End-to-end and physics-informed learning for space oceanography

R. Fablet, Prof. IMT Atlantique
Lab-STICC, INRIA Team Odyssey

Joint work with B. Chapron, L. Drumetz, F. Rousseau, E. Mémin, S. Ouala, M. Beauchamp, Q. Febvre, A. Pascual
…

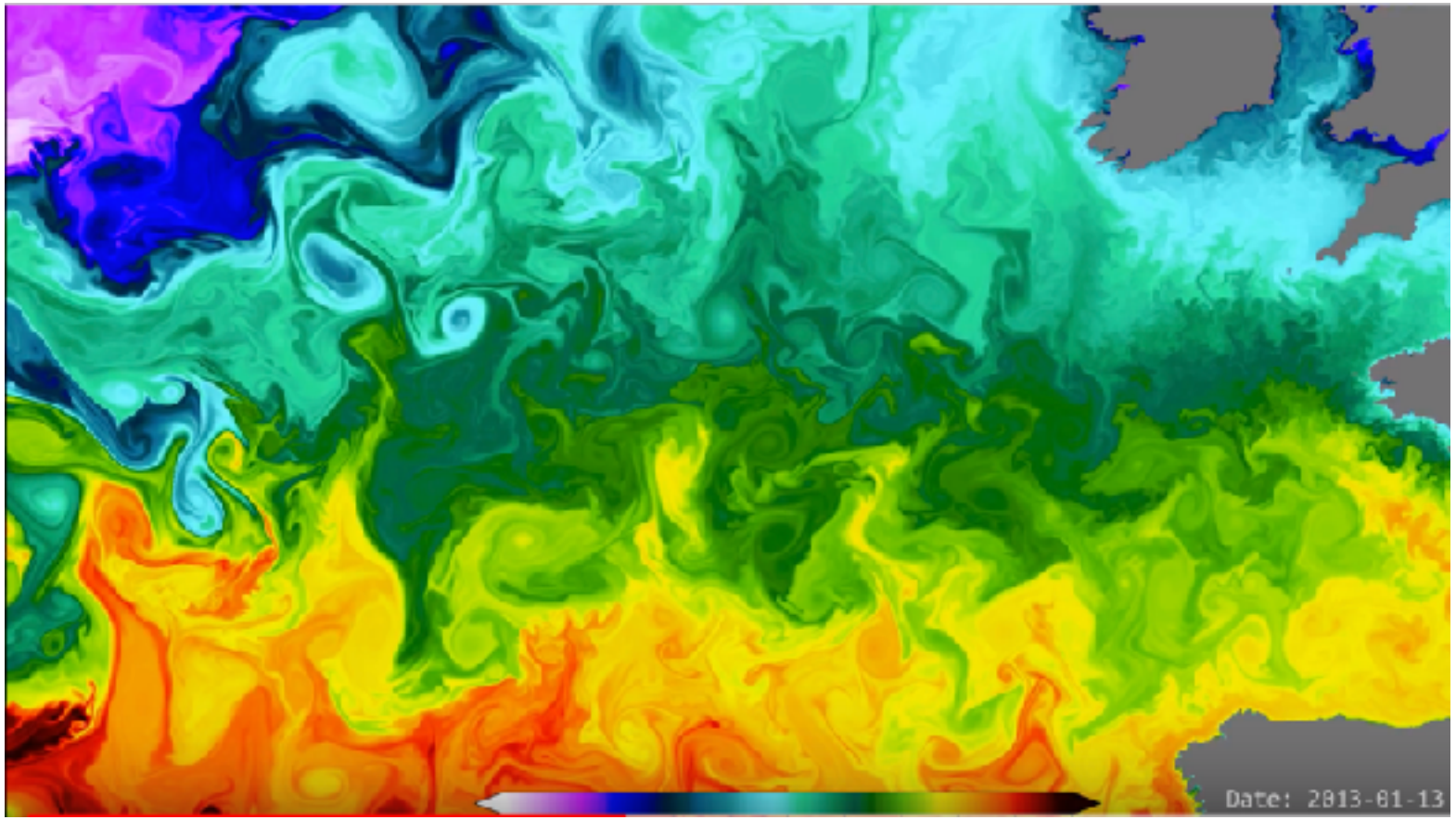
March 2022

Overview of the talk

- ***Challenges in ocean observation, modeling and forecasting***
- ***Beyond Neural Networks regarded as black boxes***
- ***End-to-end learning can make a difference***

Challenges in Ocean Modeling and Forecasting

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



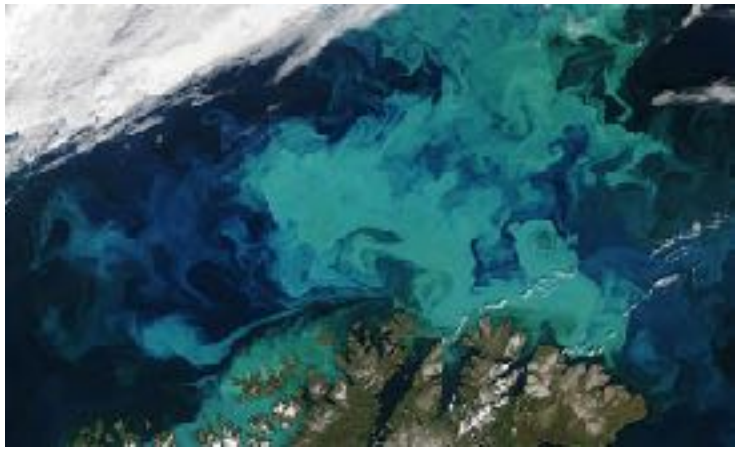
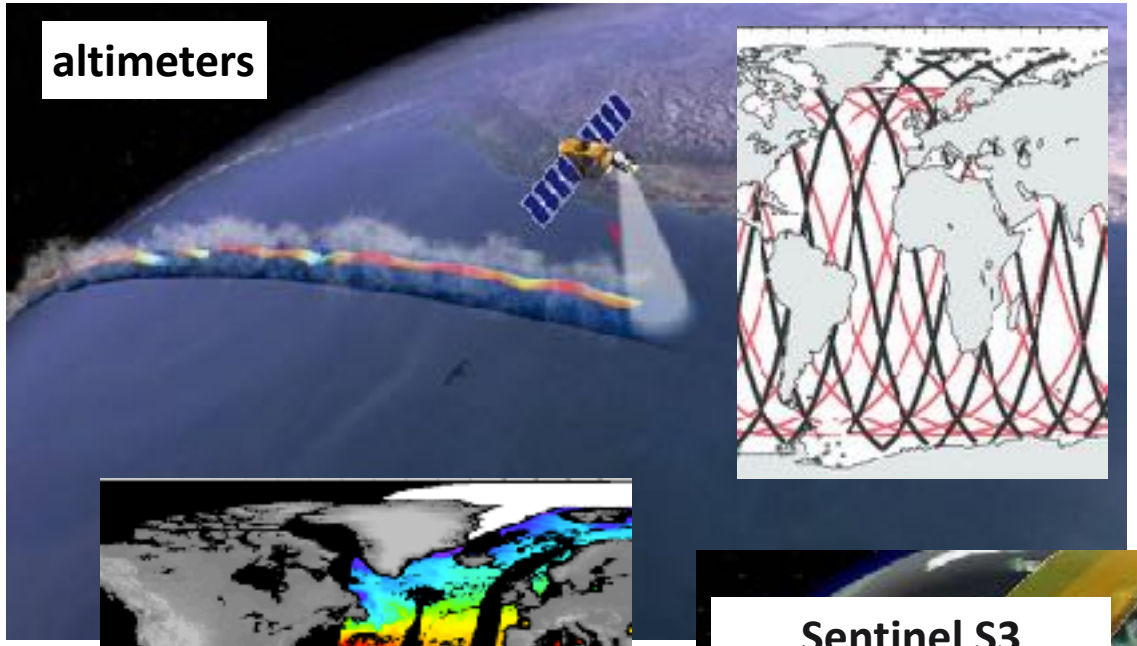
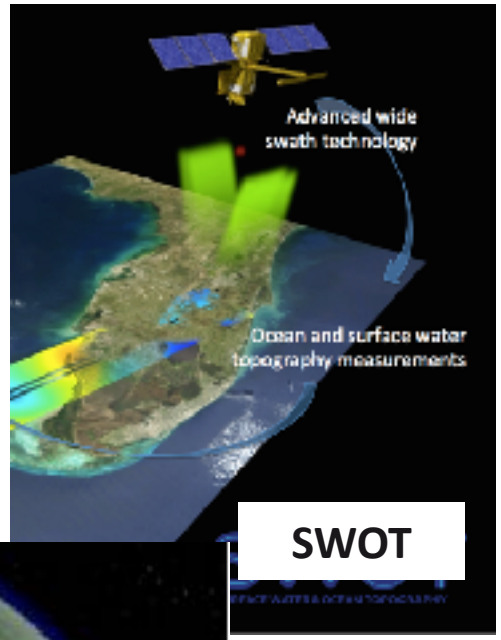
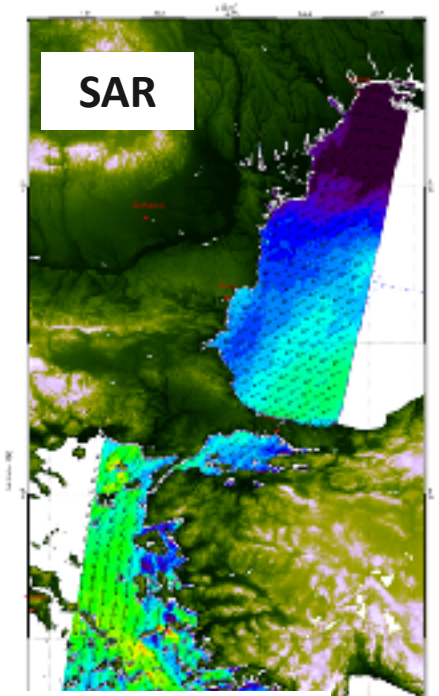
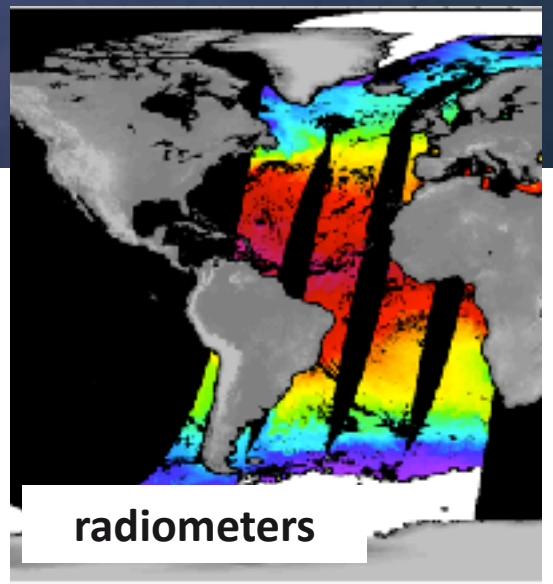


