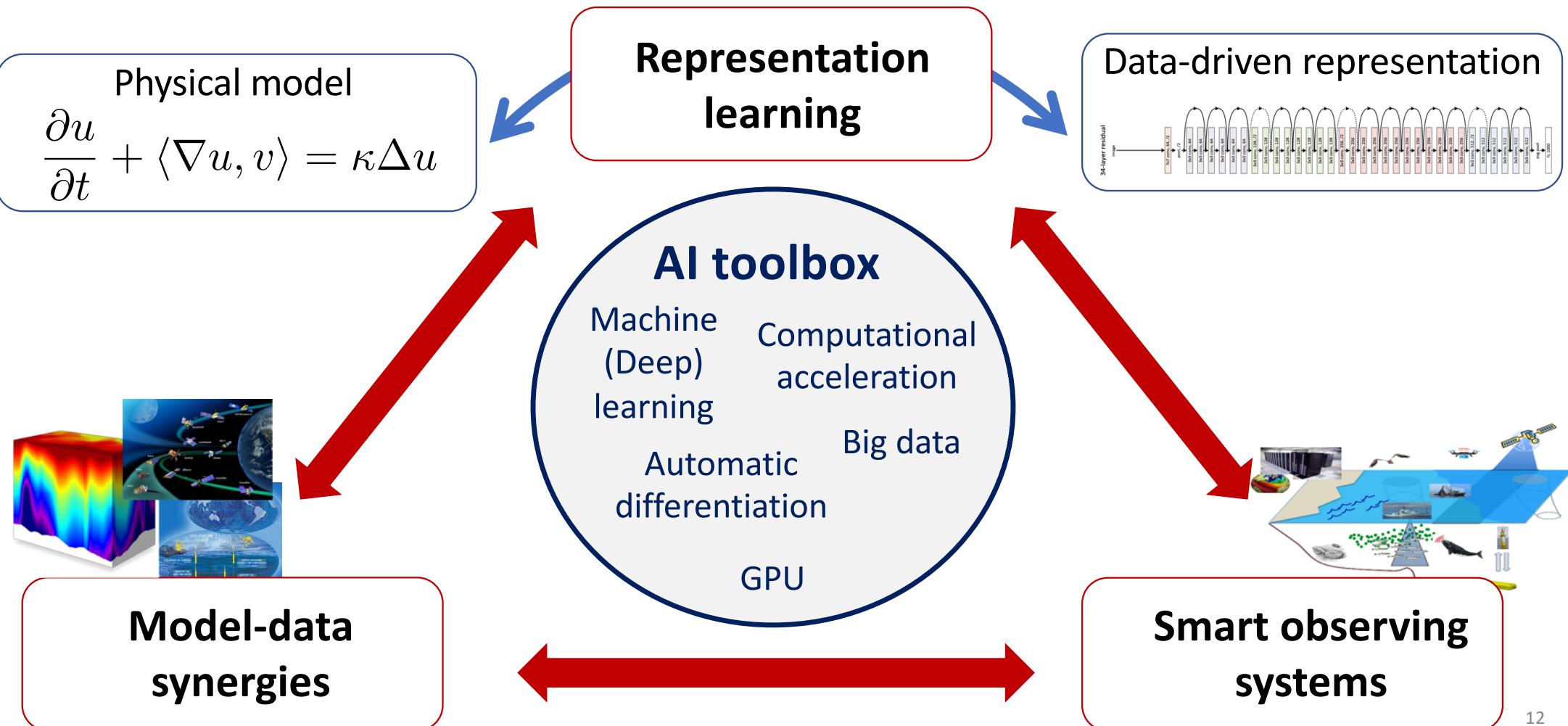


Model-data
synergies

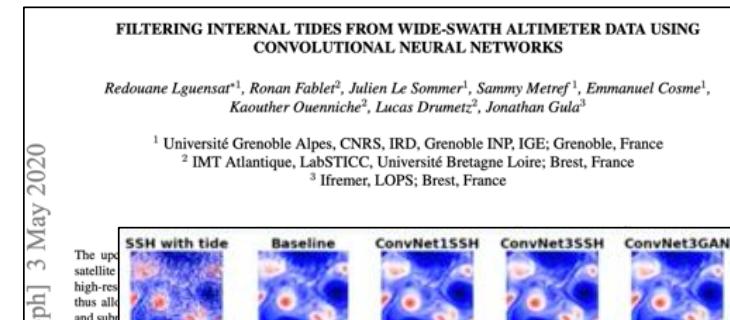
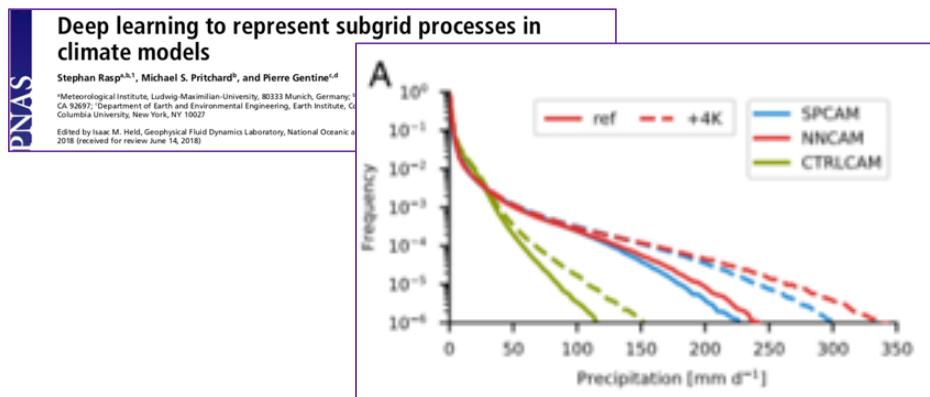
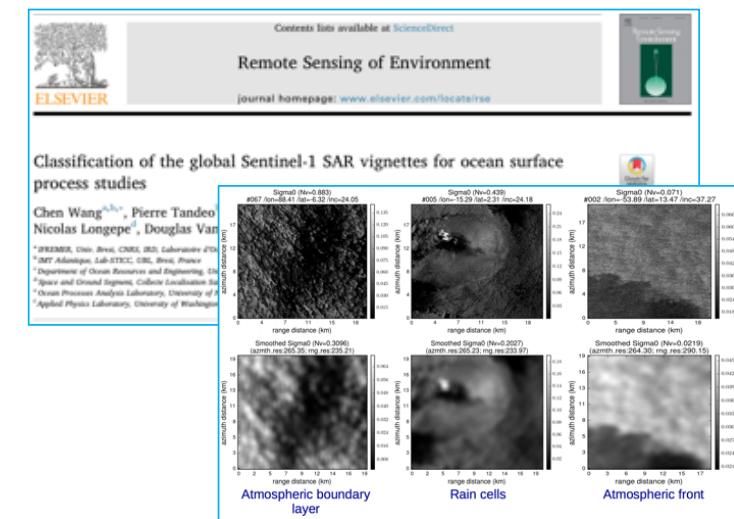
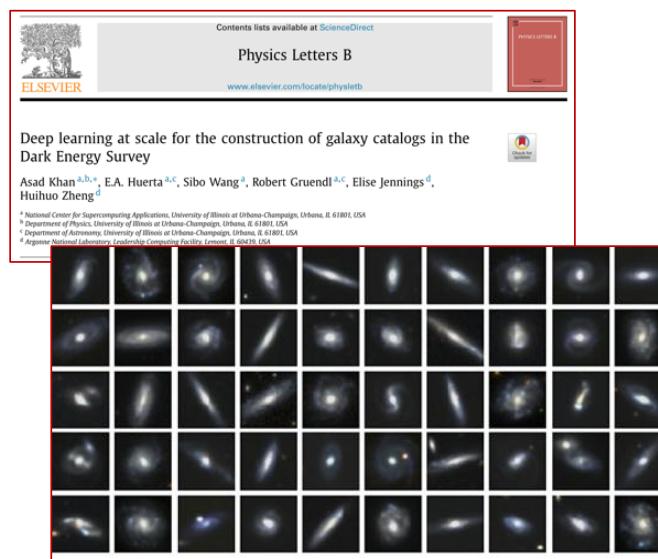
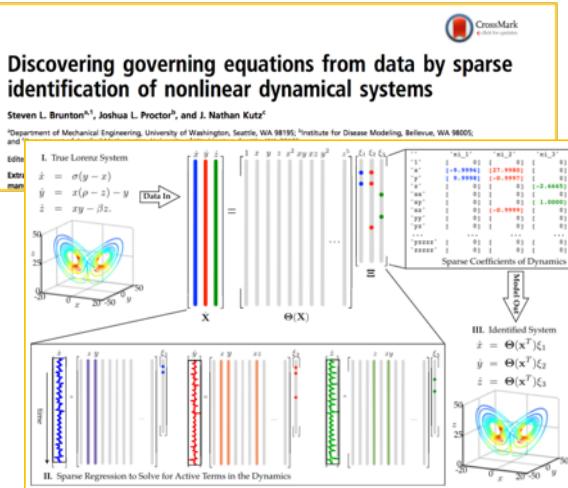


