A Multi-source Generative Model for Misaligned Geospatial Data

With an application to tropical cyclones

Clément Dauvilliers[1], Claire Monteleoni[1,2]

[1]INRIA, Paris, Île-de-France, France [2]University of Colorado Boulder, Boulder, CO, United States

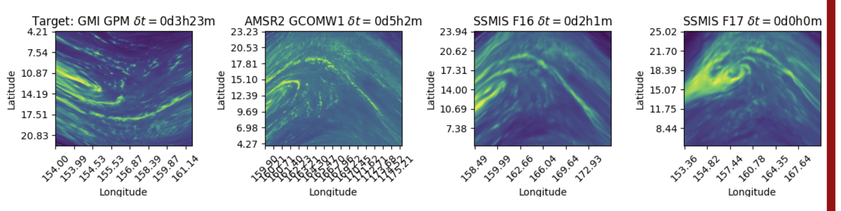
Problem

Reconstruction problem: Given input geospatial sources $S_1,\dots,S_K$, reconstruct a target source S_0. Geospatial sources are data for which each point is associated with a latitude and longitude.

Multi-source setting: we consider a setting more general than previous works, in which geospatial sources can
- Measure different variables
- Cover different geographical areas
- Occur at irregular time intervals
- Be on different (possibly irregular) grids with varying resolutions and geometries.

We present MOTIF, a framework and architecture to learn the reconstruction task in this general multi-source setting. The model can be trained either as deterministic or generative, using the same self-supervised training task.

Motivations: data assimilation, weather forecasting, cross-sensor harmonization, among others.



Example: microwave images of the same cyclone from 4 different satellites.

Multi-source training

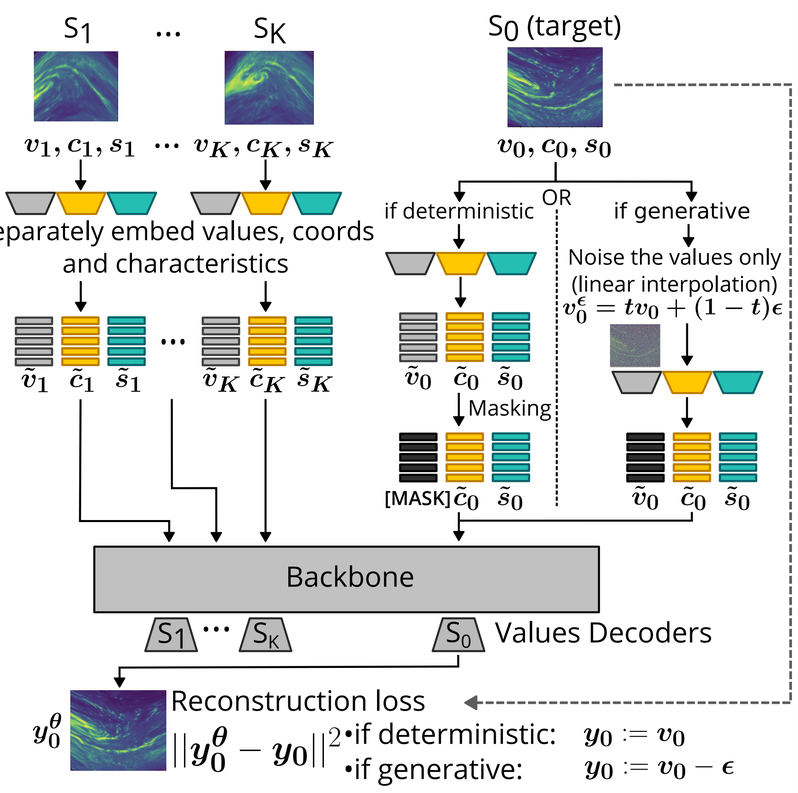
Each source S_i includes three components:

$\boldsymbol{v_i} \in \mathbb{R}^{H_i \times W_i \times C_i}$ values (e.g. pixels)

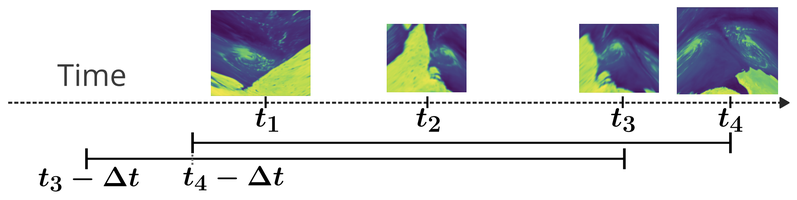
$\boldsymbol{c_i} \in \mathbb{R}^{H_i \times W_i \times 3}$ spatio-temporal coordinates (lat/lon/time) at every point

$\boldsymbol{s_i} \in \mathbb{R}^{d_i}$ source characteristics (e.g. IFOV, frequency)

Training pipeline: For a given target source, mask or noise the values only, and task the model with reconstructing them at the corresponding coordinates.