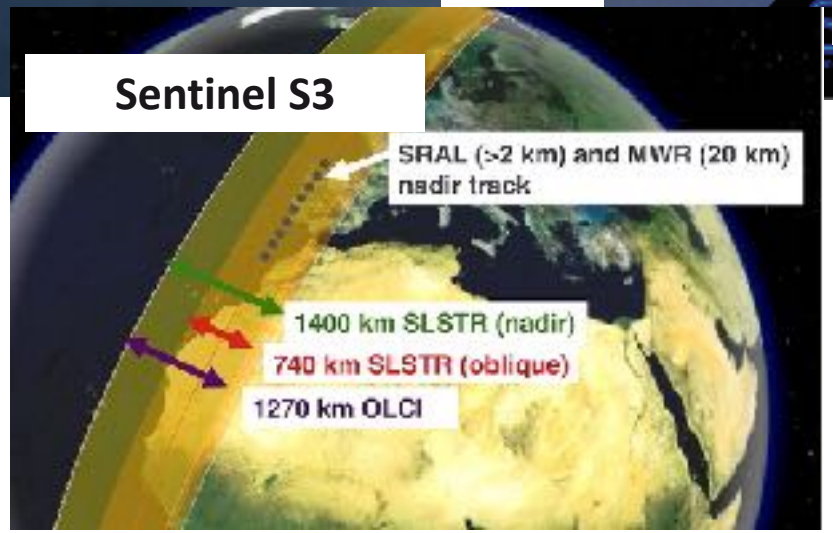
Illustration of satellite remote sensing observations for ocean dynamics

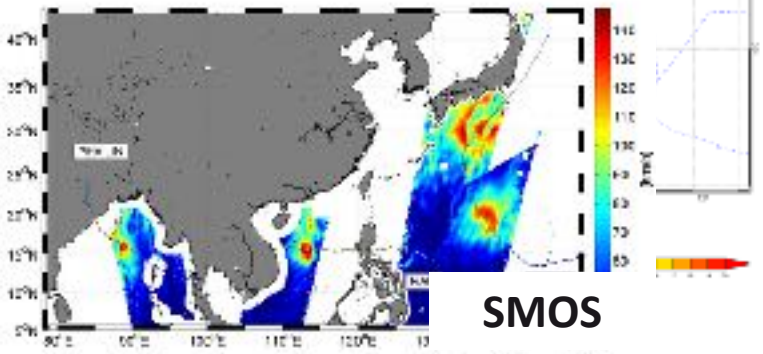






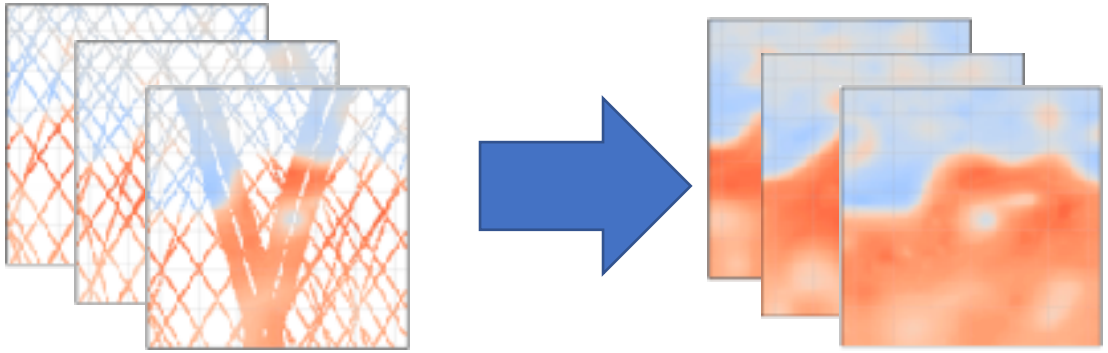






Challenges in space oceanography

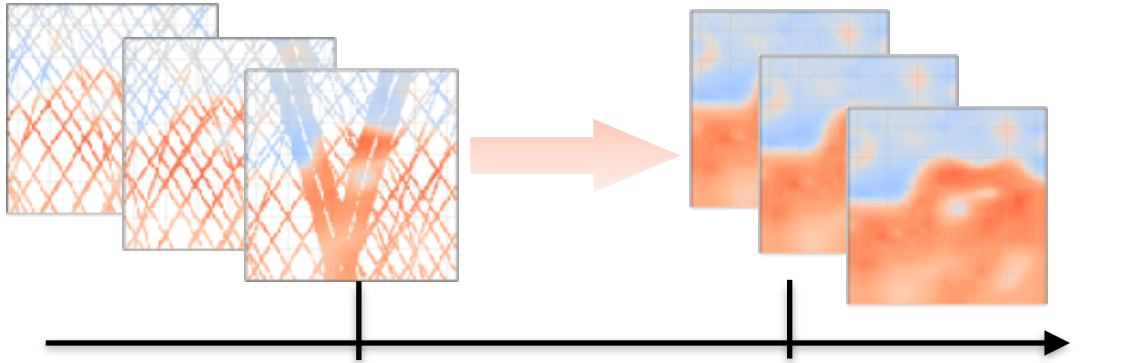
How to reconstruct from observations ?



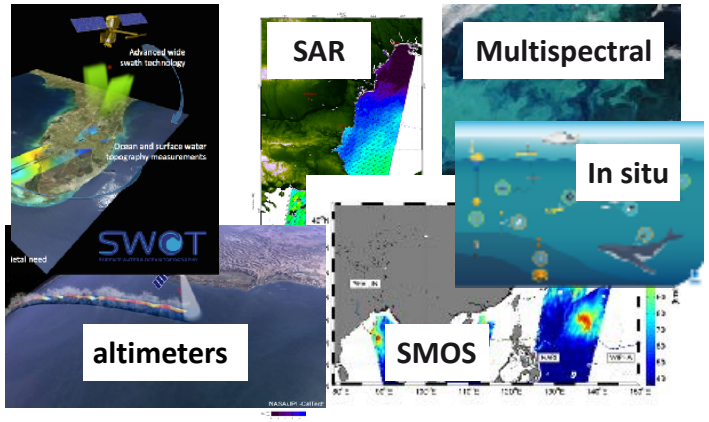
Partial observations y

Gap-free states x

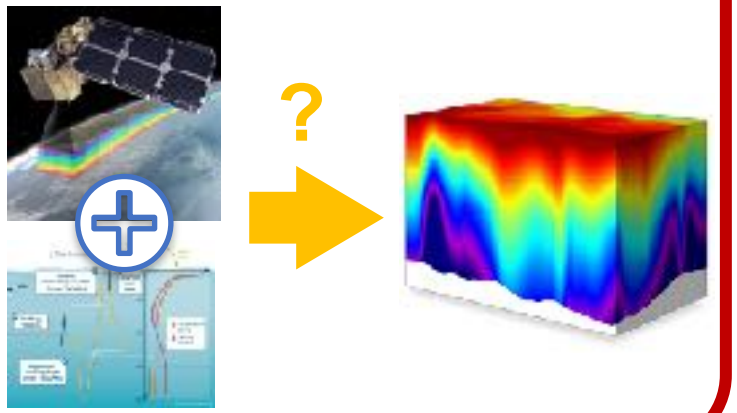
How to model/forecast from observations ?