Smart observing
systems

Bridging Physics & AI: Expected breakthroughs



Direct applications of DL schemes to physics-related issues



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)

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

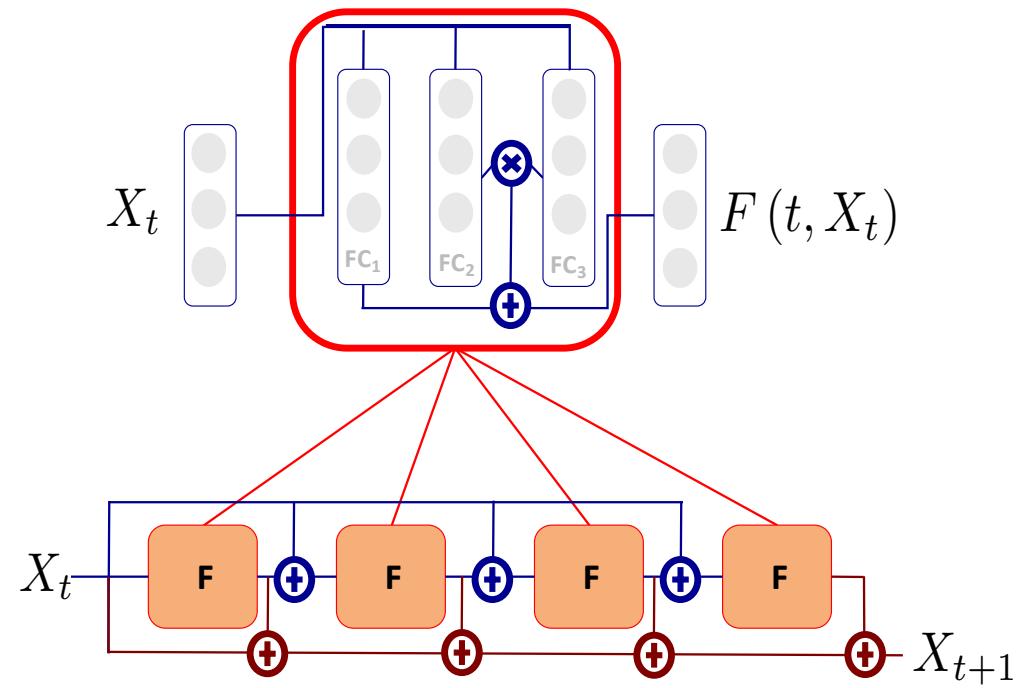


Making the most of AI and Physics Theory

- **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]



$$\begin{aligned}\frac{dx(t)}{dt} &= \sigma(y(t) - x(t)) \\ \frac{dy(t)}{dt} &= x(t)(\rho - z(t)) - y(t) \\ \frac{dz(t)}{dt} &= x(t)y(t) - \beta z(t)\end{aligned}$$

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)

$$\partial_t u + u \partial_x u = \kappa \partial_x^2 u$$



Symbolic calculus
(*Sympy*)



PDE-GenNet
(*keras*)

