Pre-training and fine-tuning neural networks to reconstruct sparse satellite fields

Théo Archambault

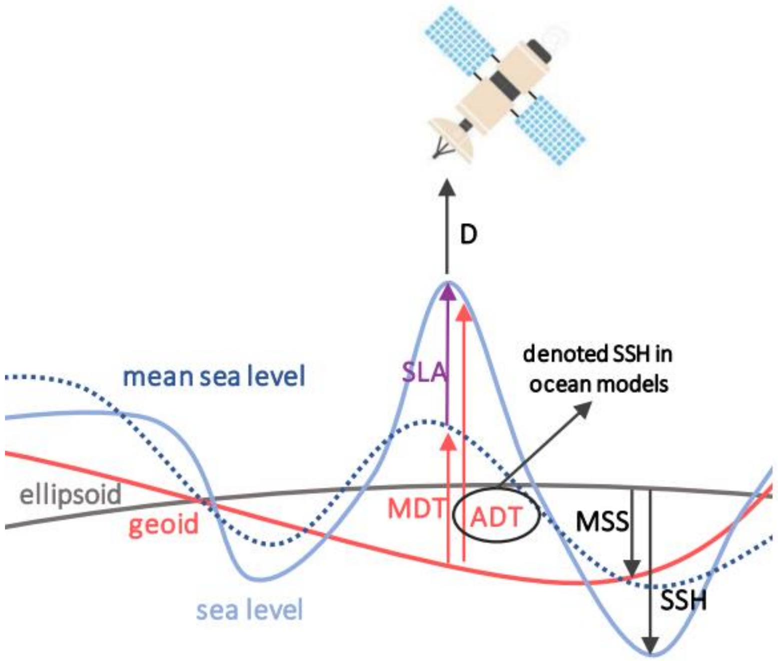
LIP6, Sorbonne University

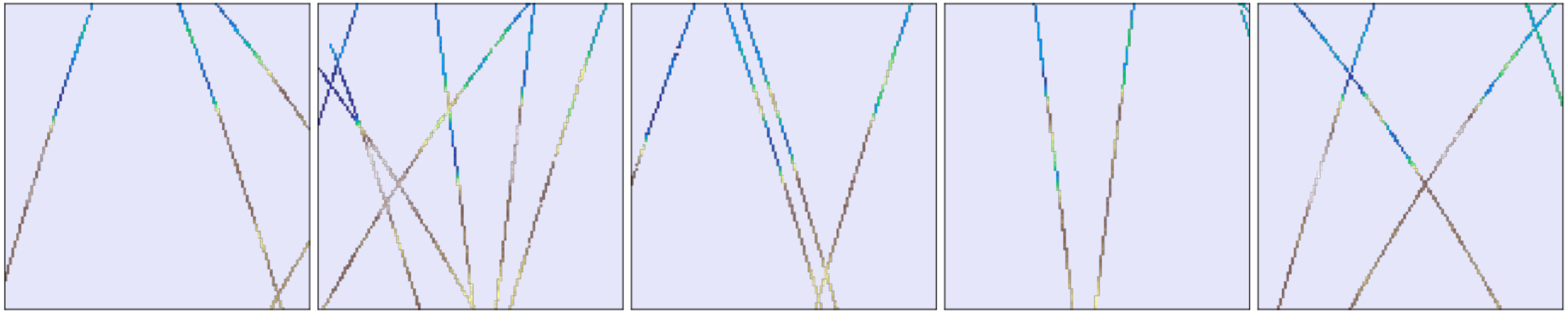
21/05/2024

Introduction

Objective: Sea Surface Height interpolation

- Used to derive surface currents to geostrophic approximation
- Measurement principle : return time of a radar pulse
- Nadir-pointing altimeters : only measure data along their ground tracks