Using multi-tracer observations ?



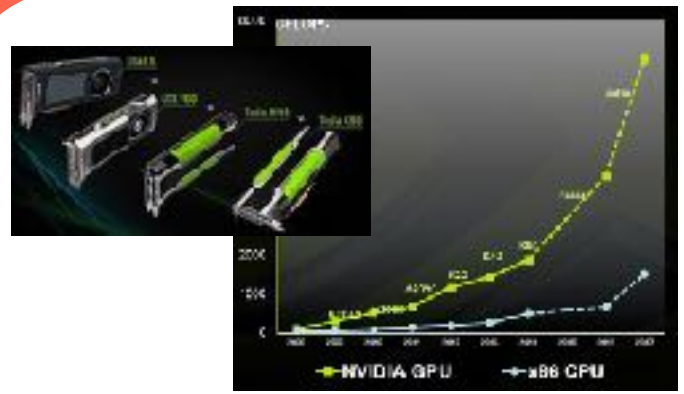
From surface to interior ?



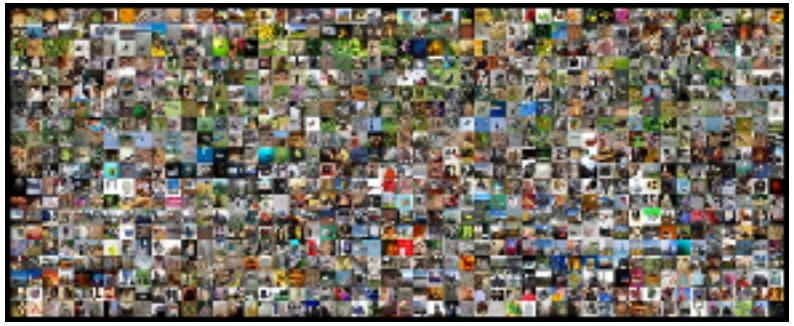
Where and what to sample ?

Deep Learning and Earth Sciences

Key reasons for the emergence of DL



High-performance computing (GPU)



Large annotated dataset (> 1M)

