Explainable and physics-informed AI for regional weather prediction

Laure RAYNAUD[1] • Laurent RISSER[2] • Valérie MASSON[1] • Luciano DROZDA[3] • Christophe BOVALO[4] • Eliott LUMET[1] • Younes ESSAFOURI[2]
Julia GARCÍA CRISTÓBAL[1] • Jean WURTZ[1] • Adrien AUDREN[14] • Mikaël JACQUEMONT[4]
[1]Météo-France/CNRS, CNRM, Toulouse, France [3]CERFACS/CNRS, CECI, Toulouse, France
[2]Université de Toulouse/CNRS, IMT, Toulouse, France [4]Eviden, France

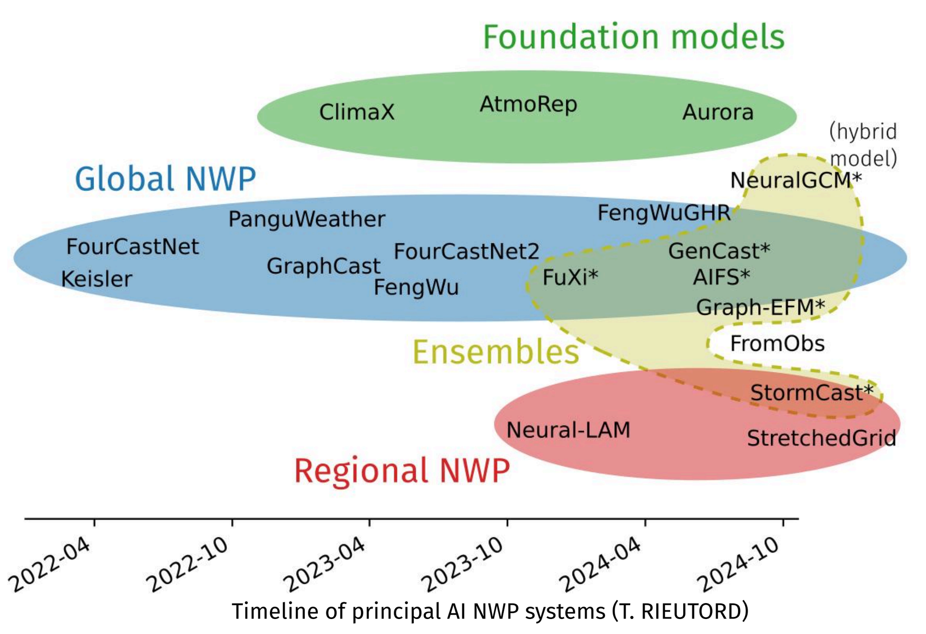
ANITI
Communauté d'universités et établissements de Toulouse

Context

Weather prediction plays a crucial role in many aspects of modern society, from agriculture and disaster management to aviation and renewable energy planning. **Traditional numerical weather prediction** (NWP) relies on solving **complex physical equations** that describe atmospheric processes. However, these models are **computationally expensive** and often struggle with small-scale phenomena and long-term forecasts.

Recent advances in **artificial intelligence (AI)** offer promising alternatives or complements to traditional NWP methods. **By learning directly from large datasets** of historical weather observations and simulations, AI-based NWP models can produce **faster and potentially more accurate forecasts**.

Beyond these potential methodological and technological breakthroughs, the **ANITI ExplEarth** chair aims to qualify the **robustness** and **explainability** of AI-based NWP models. This includes exploring how standard **Explainable AI** (XAI) techniques can shed light on the inner workings of these models, revealing how they encode physical relationships—insights that are essential for building trust and guiding further scientific improvements. The integration of **physical constraints** into the construction and training of AI NWP models will be another way to ensure their quality and strengthen user confidence in these new tools..