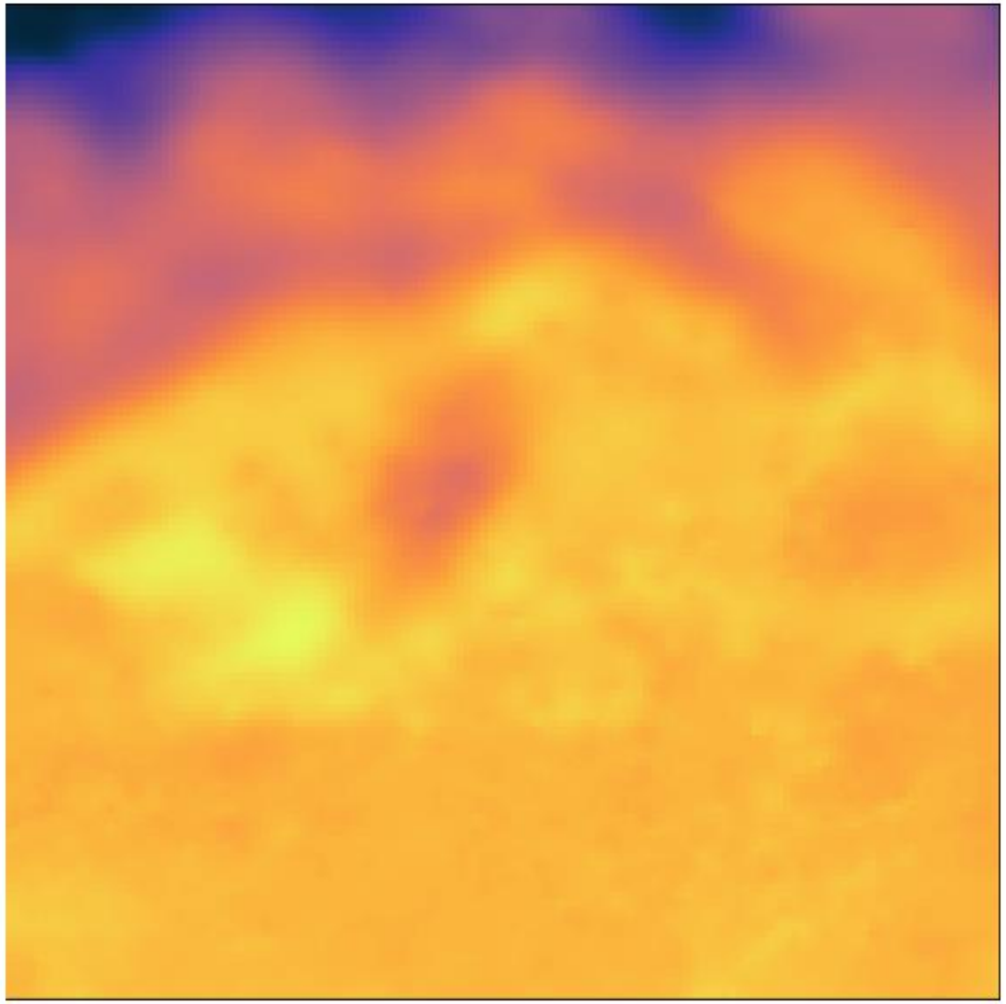




Contextual information: Sea Surface Temperature

- Measurement principle : Direct infra-red image with high resolution (1/25°)
- Clouds introduce gaps in data
- L4 SST fields : obtained through linear Optimal Interpolation combining several satellites and in-situ data.
- Noise : high-frequency instrumental errors and blurring.
- Advected by the currents