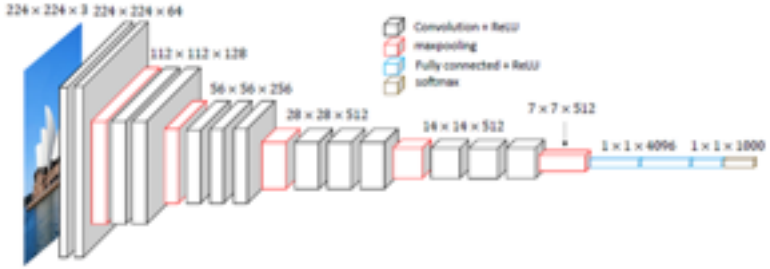
+

TensorFlow

Keras

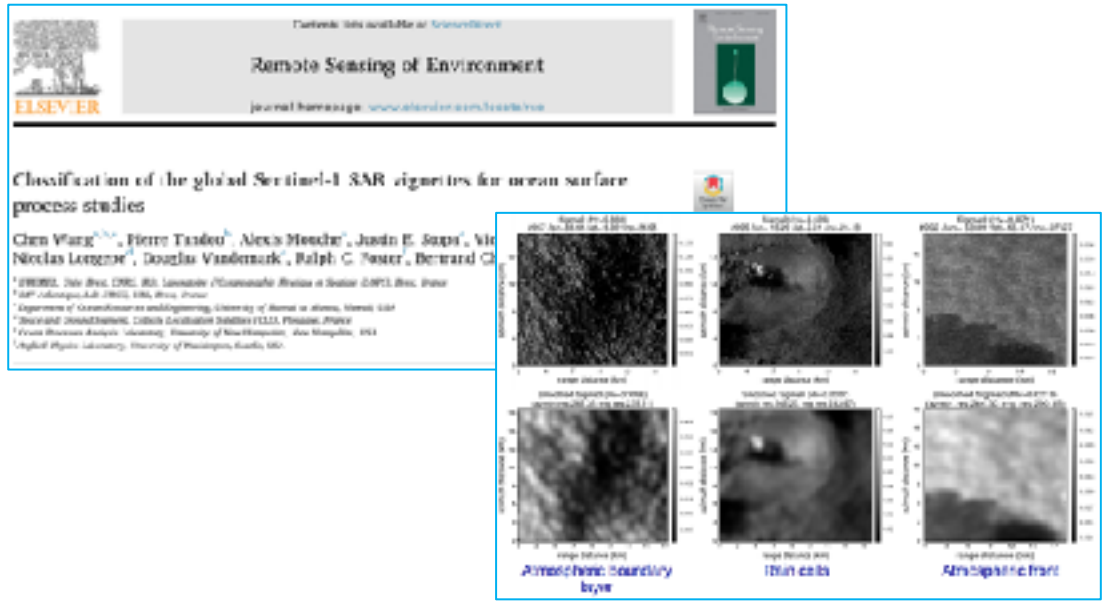
PYTORCH

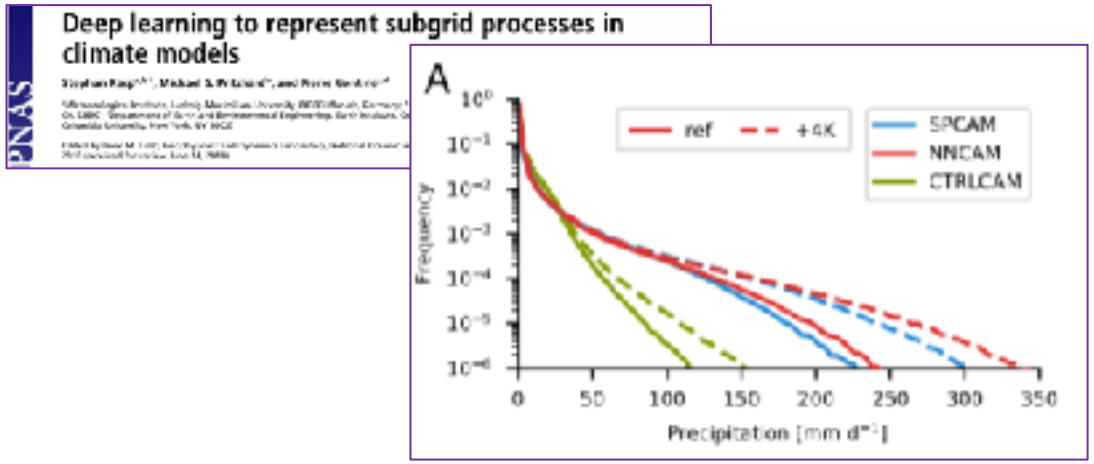
Efficient & easy-to-use frameworks

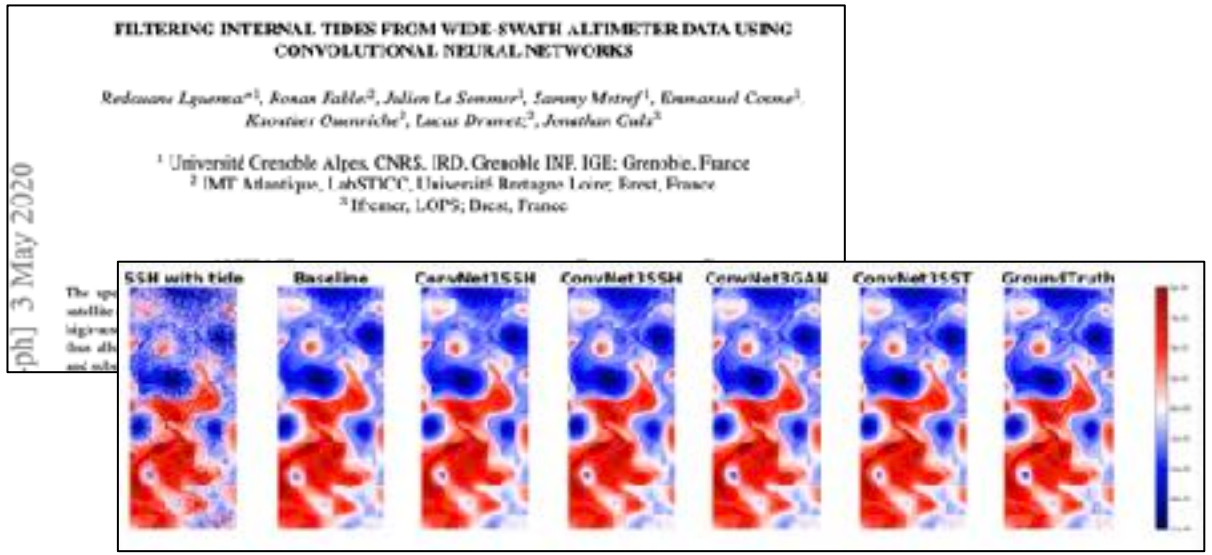


End-to-end learning

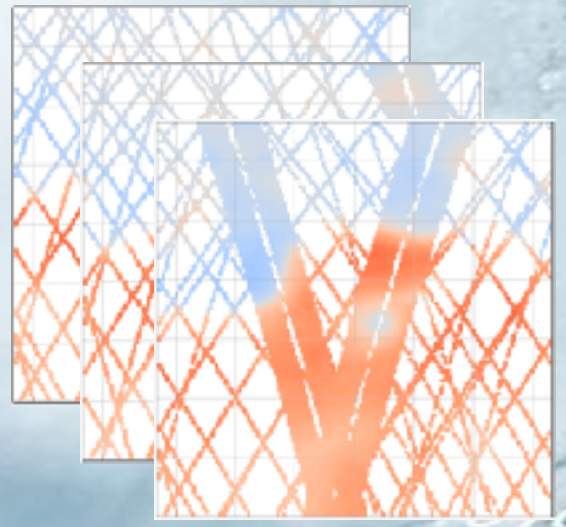
“Off-the-shelf” DL schemes applied to physics-related issues





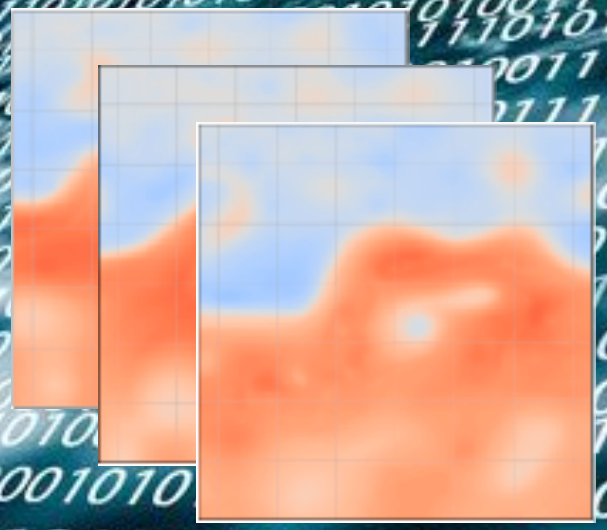


End-to-end physics-informed Learning for reconstruction and forecasting problems



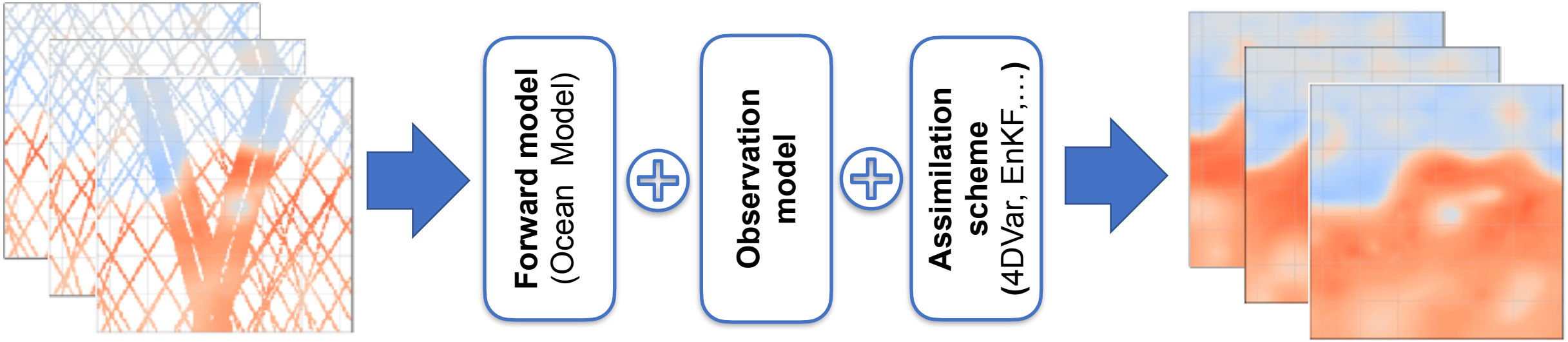
Partial observations y

?



True states x

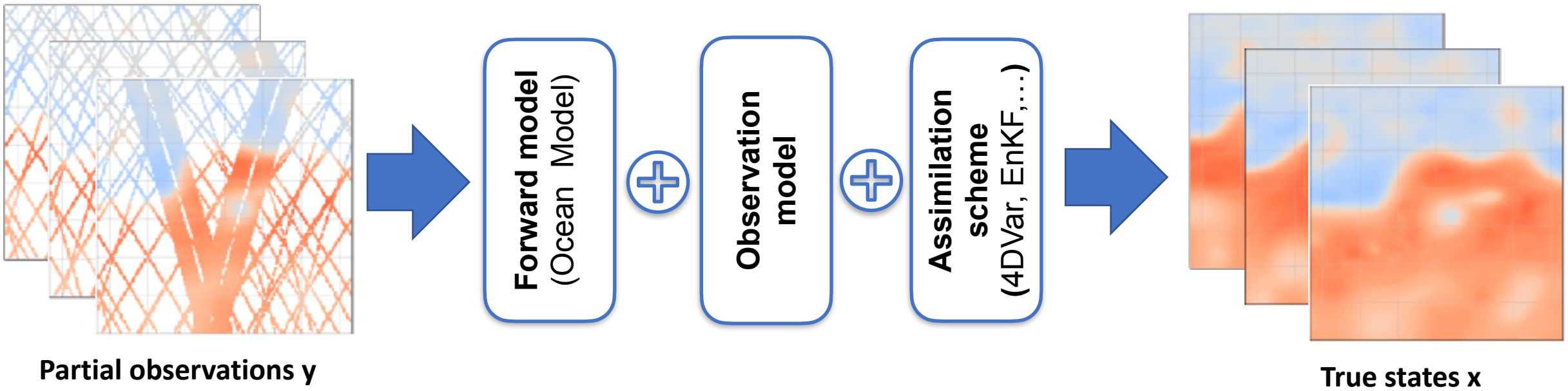
Data assimilation in earth sciences [Evensen, 2000]



Partial observations y

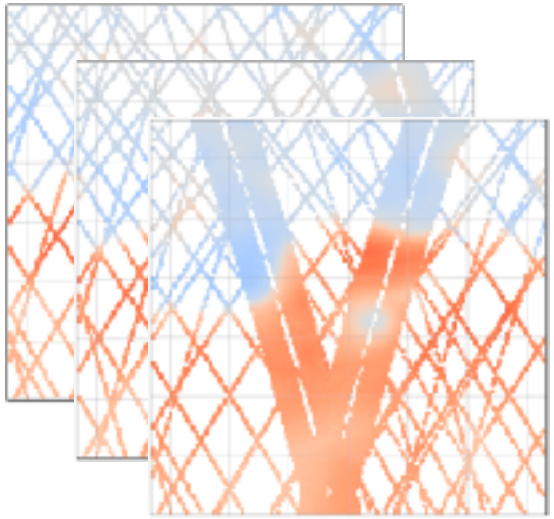
True states x

Data assimilation in geoscience [Evensen, 2000]



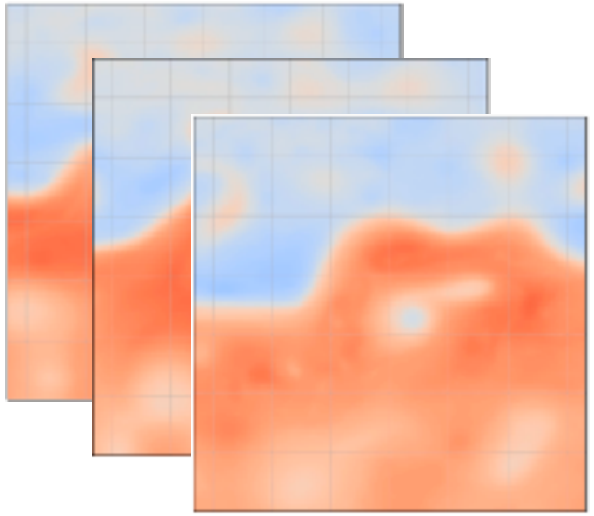
Each component designed using model-driven principles and mostly independently…. But limited ability to fully exploit observation datasets.