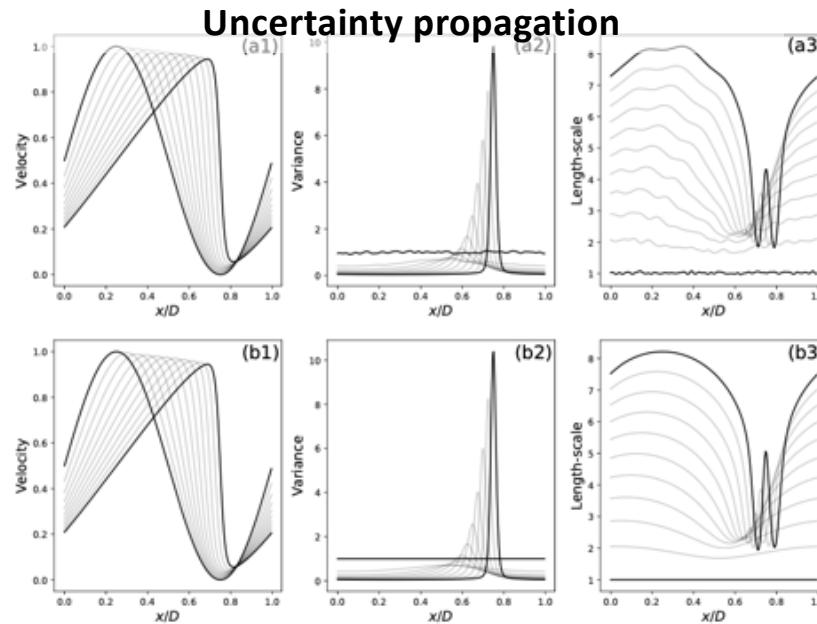


$$\begin{pmatrix} \mu_u(t) \\ \Sigma_u(t) \end{pmatrix} \rightarrow \boxed{\text{ResNet}} \rightarrow \begin{pmatrix} \mu_u(t+1) \\ \Sigma_u(t+1) \end{pmatrix}$$

```
# Example of computation of a derivative
kernel_Du_x_o1 = np.asarray([[0.0,tes.index('x')],0.0],
[0.0,0.0,0.0],
[0.0,1/(2*self.dx[self.coordinates.index('x')]),0.0]]).reshape((3, 3)*(1,1))
Du_x_o1 = DerivativeFactory((3, 3),kernel=kernel_Du_x_o1,name='Du_x_o1')(u)

# Computation of trend_u
mul_1 = keras.layers.multiply([Dkappa_11_x_o1,Du_x_o1],name='MulLayer_1')
mul_2 = keras.layers.multiply([Dkappa_12_x_o1,Du_y_o1],name='MulLayer_2')
mul_3 = keras.layers.multiply([Dkappa_12_y_o1,Du_x_o1],name='MulLayer_3')
mul_4 = keras.layers.multiply([Dkappa_22_y_o1,Du_y_o1],name='MulLayer_4')
mul_5 = keras.layers.multiply([Du_x_o2,kappa_11],name='MulLayer_5')
mul_6 = keras.layers.multiply([Du_y_o2,kappa_22],name='MulLayer_6')
mul_7 = keras.layers.multiply([Du_x_o1,y_o1,kappa_12],name='MulLayer_7')
sc_mul_1 = keras.layers.Lambda(lambda x: 2.0*x,name='ScalarMulLayer_1')(mul_7)
trend_u = K
```

Generated code



Ensemble-based prediction

NN prediction

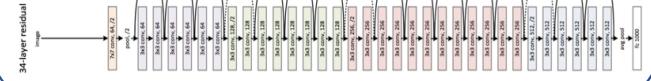
Bridging physics & AI: Expected breakthroughs

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)



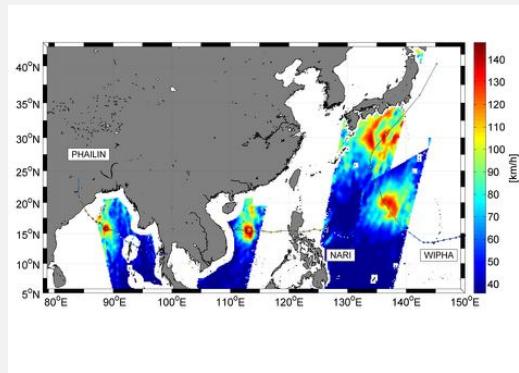
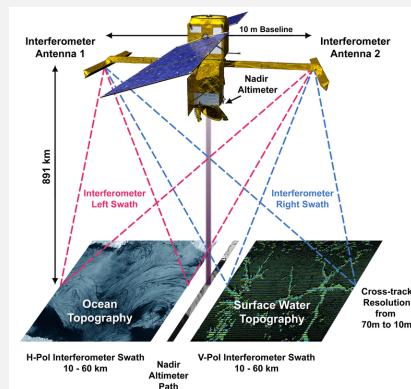
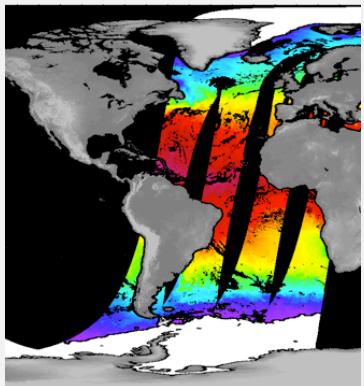
**Dealing with real systems, including
Irregularly-sampled, noisy and/or
partially-observed systems ?**



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

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

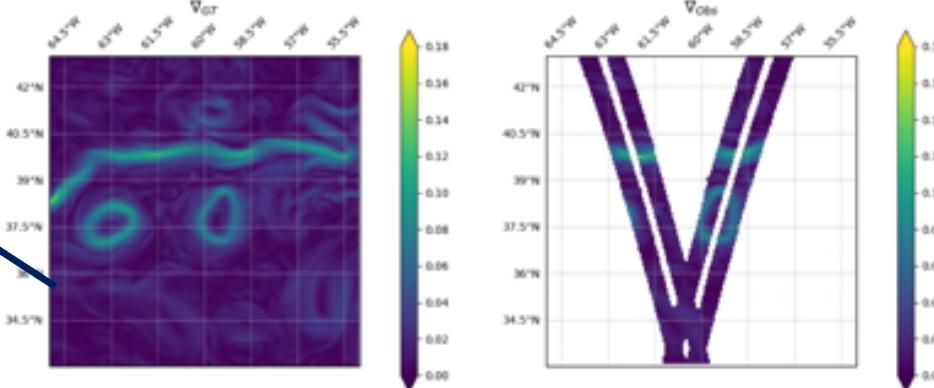


Observation model

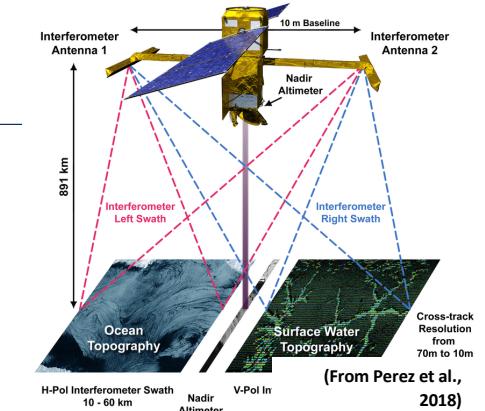
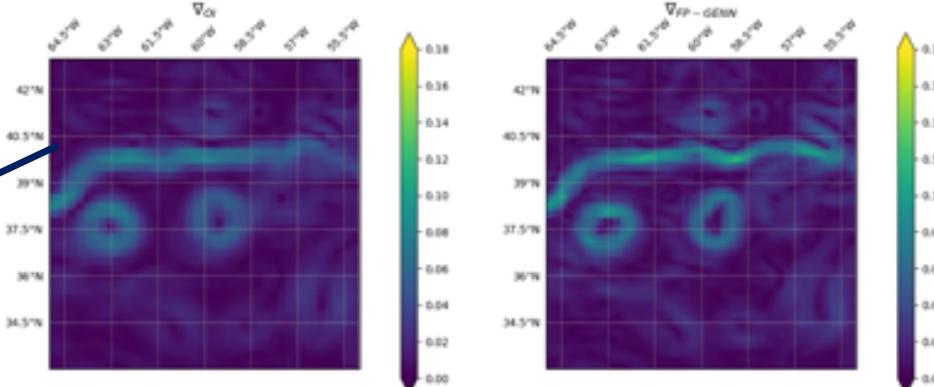
$$Y_t = H(X, \zeta, t, \phi)$$

An example for upcoming SWOT mission

Groundtruth

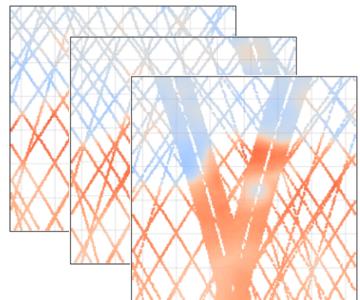


State-of-the-art
operational processing

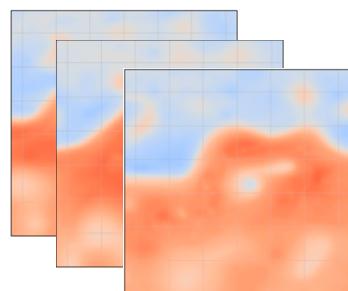


Proposed NN framework
(Fablet et al., 2019)