SST image

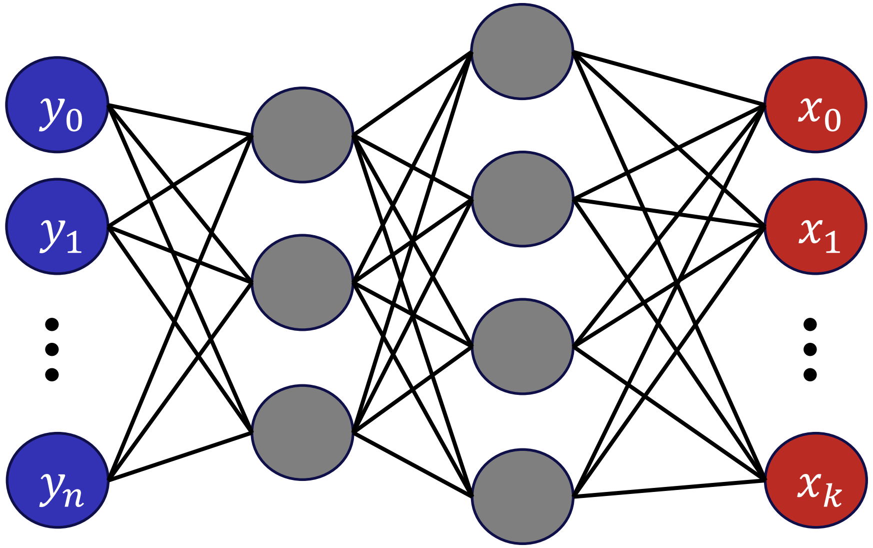


Method

Artificial Neural Network (ANN)

- Machine learning algorithm
  - Composed of layers : linear combination + unlinear activation
  - The connections ("weights") between inputs, layers and outputs are trainable parameters

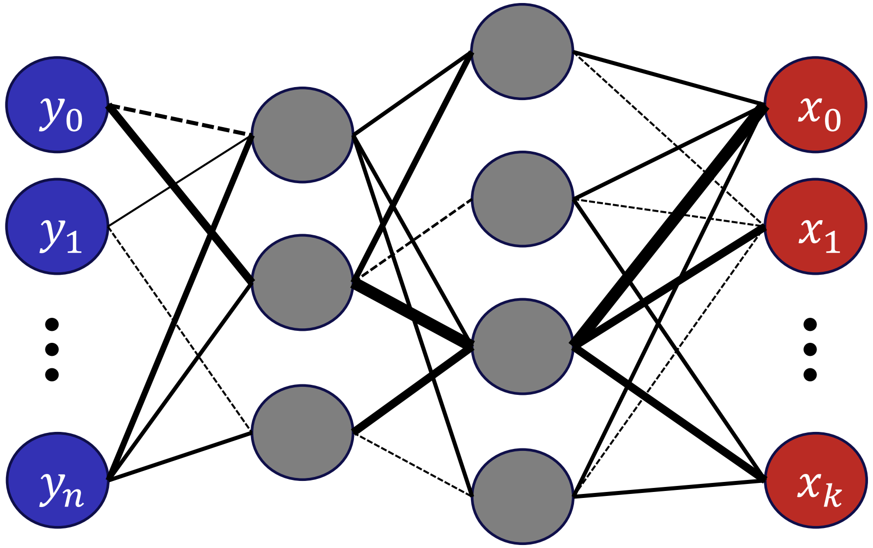
Artificial Neural Network



Artificial Neural Network (ANN)

- Machine learning algorithm
  - Composed of layers : linear combination + unlinear activation
  - The connections ("weights") between inputs, layers and outputs are trainable parameters