What about end-to-end learning for data assimilation?



Partial observations y

End-to-end neural architecture

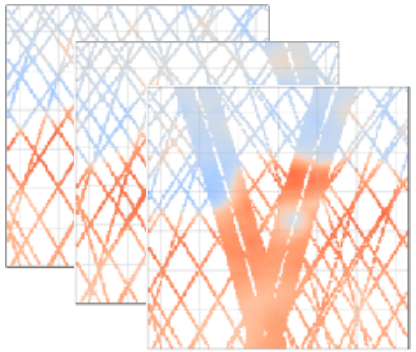


True states x

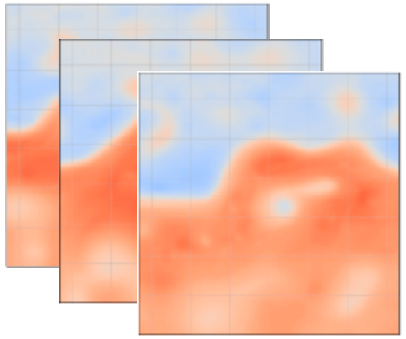
Can we exploit prior knowledge ?

Can we calibrate all the components of a DA scheme at once?

(Weak constraint) **4DVar Data Assimilation (DA) formulation**



Partial observations y



True states x

State-space formulation:

$$\begin{cases} \dfrac{\partial x(t)}{\partial t} & = & \mathcal{M}\left(x(t)\right) \\ \\ y(t) & = & x(t) + \epsilon(t), \forall t \in \{t_0, t_0 + \Delta t...., , t_0 + N\Delta t\} \end{cases}$$

Associated variational formulation:

$$\arg\min_x \lambda_1 \sum_i \|x(t_i) - y(t_i)\|^2_{\Omega_{t_i}} + \lambda_2 \sum_n \|x(t_i) - \Phi(x)(t_i)\|^2$$

with $$\Phi(x)(t) = x(t - \Delta) + \int_{t-\Delta}^{t} \mathcal{M}\left(x(u)\right) du$$

$$\boxed{\arg\min_x \lambda_1 \|x - y\|^2_\Omega + \lambda_2 \|x - \Phi(x)\|^2}$$

4DVarNet: Learning 4DVar models and solvers

Trainable Variational DA formulation

$$\widehat{x} = \arg\min_{x} \|y - H(x)\|^2 + \lambda \|x - \Phi(x)\|^2$$

Trainable or pre-defined observation model

Trainable or pre-defined prior

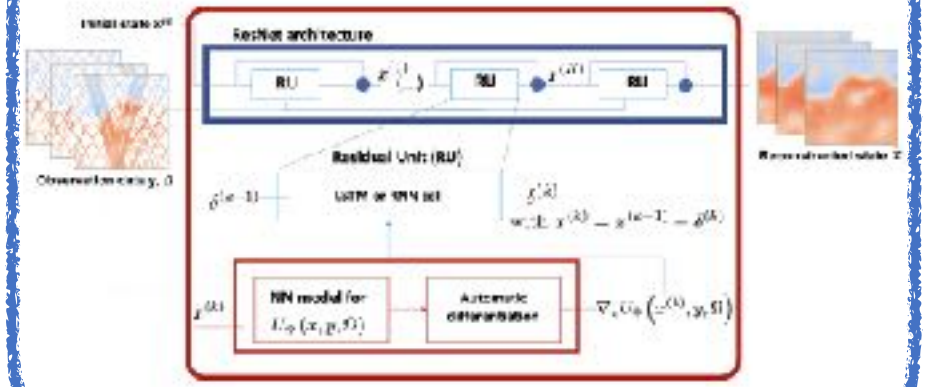
Trainable solver

$$x^{(k+1)} = x^{(k)} - \mathcal{H}\left[\nabla_x U_\Phi\left(x^{(k)}, y\right)\right]$$

RNN

Automatic differentiation

End-to-end architecture



Preprint: https://arxiv.org/abs/2006.03653
Code: https://github.com/CIA-Oceanix/DinAE_4DVarNN_torch

4DVarNet: Learning 4DVar models and solvers

Trainable Variational DA formulation

$$\widehat{x} = \arg\min_{x} \|y - H(x)\|^2 + \lambda \|x - \Phi(x)\|^2$$

Trainable or pre-defined observation model

Trainable or pre-defined prior

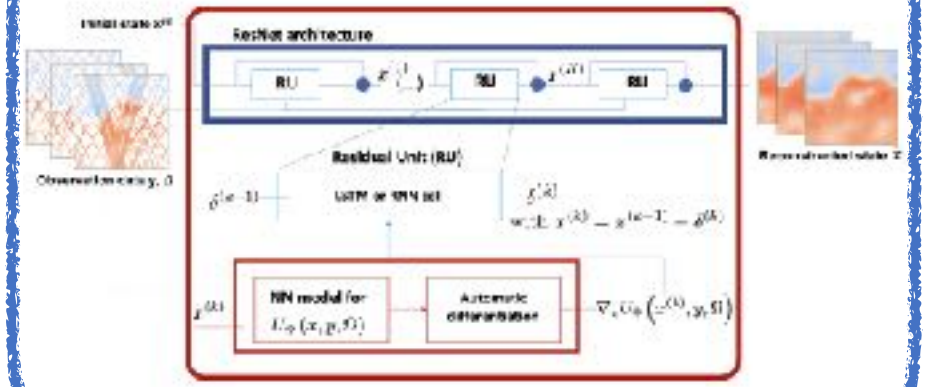
Trainable solver

$$x^{(k+1)} = x^{(k)} - \mathcal{H}\left[\nabla_x U_\Phi\left(x^{(k)}, y\right)\right]$$

RNN

Automatic differentiation

End-to-end architecture



Learning criterion

Variational cost (non-supervised)

Reconstruction error (supervised)

Preprint: https://arxiv.org/abs/2006.03653
Code: https://github.com/CIA-Oceanix/DinAE_4DVarNN_torch