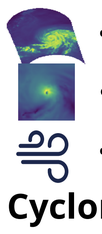
Self-supervised task: given a maximum time window Δt
- Select a source S observed at time t
- Gather all sources available within $t - \Delta t$
- Randomly choose one of the sources as target



Experiments and Results

Dataset

TC-PRIMED dataset (Razin et al. 2023), covering 3,500 tropical storms from 1987 to 2024.

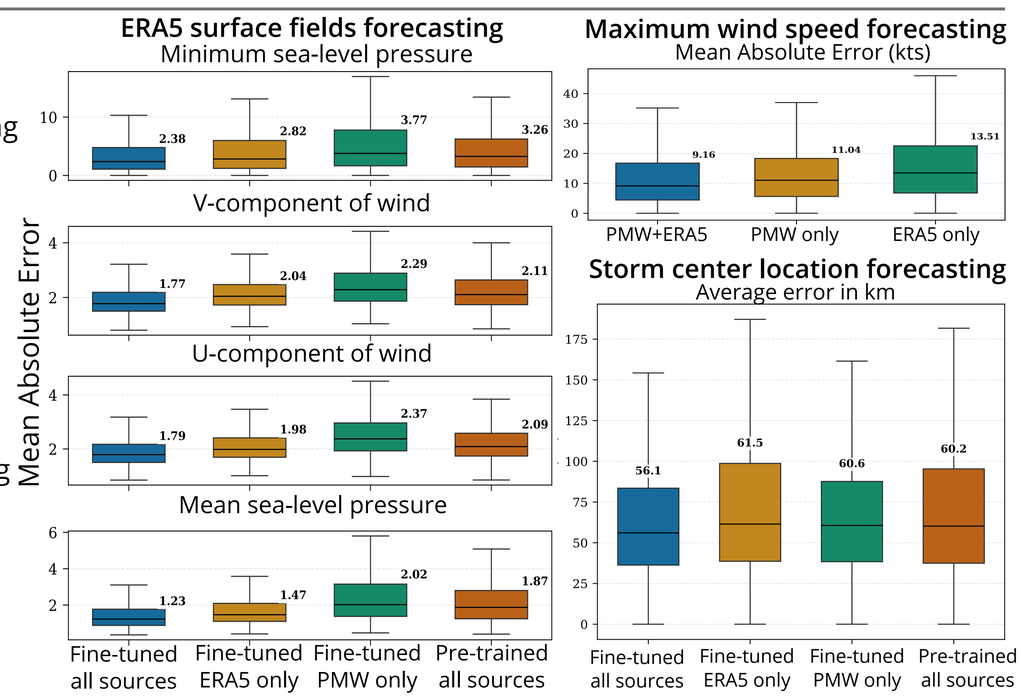


- **PMW**: Passive Microwave images from 11 satellites with different characteristics (frequency, geometry, resolution...)
- **ERA5:** patches around the cyclone's center, as boxes of 20°x20° at a resolution of 0.25° every 6 hours.
- Best-track estimates of the storms' 1-minute Maximum Sustained Wind speeds (MSW) every 6 hours.

Cyclones forecasting

Comparisons (with fixed architecture):
- Pre-trained on ERA5 and PMW sources using the self-supervised reconstruction task.
- Fine-tuned on +24h forecasting, using only PMW sources as input.
- Fine-tuned using only ERA5 as input.
- Fine-tuned using PMW and ERA5 as input.

Conclusions
- ERA5 + PMW input > ERA5 or PMW only.
- Unsurprisingly, ERA5 > PMW for forecasting ERA5.
- PMW > ERA5 for intensity and location forecasting.



Cross-sensor harmonization

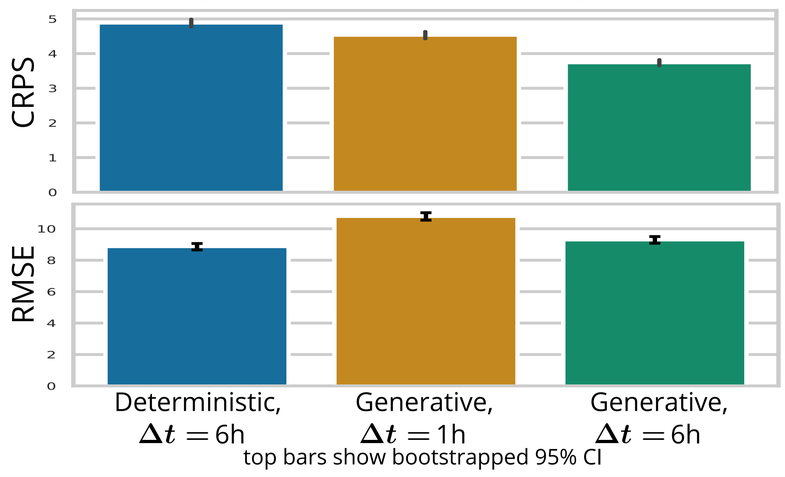
Task: (A) Satellites within the GPM constellation — Domain adaptation → (B) GPM Core observatory

Going beyond co-located images

- (A) ~~Supervised training~~ (B) — Not enough co-timed and co-located samples.
- Instead: (A & B) — Self-supervise training using all sources as targets or inputs.