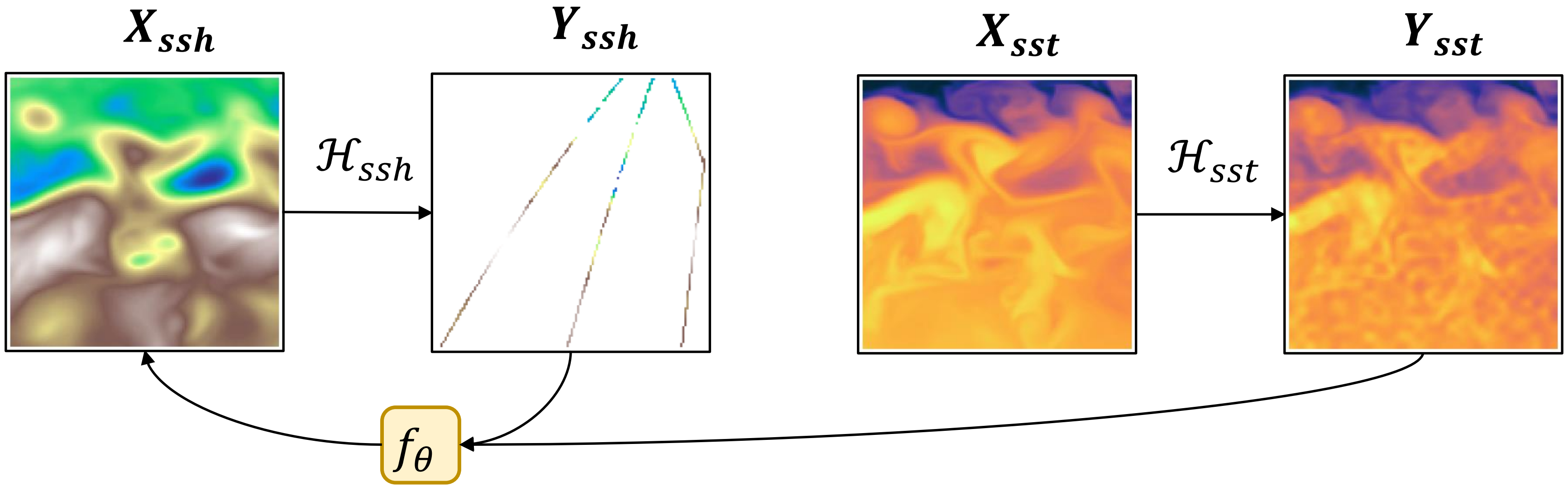
Artificial Neural Network



- The weights are trained :
  - By showing multiple examples to the neural network
  - Computing a "loss" function comparing ANN estimations and "ground truth"
  - The weights are adjusted to learn the task
  - **Supervised Learning**: requires example of **ground truth**.