4DVarNet: Application to sea surface dynamics

Method

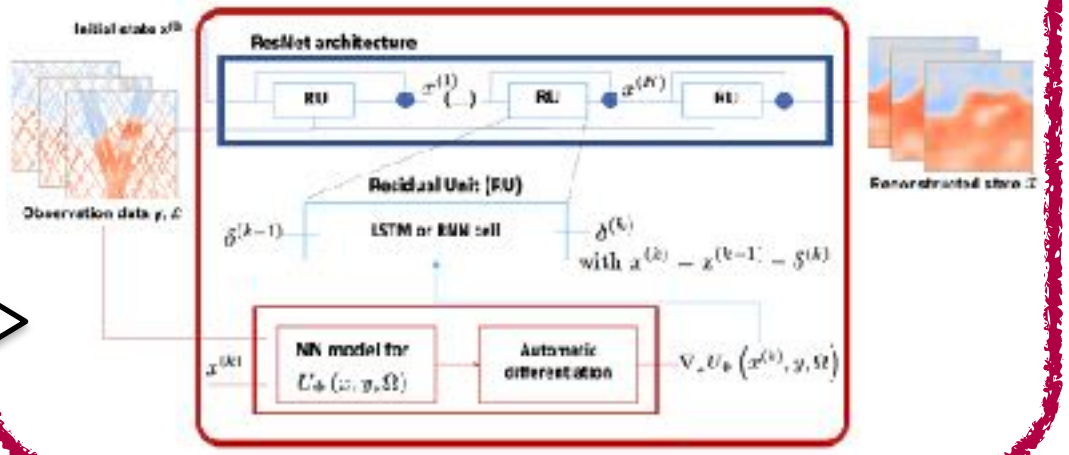
From a Variational DA formulation

$$\widehat{x} = \arg\min_{x} \|y - H(x)\|^2 + \lambda \|x - \Phi(x)\|^2$$

Trainable variational model

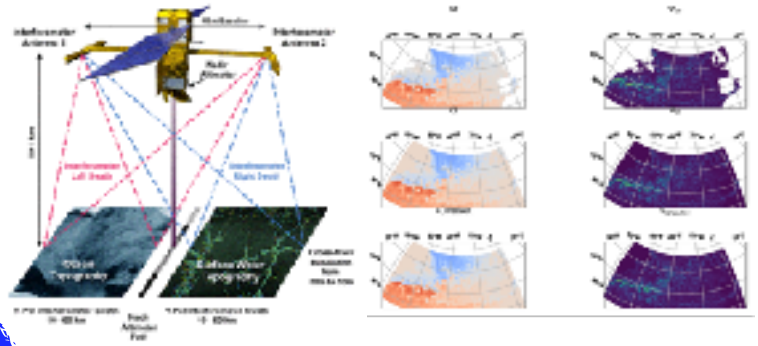
Trainable gradient-based solver

Associated end-to-end scheme

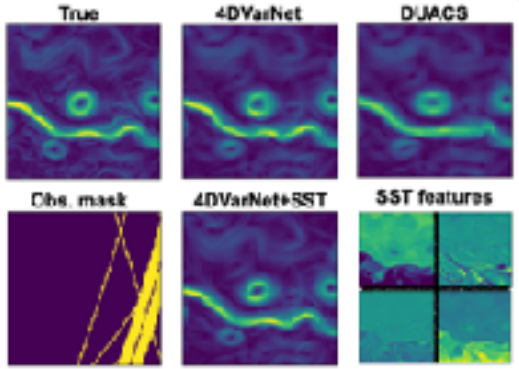


Applications

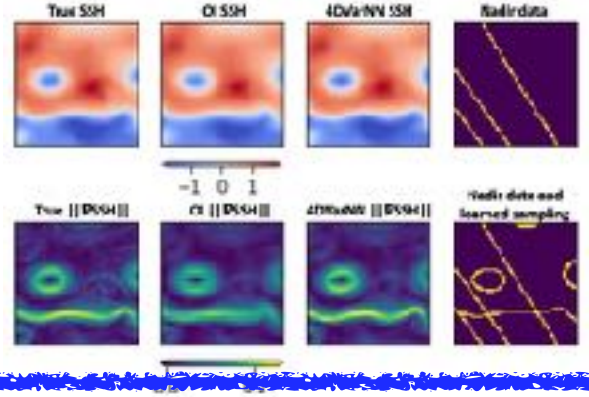
Interpolation & Forecasting



Multimodal DA

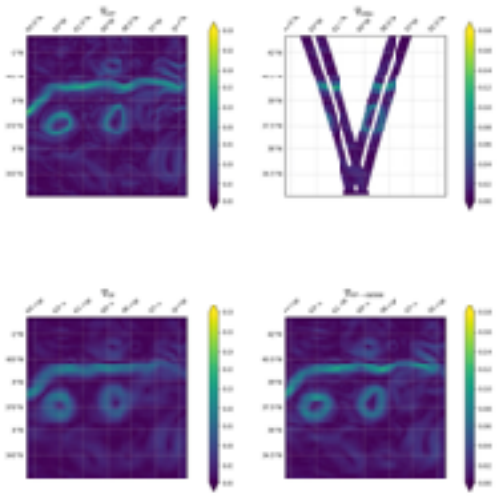


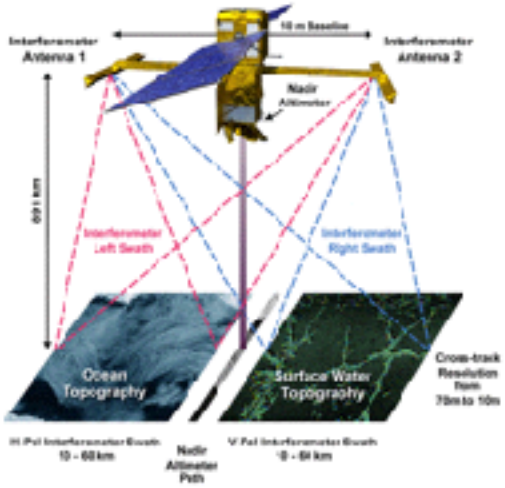
Learning where to sample ?



Space-time interpolation of sea surface geophysical fields

Satellite altimetry





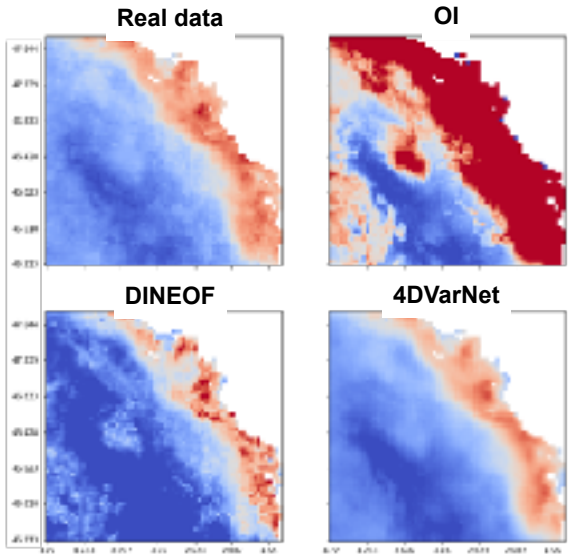
Best score for BOOST-SWOT SLA Data Challenge

duacs: 1 swot + 4 nadirs	0.82	0.22	1.22	11.16	Covariances DUACS	eval_duacs.ipynb
bfn: 1 swot + 4 nadirs	0.83	0.22	0.8	10.89	QG Nudging	eval_bfn.ipynb
dymos: 1 swot + 4 nadirs	0.83	0.22	1.2	12.07	Dynamic mapping	eval_dymos.ipynb
miost: 1 swot + 4 nadirs	0.84	0.01	1.18	10.14	Multiscale mapping	eval_miost.ipynb
4DVarNet: 1 swot + 4 nadirs	0.86	0.01	0.82	6.67	4DVarNet mapping	eval_4dvarnet.ipynb

https://github.com/ocean-data-challenges/2020a_SSH_mapping_NATL60

Sea surface suspended sediments

Metric	Dataset	Unit	Samp. Strat	OI	DinEOF	4DVarNet
RMSE	OSSE	$\log_{10}[g/L]/m$	-	0.176	0.167	**0.104**
	MODIS	$\log_{10}[g/L]/m$	Random	0.304	0.237	**0.156**
	MODIS	$\log_{10}[g/L]/m$	Patch	0.346	0.253	**0.168**
R-score	OSSE	%	-	90.4	91.3	**96.6**
	MODIS	%	Random	60.5	76.4	**89.5**
	MODIS	%	Patch	56.5	73.8	**87.3**



Mean gradient norm

Simulation and real data

Learning from real gappy data only

Collab. IMT Atl./SHOM/LGO

Trainable observation operators

Learning where to sample ?

Learning what to measure ?

4DVarNet models with trainable observation models

$$\hat{x} = \arg\min_x \|x - y\|^2 + \lambda \|x - \phi(x)\|^2$$

Sparse sampling operator

$$\|H(z) * (x - y)\|^2$$

$$\text{s.t. } \forall z, \ \|H(z)\|_1 < \epsilon$$

Multimodal observation

$$\|x - y\|^2 + \alpha \|G * x - F * z\|^2$$

Trainable observation operators

4DVarNet models with trainable observation models

Spase sampling operator

$$\|H(z) * (x - y)\|^2$$

$$\text{s.t. } \forall z, \ \|H(z)\|_1 < \epsilon$$

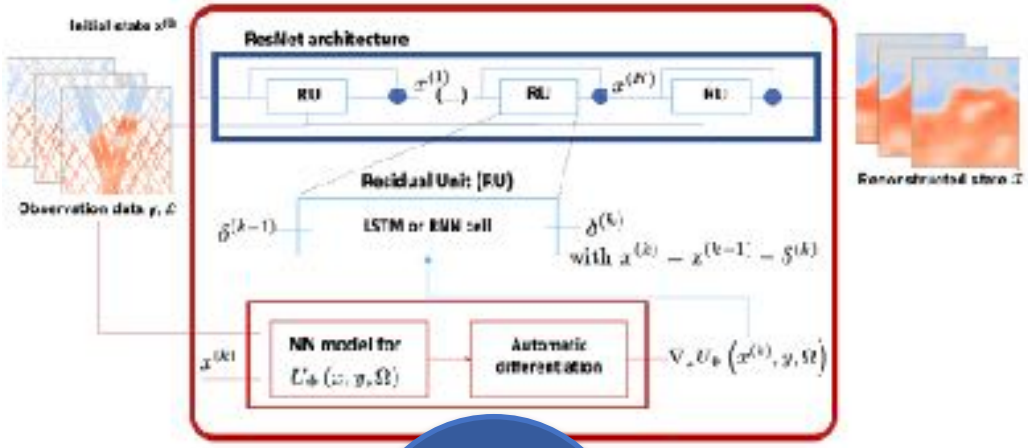
Multimodal observation

$$\|x - y\|^2$$

$$+$$

$$\alpha \|G * x - F * z\|^2$$

End-to-end 4DVarNet



Supervised training loss

(under sparsity constraint for the optimal sampling case)