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



Partial observations y



True states x

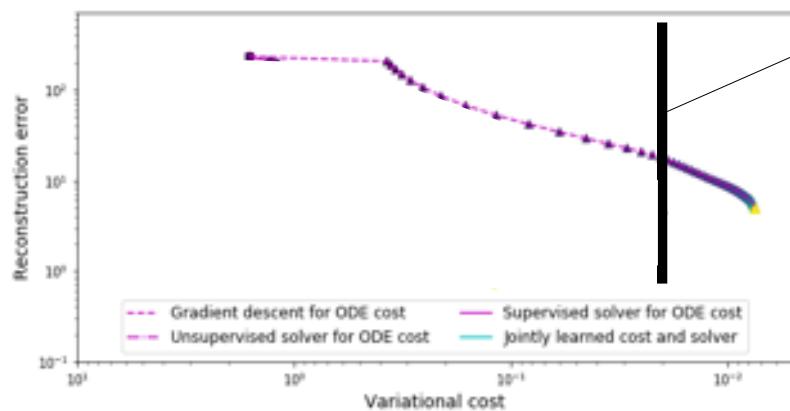
Model-driven schemes: $\hat{x} = \arg \min_x \underbrace{\lambda_1 \|x - y\|_{\Omega}^2 + \lambda_2 \Phi(x)}_{U_{\Phi}(x^{(k)}, y, \Omega)}$

Gradient-based solver (adjoint/Euler-Lagrange method): $U_{\Phi}(x^{(k)}, y, \Omega)$

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

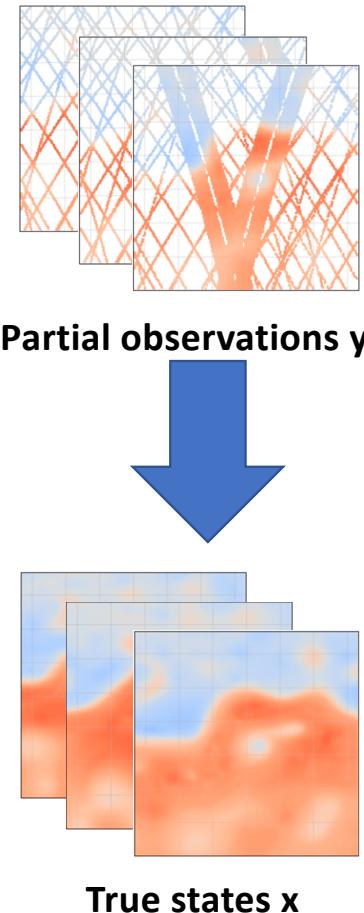
No control on the reconstruction error

$$x^{true} \neq \arg \min_x U_{\Phi}(x^{(k)}, y, \Omega)$$



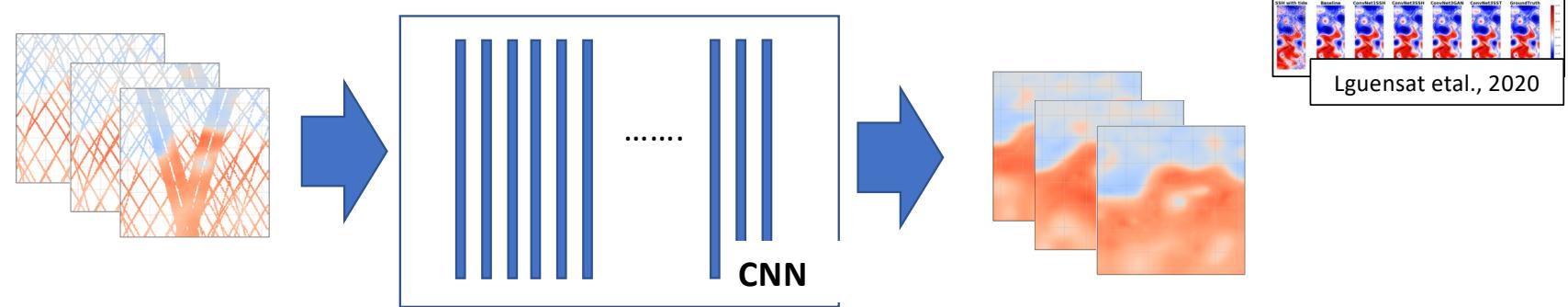
Variational cost
for the true state

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



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

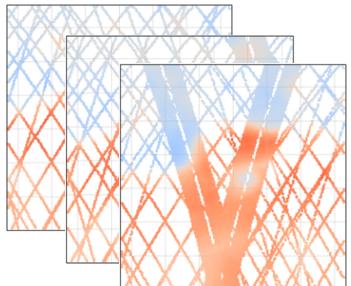
Direct learning for inverse problems: $\hat{x} = \Psi(y)$ $y \rightarrow \text{CNN} \rightarrow x$



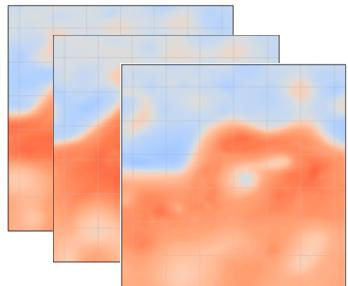
Examples of CNN architectures: Reaction-Diffusion architectures, ADMM-inspired architectures,...

Good performance but possibly weak interpretability/generalization capacities of the solution beyond the training cases

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



Partial observations y



True states x

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

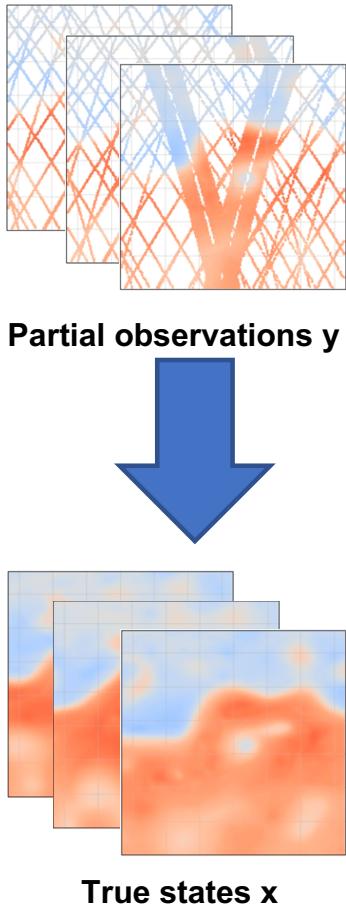
Direct learning for inverse problems: $\hat{x} = \Psi(y) \quad y \rightarrow \text{CNN} \rightarrow x$