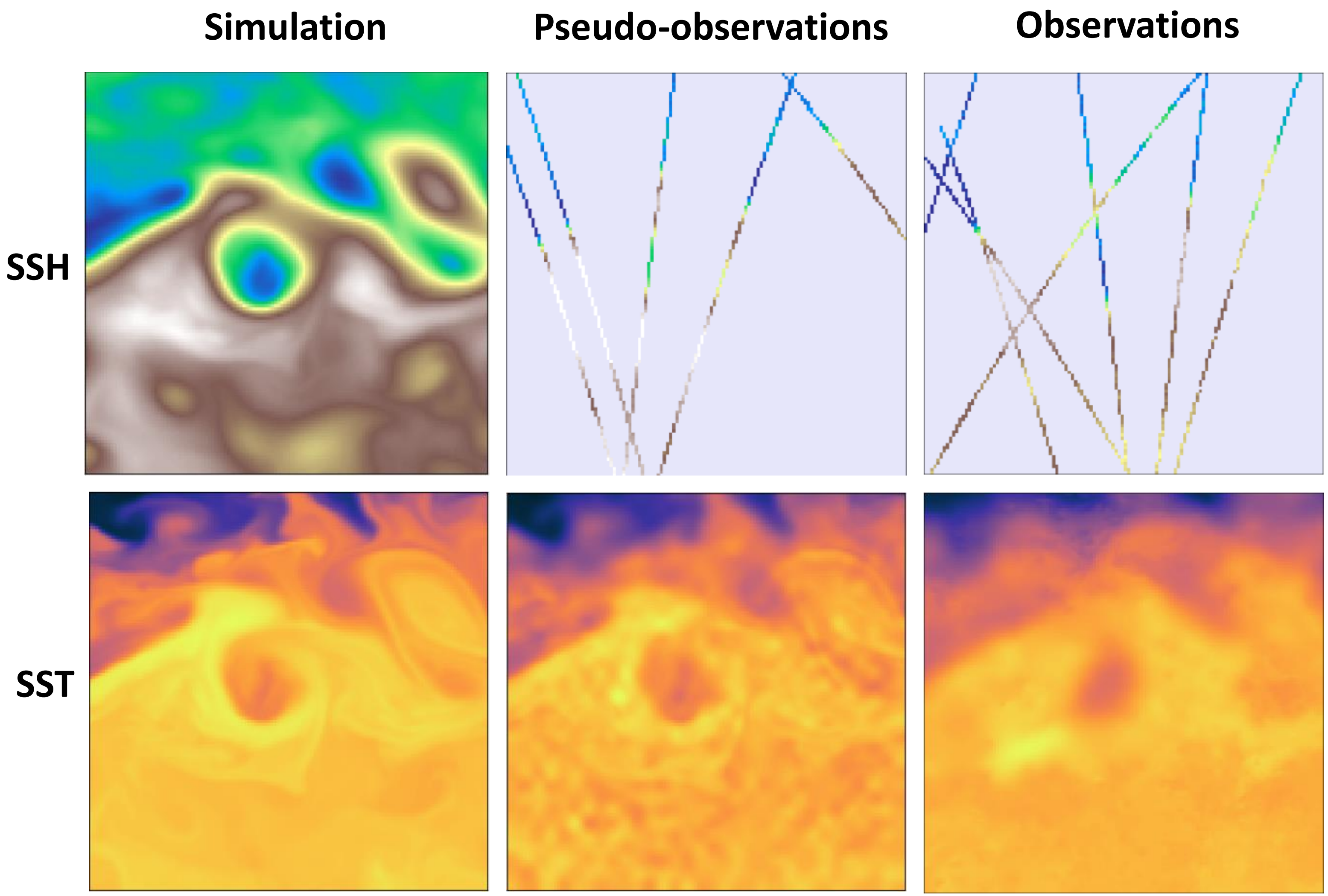
Oberving System Simulation Experiment

- Observing System Simulation Experiment (OSSE)
  - In geosciences the ground truth is not accessible
  - To train a NN, we replicate satellite observations on the outputs of a physical model (GLORYS12): $\boldsymbol{Y} = \mathcal{H}(\boldsymbol{X}) + \varepsilon$
    - $\mathcal{H}_{ssh}$: SSH along the satellite path + noise
    - $\mathcal{H}_{sst}$: SST blurred in cloudy area + noise
    - f_θ : CNN inversion of $\mathcal{H}_{ssh}$ using SSH and SST observations