Multimodal data assimilation

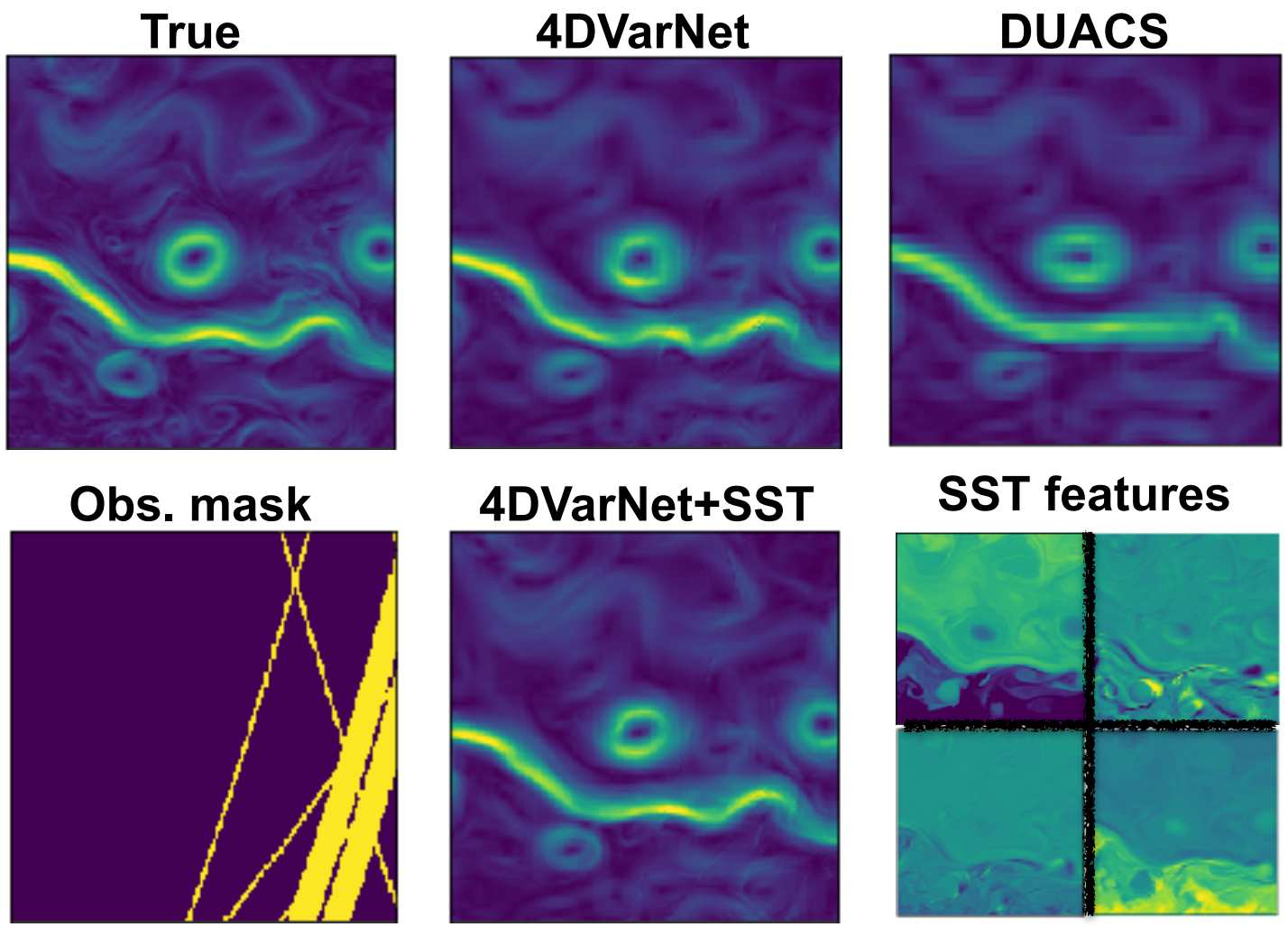
SSH-SST case-study

OSSE with NATL60 data

4-nadir-altimeter + SWOT + DUACS baseline

Gulf Stream area (10°x10°)

63% vs. 53% gain in SSH MSE w.r.t. DUACS with/ without SST (Winter period)



Optimal sampling

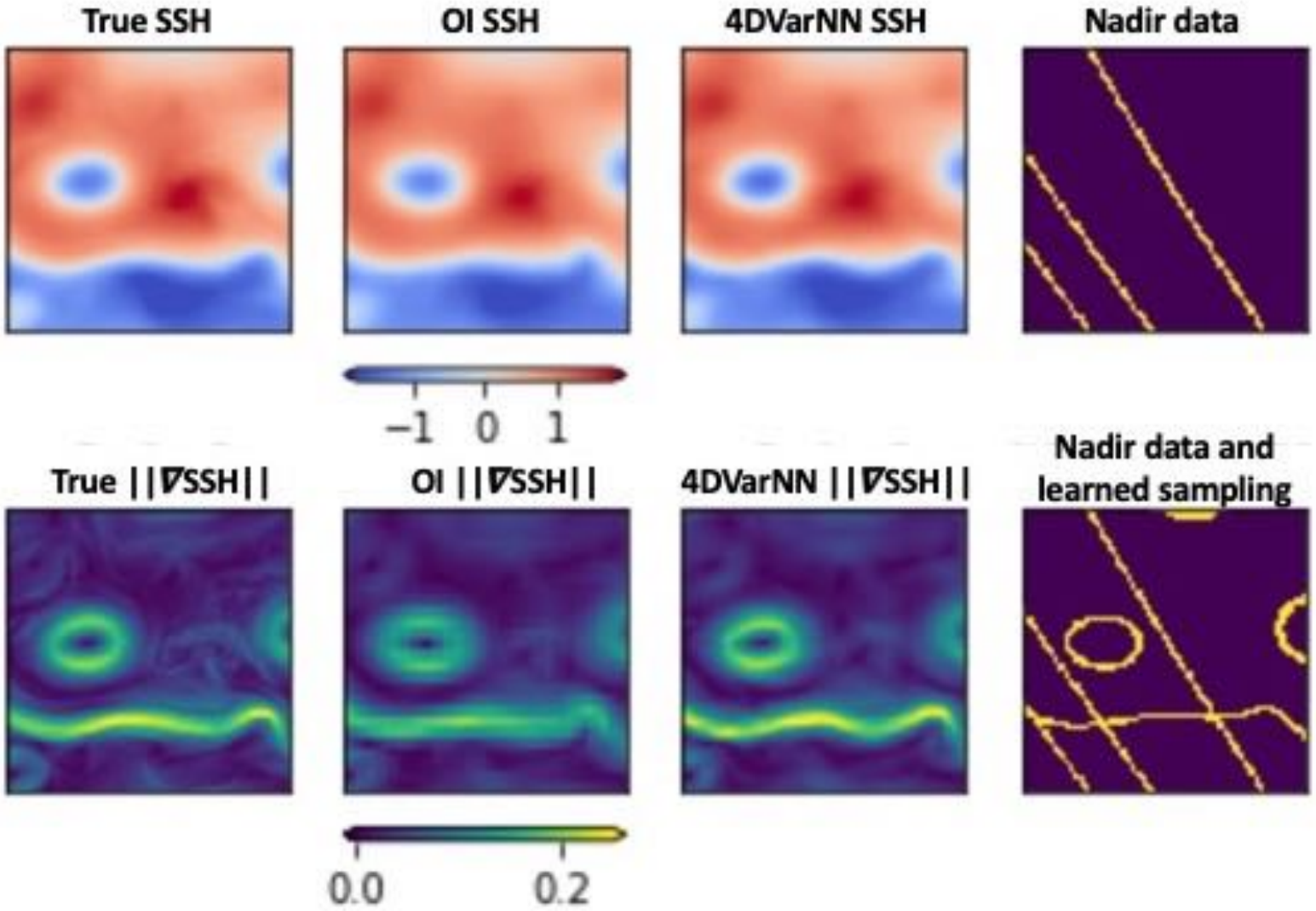
SSH case-study

OSSE with NATL60 data

Available 4-nadir SLA observation

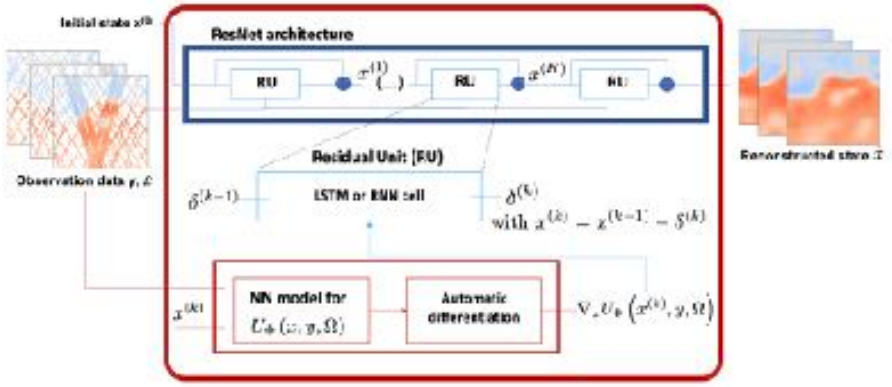
Gulf Stream area (10°x10°)