Proposed scheme: joint learning of the variational model and solver

- **Theoretical bi-level optimization**

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

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



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

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

Proposed scheme: joint learning of the variational model and solver

- **Theoretical bi-level optimization**

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

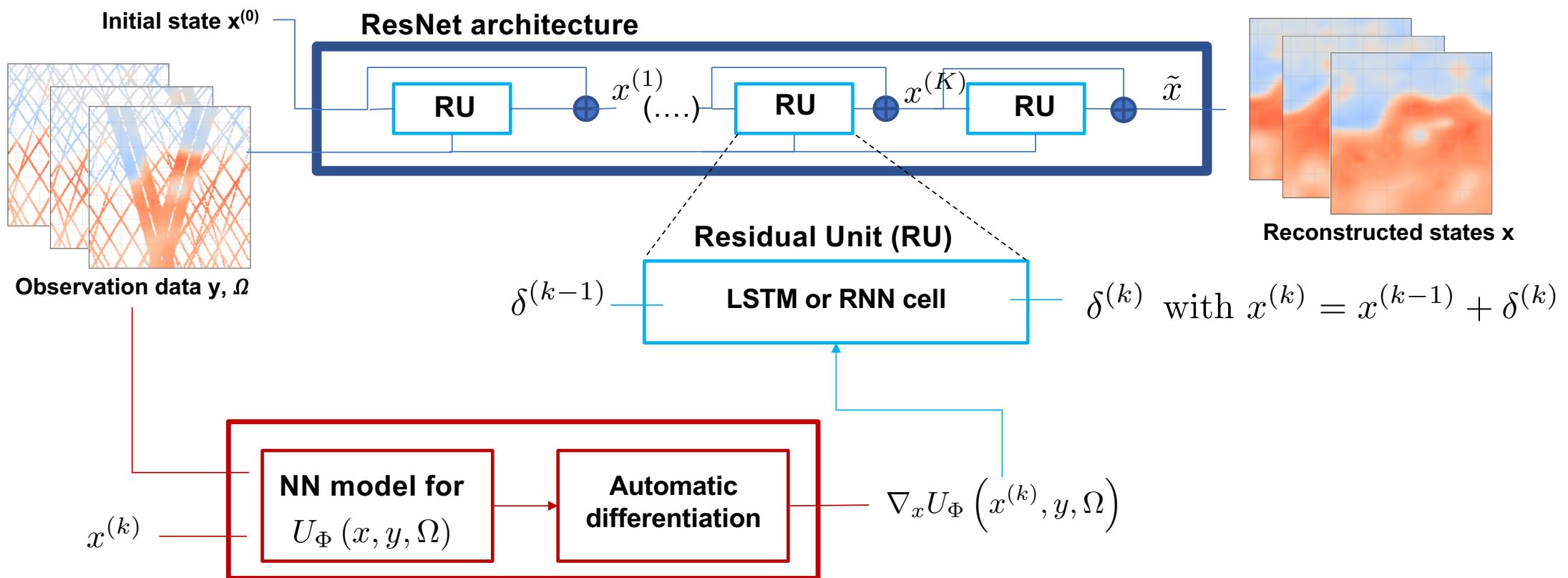
- **Restated with a gradient-based NN solver for inner minimization**

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

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

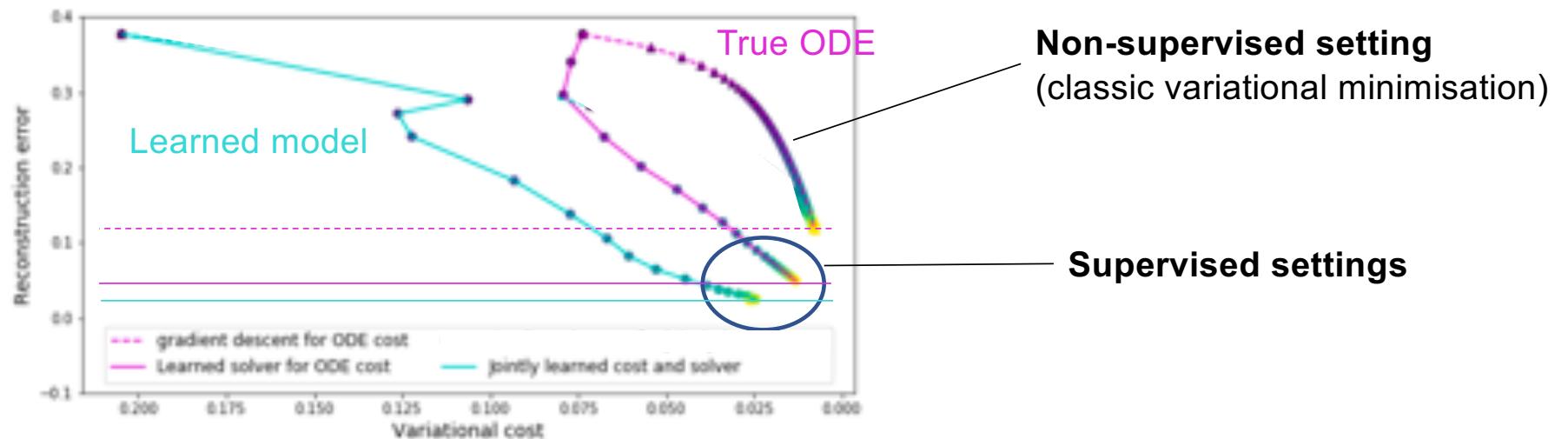
End-to-end learning for inverse problems (Fablet et al., 2020)

Proposed scheme: associated NN architecture

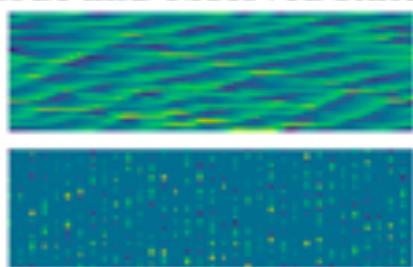


End-to-end learning for inverse problems (Fablet et al., 2020)

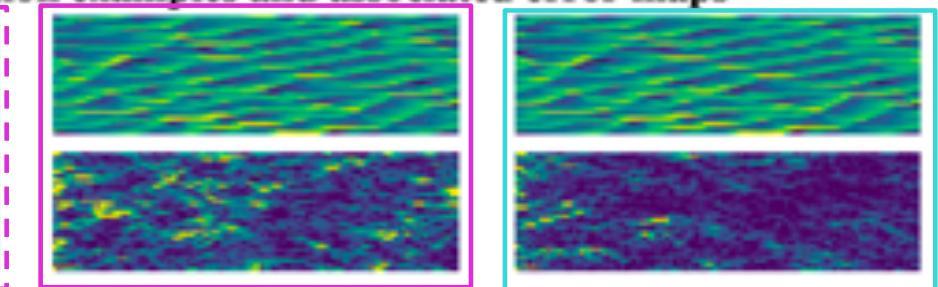
Illustration on Lorenz-96 dynamics (Bilinear ODE)