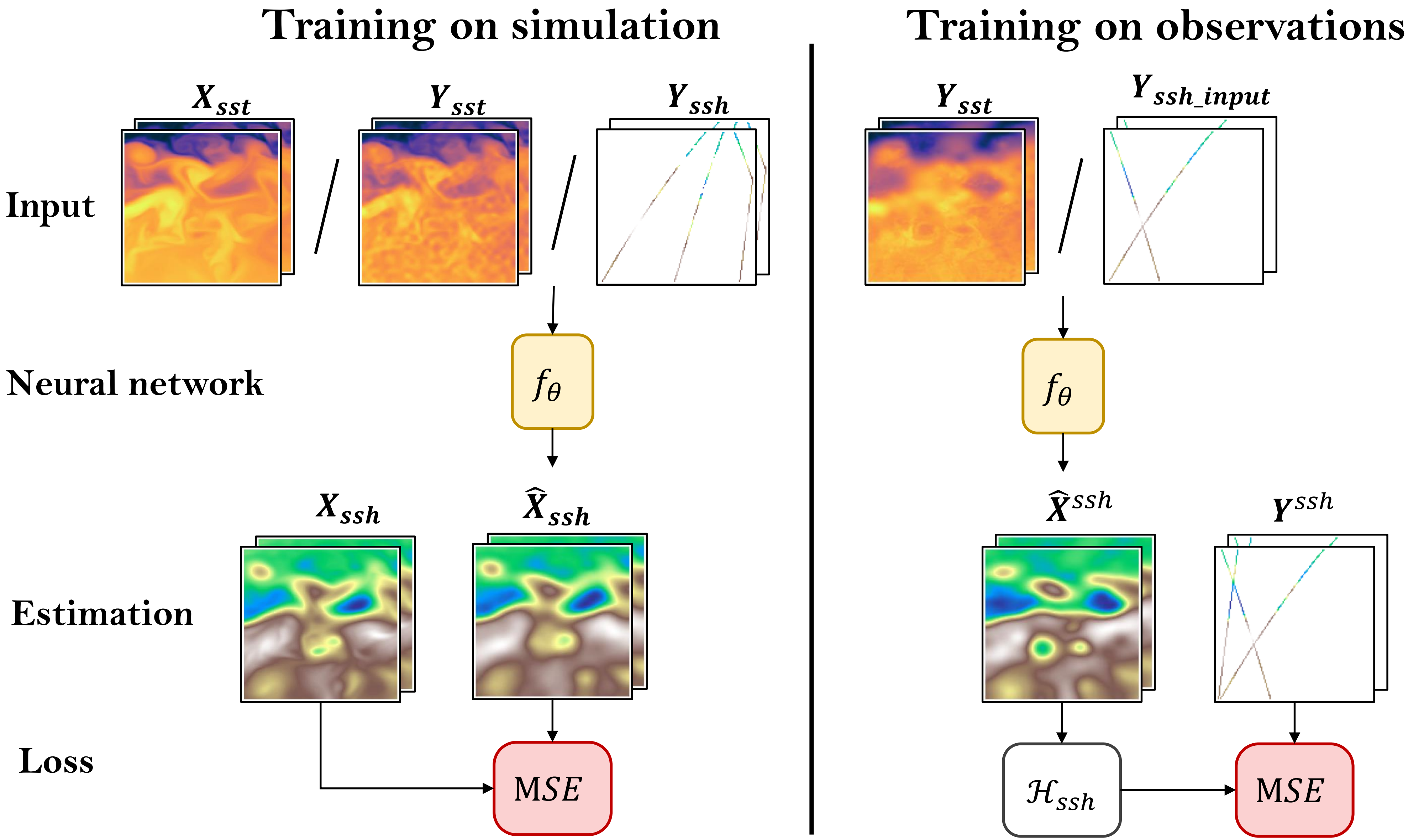


Domain Gap

- We have a 2 datasets:
  - OSSE (ground truth and pseudo observations)
  - Real-world observations
- Domain gap issue
  - Observing system is not perfectly simulated
  - The physical simulation is not perfect
- Adapt NN trained on OSSE to real data: transfer learning



Training method: pre-training and fine-tuning



Results

Real Observations dataset

Evaluation on the *Ocean Data Challenge 2021*

- 1 year of data on the Gulf Stream area
- Provides state-of-the-art reconstruction methods
- Evaluation on independent data
- Metrics : μ the RMSE score (in cm), σ_t its temporal std (in cm), λ_x the half-resolved spatial wavelength (in km)

We want to test :