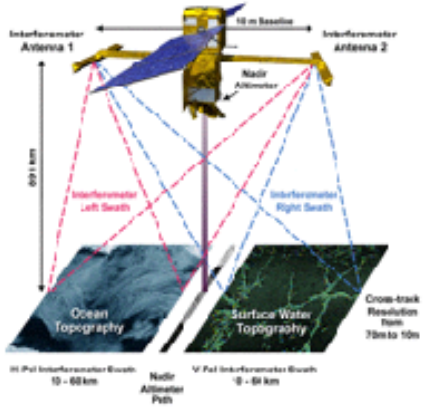
Mean relative gain of 60% in the reconstruction of the SSH using the learned sampling (~6% of the pixels vs. 1.3% for nadir altimeters)

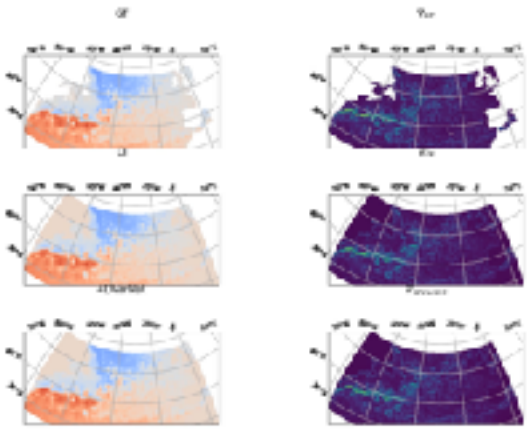


Key messages

- **Beyond NNs as data-driven-only black bloxes**
- **Trainable variational DA models** (observation model, prior, solver)
- **Application to interpolation, forecasting sampling and multimodal synergies**
- **End-to-end learning makes it easier**
- **Scaling up to the global scale**



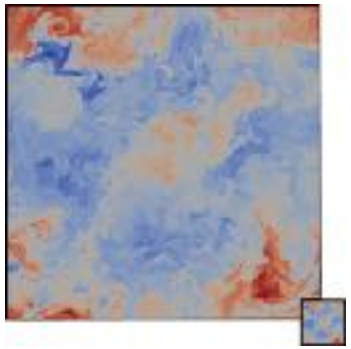




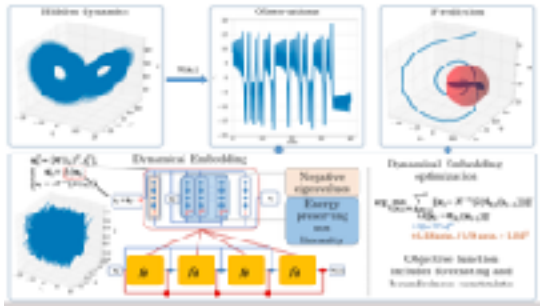
Paper (doi): [10.1029/2021MS002572](https://doi.org/10.1029/2021MS002572)
Code: https://github.com/CIA-Oceanix/4dvarnet-core

End-to-end deep and physics-informed learning and dynamical systems (cia-oceanix.github.io)

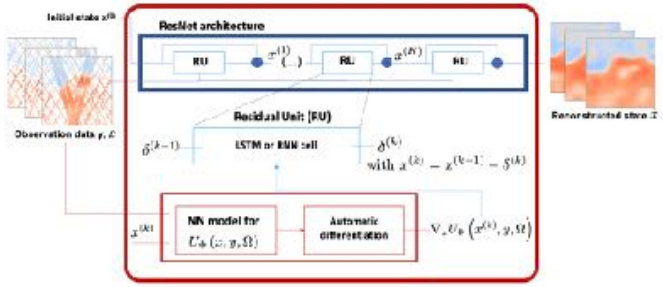
Learning & Simulation



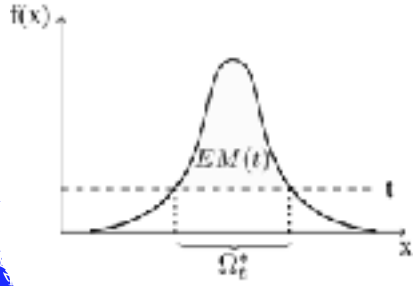
Observation-driven forecasting



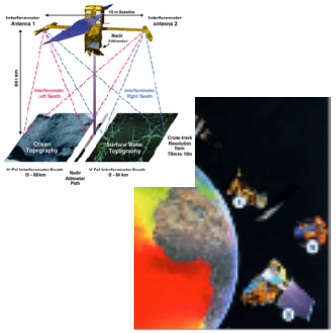
Learning & Data Assimilation

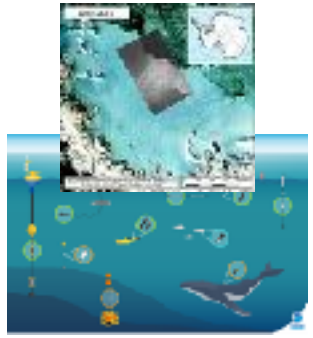


Learning for geophysical extremes

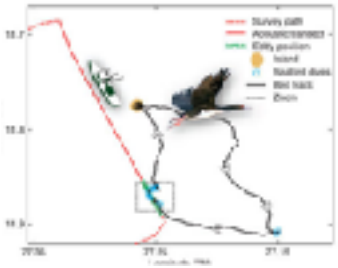


Multimodal observations





Trajectory data modelling and analysis



AI Chair OceaniX 2020-2024

Physics-informed AI for Observation-Driven Ocean AnalytiX

R. Fablet, Prof. IMT Atlantique, Brest
Web: https://cia-oceanix.github.io/

Thank you.

References

- **Model calibration/identification**
  - Frezat et al. Physics-informed neural networks for sub-grid scale modeling in filtered turbulence. PRF, 2021. https://arxiv.org/abs/2010.04663
  - Pannekoucke et al. PDE-NetGen 1.0: from symbolic PDE representations of physical processes to trainable neural network representations. GMD, 2020. https://gmd.copernicus.org/articles/13/3373/2020/
  - Ouala et al. Learning Runge-Kutta Integration Schemes for ODE Simulation and Identification, aRxiv, 2021. https://arxiv.org/abs/2105.04999

- **Data assimilation and Forecasting**
  - Ouala et al. Learning Latent Dynamics for Partially-Observed Chaotic Systems. Chaos, 2020. https://arxiv.org/abs/1907.02452
  - Fablet et al. Joint interpolation and representation learning for irregularly-sampled satellite-derived geophysical fields. FAMS, 2021. https://doi.org/10.3389/fams.2021.655224
  - Fablet et al. Learning Variational Data Assimilation Models and Solvers. JAMES, 2021. https://arxiv.org/abs/2007.12941

- **Trajectory data analysis and modeling**
  - Nguyen et al. GeoTrackNet-A Maritime Anomaly Detector using Probabilistic NN Representation of AIS Tracks and A Contrario Detection. IEEE TITS, 2020. https://arxiv.org/abs/1912.00682
  - Roy et al. Deep Learning and Trajectory Representation for the Prediction of Seabird Diving Behaviour. biorRxiv, 2020. https://www.biorxiv.org/content/10.1101/2021.04.19.438554v1
  - Nguyen et al. TrAISformer-A generative transformer for AIS trajectory prediction, arXiv, 2021. https://arxiv.org/abs/2109.03958

- **Applications to upper ocean dynamics**
  - Cazau et al. Multimodal deep learning for cetacean distribution modeling of fin whales in the western Mediterranean Sea. ML, 2021. https://link.springer.com/article/10.1007/s10994-021-06029-z
  - Benaichouche et al. Unsupervised Reconstruction of Sea Surface Currents from AIS Maritime Traffic Data Using Trainable Variational Models. RS, 2021. https://www.mdpi.com/2072-4292/13/16/3162
  - Vient et al. Data-Driven Interpolation of Sea Surface Suspended Concentrations Derived from Ocean Colour Remote Sensing Data. RS 2021. https://www.mdpi.com/2072-4292/13/17/3537
  - Beauchamp, et al. Intercomparison of data-driven and learning-based interpolations of along-track nadir and wide-swath altimetry observations. RS, 2020. https://www.mdpi.com/2072-4292/12/22/3806