True and observed states



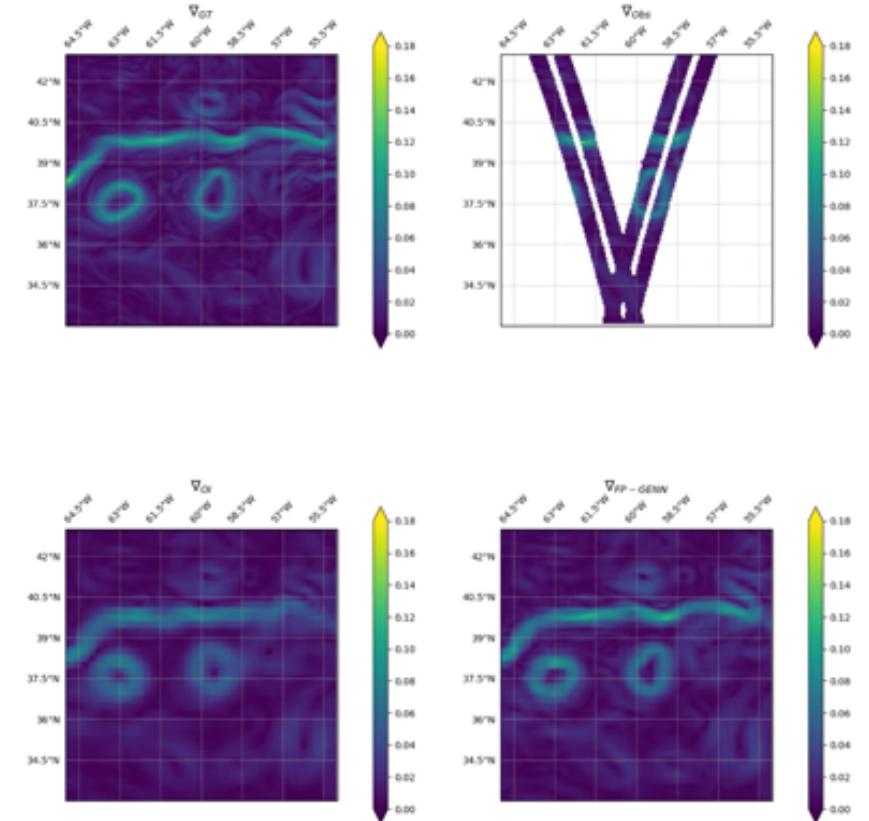
Reconstruction examples and associated error maps



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

Applications to the reconstruction of sea surface current from SWOT data

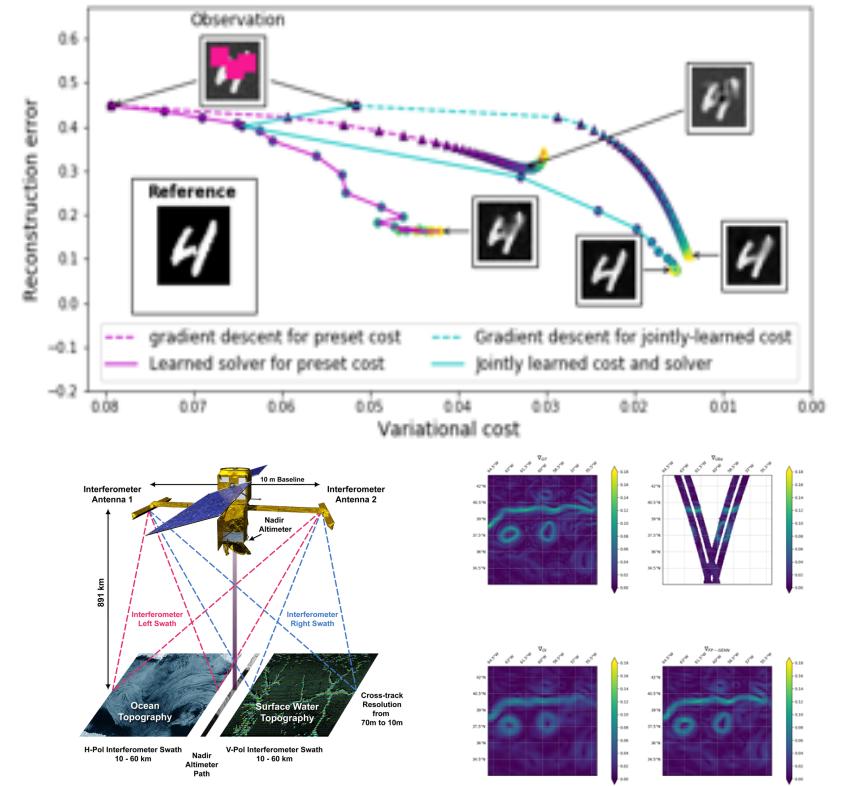
NB: preliminary results with a fixed-point Solver rather than a gradient-based solver



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

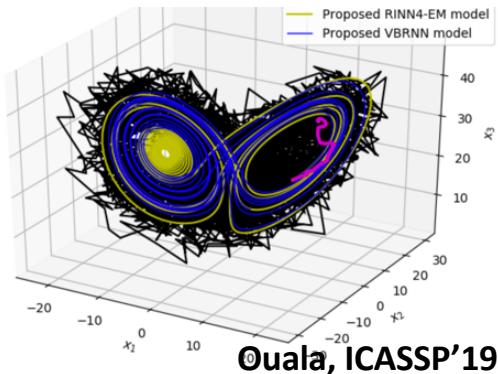


Preprint: <https://arxiv.org/abs/2006.03653>

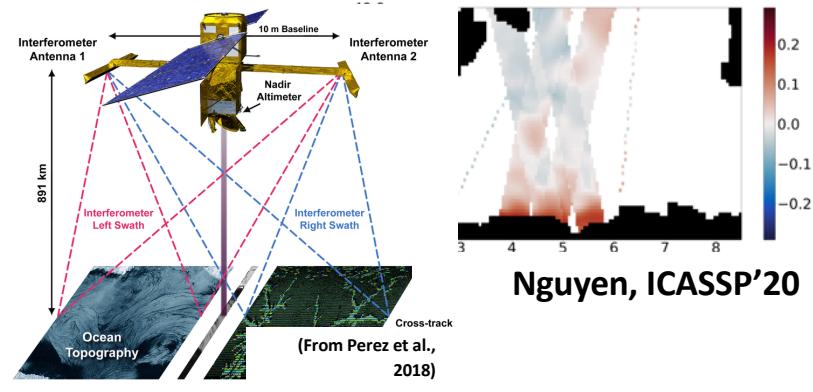
Code: <https://github.com/CIA-Oceanix>

End-to-end learning from real observation data ?