- SST impact : training using SSH, SSH + nSST, SSH + SST
- The learning strategy

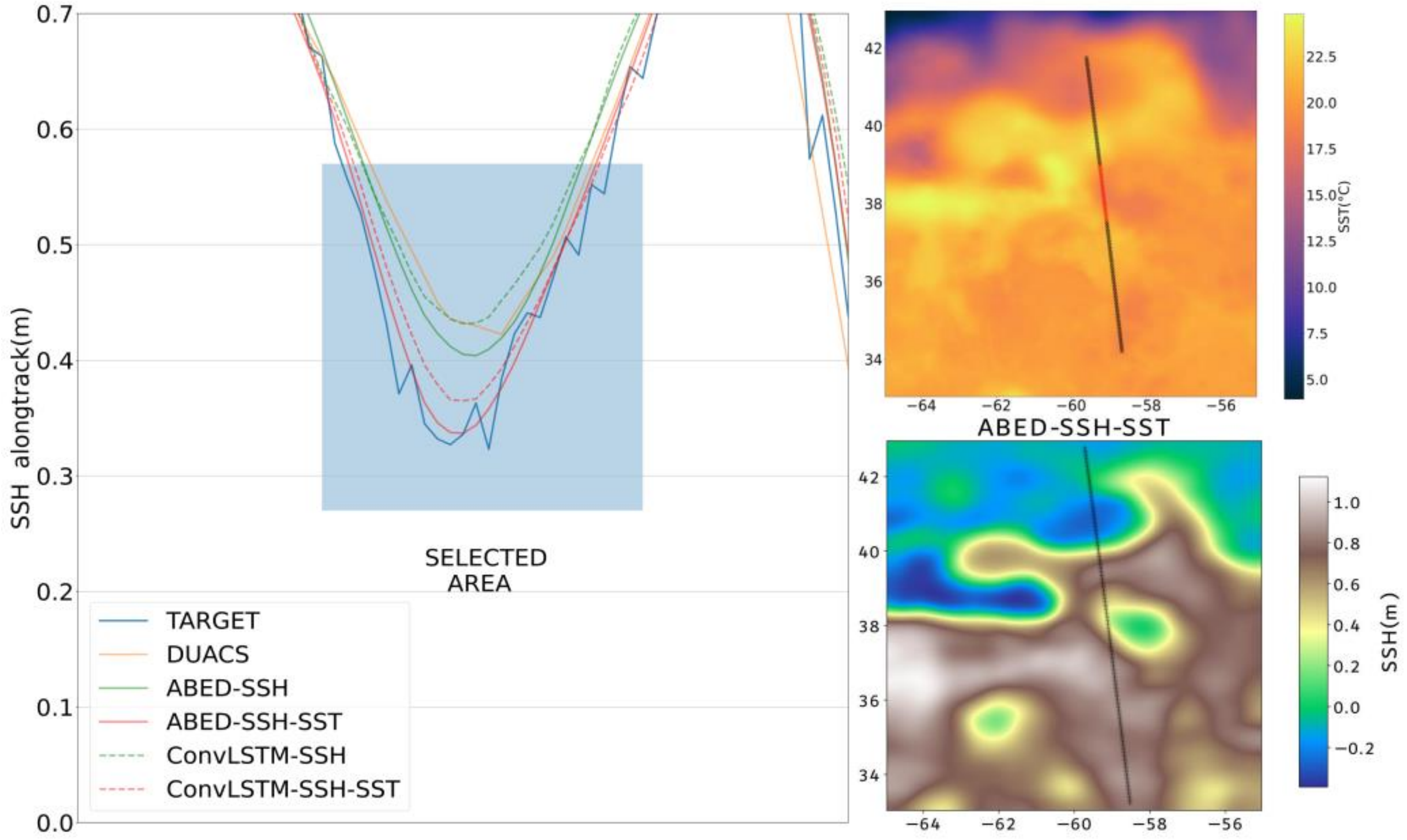
Comparing training strategies

Given the supervised and the unsupervised learning, we derive 3 strategies:

- **Observations only :** unsupervised training on real-world observations.
- **Simulation only :** supervised training on simulation and direct inference on real-world data.
- **Both :** Supervised pre-training on simulation and unsupervised fine-tuning on real-world observations.

Input data	SSH			SST+nSST			SSH+SST		
Learning method	μ	σ_t	λ_x	μ	σ_t	λ_x	μ	σ_t	λ_x
Observations	6.52	1.95	111	6.13	1.84	104	x	x	x
Simulation	6.35	1.9	112	6.2	1.87	108	6.85	2.22	111
Both	6.27	1.85	110	**5.77**	1.64	**102**	**5.77**	1.6	103

Example of improvement brought by the SST



Conclusion