We benefit from training on images that are further apart in time

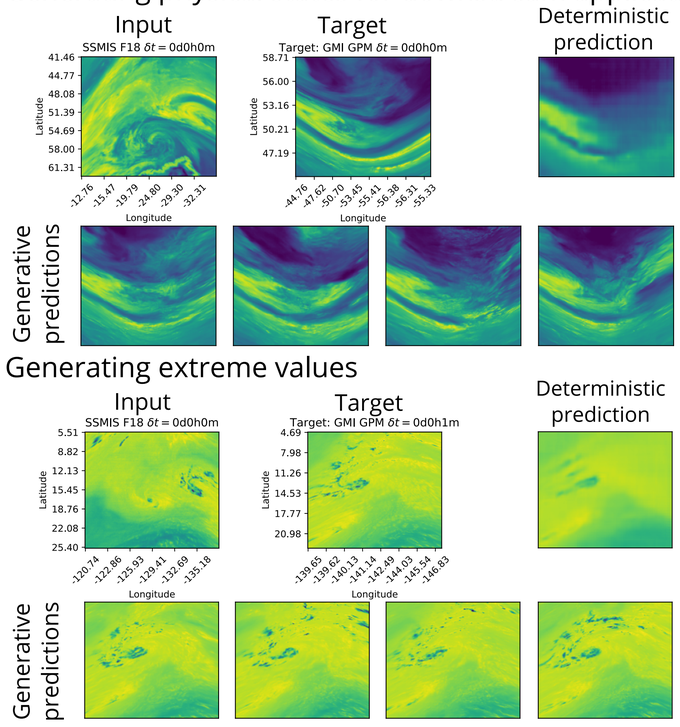
Impact of the maximum time between sources during training, evaluated on co-timed images



top bars show bootstrapped 95% CI

Why use a generative model ?

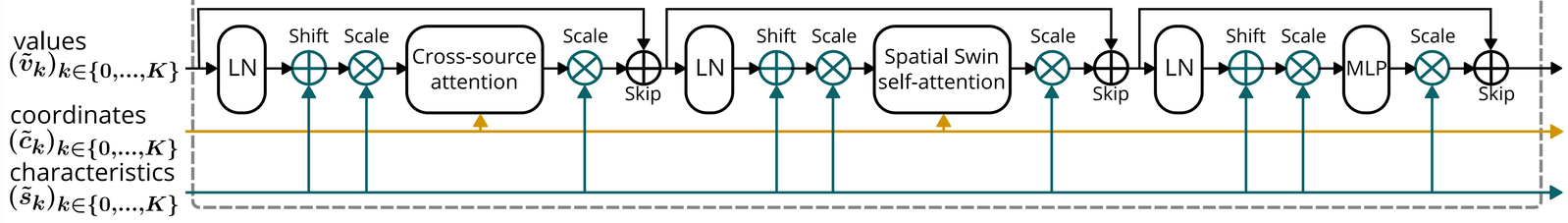
- Resolving small-scale features
- Generating physical states for downstream applications



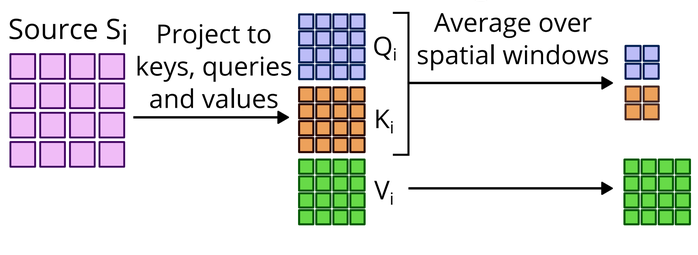
- Generating extreme values

Architecture

Backbone



Cross-source attention for misaglined sources



- Apply to all sources
- Concatenate the queries, keys and values of all sources
- Compute the attention weights and updated values following:

$$V' = \mathrm{Softmax}\left(\frac{Q_V (K_V)^T}{\sqrt{d_V}} + \frac{Q_C (K_C)^T}{\sqrt{d_C}}\right) V$$

Values attention — Coordinates attention

- Split back V' into the sources

Ongoing & Future work

Two main directions to explore:
- As a multi-source generative forecasting model for tropical cyclones
  - Using ERA5, satellite imagery and aircraft data as input
  - Predicting ERA5 and real intensity, contrary to data-driven GCMs (e.g. GenCast, Aurora).
- As a general multi-source domain adaptation method for remote sensing
  - Requires multiple, more general datasets than used here (suggestions are welcome !).

References

Razin, M. N., C. J. Slocum, J. A. Knaff, P. J. Brown, and M. M. Bell, 2023: Tropical Cyclone Precipitation, Infrared, Microwave, and Environmental Dataset (TC PRIMED). Bull. Amer. Meteor. Soc., 104, E1980–E1998, https://doi.org/10.1175/BAMS-D-21-0052.1.

Dauvilliers C, Monteleoni C. MoTiF: a self-supervised model for multi-source forecasting with application to tropical cyclones. Environmental Data Science. 2025;4:e36. doi:10.1017/eds.2025.10014