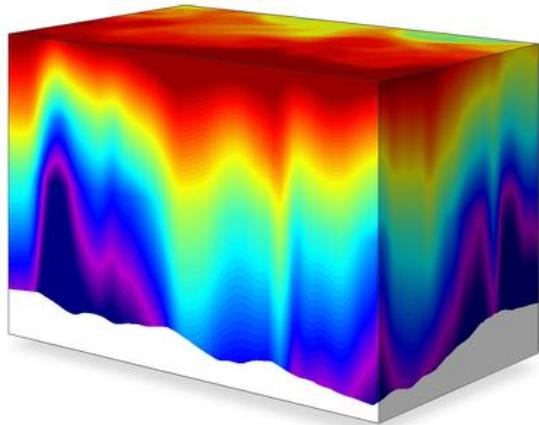
Scarce time sampling



Noisy and irregular sampling



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 et al., 2020)

Beyond Ocean Dynamics

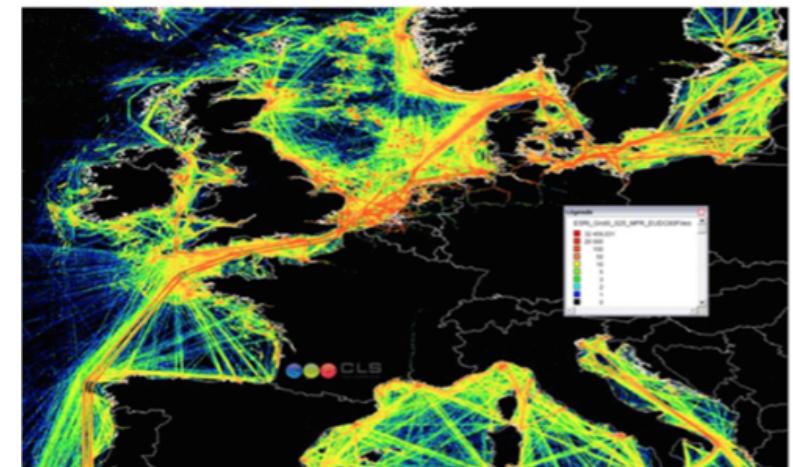
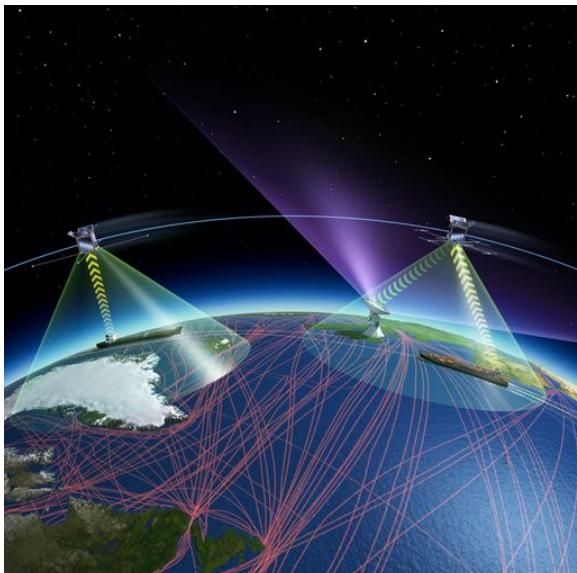
Learning stochastic hidden dynamics



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

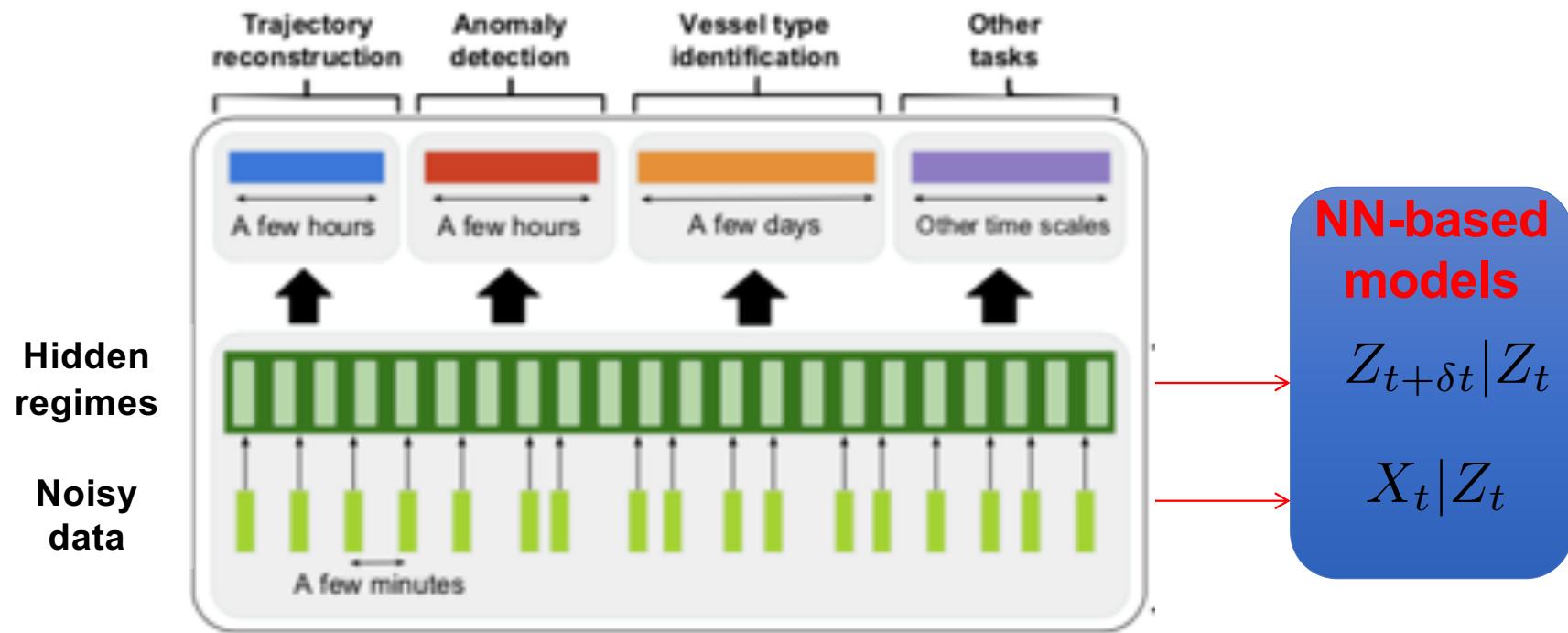
The example of AIS Vessel trajectory data

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



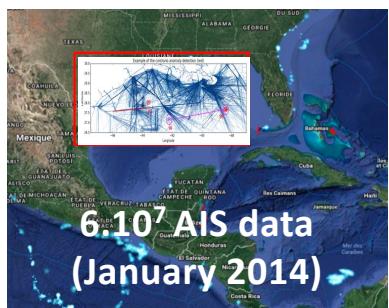
How can we learn from AIS data streams ?

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



Model training from noisy AIS streams using variational Bayesian approximation

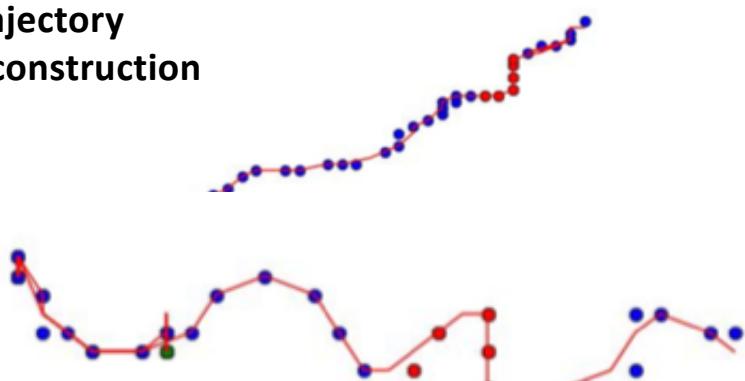
Learning stochastic hidden dynamics [Nguyen et al., 2018]



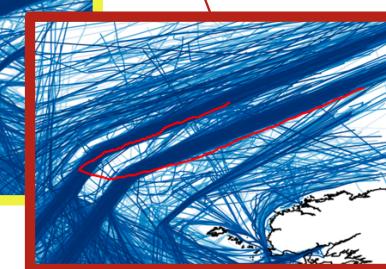
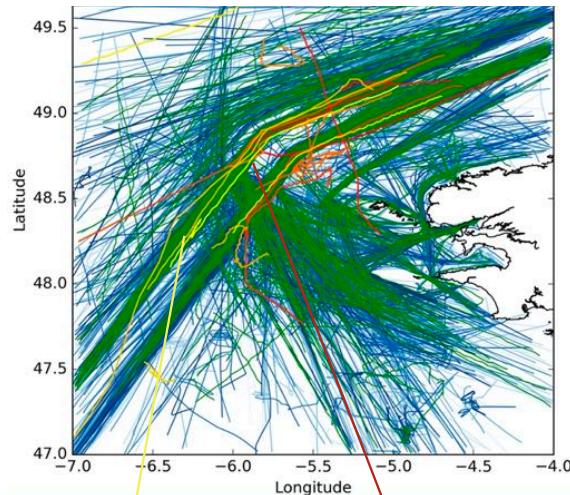
Vessel type
recognition

~88% of correct
recognition

Trajectory
reconstruction



Abnormal behaviour detection



Beyond Ocean Dynamics

Dynamical System Theory for Deep
Learning



Understanding DL models ?



88% **tabby cat**

adversarial
perturbation

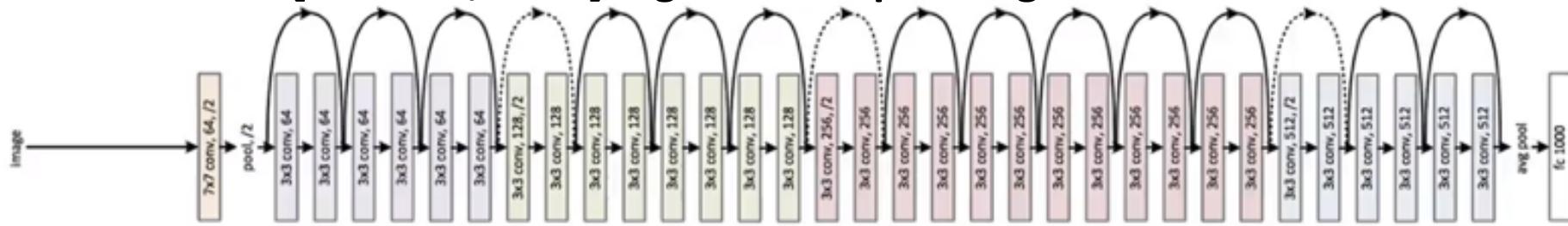


99% **guacamole**

Szegedy et al., 2015

Understanding ResNets [Rousseau et al., 2019]

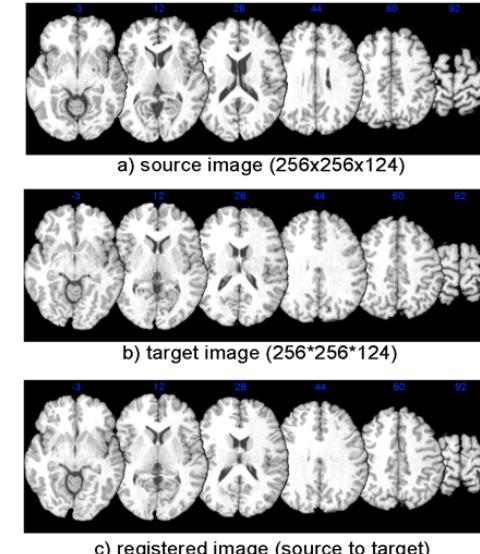
ResNet [He et al., 2015] regarded as space registration machines



- Image registration examples



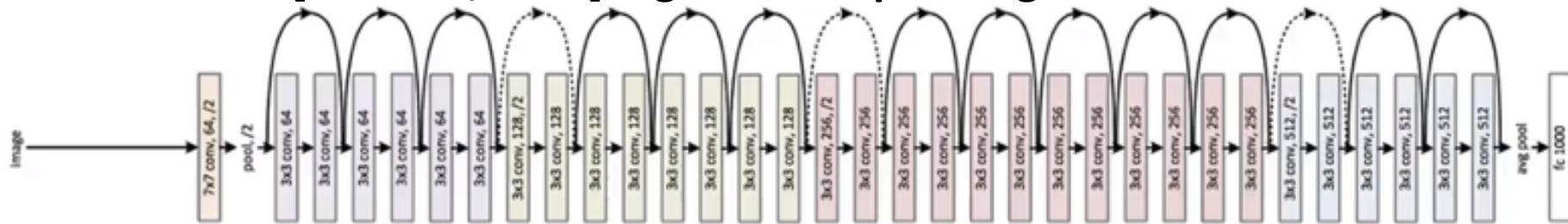
[Matlab tutorial]



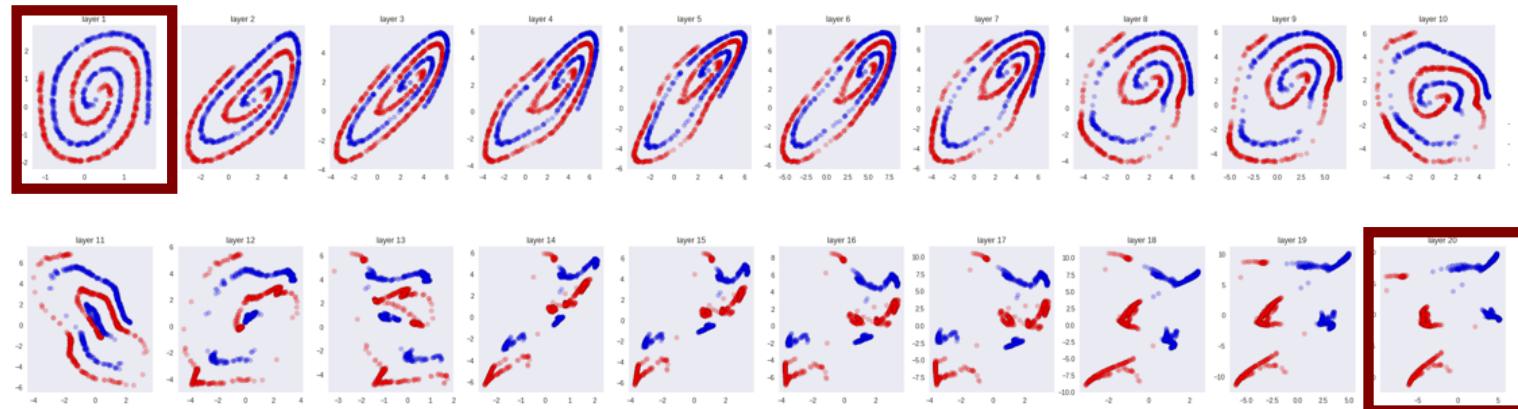
[Dramms tutorial]

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](https://www.researchgate.net/profile/Ronan_Fablet)