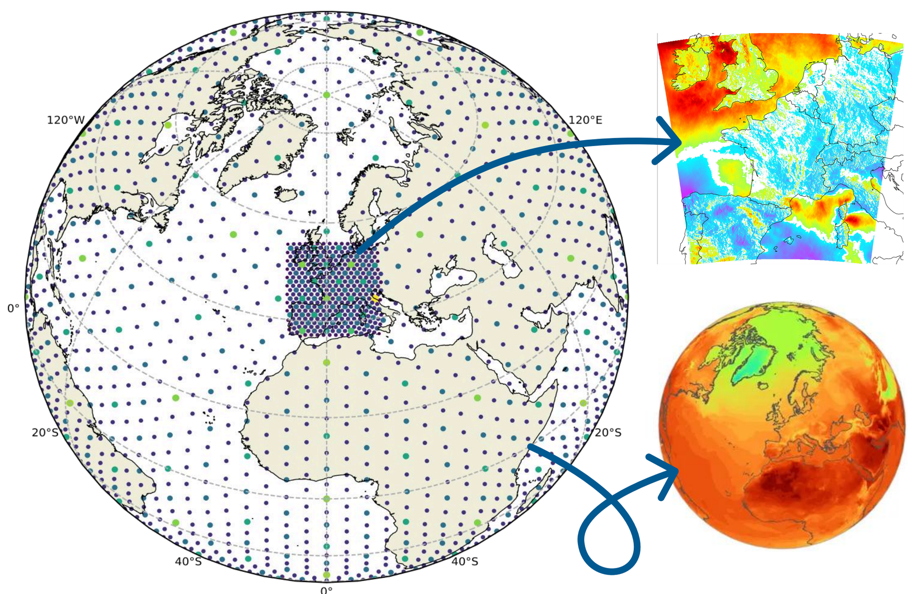


Timeline of principal AI NWP systems (T. RIEUTORD)

Diffusion downscaling

Aim: Train a diffusion model for downscaling NWP forecasts
i.e. that learns the relationship between a **global low-resolution** physical model and a **regional high-resolution** physical model.

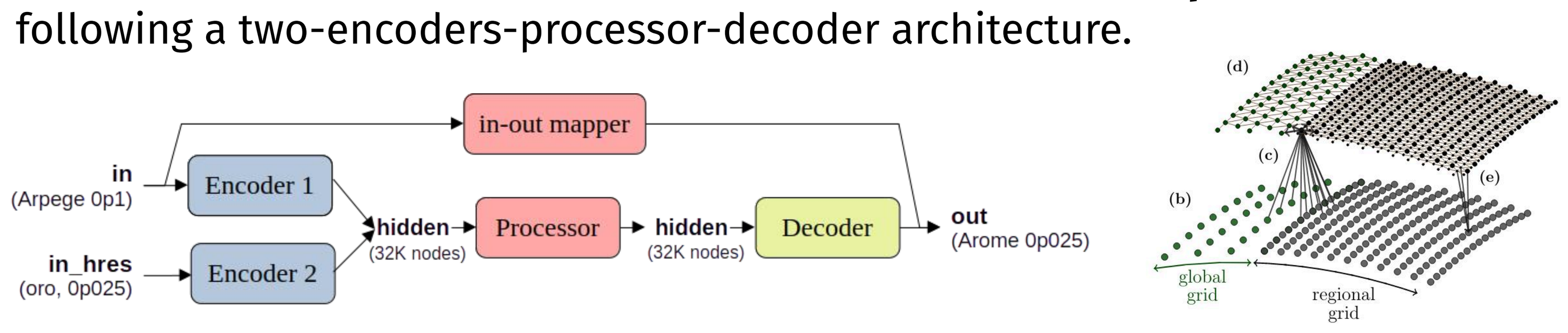
TITAN data (Training Inputs & Targets from Arome for Neural networks)
Hourly **global** (Arpege) and **regional** (Arome) **analysis** from Metéo-France archive over the 2020-2025 period.



AROME analysis
Resolution: 2.5 km
Domain: western Europe
Atmospheric variables:
- 2t, 2r, 10u, 10v, msl + tp
- t, r, u, v, z at 5 pressure levels

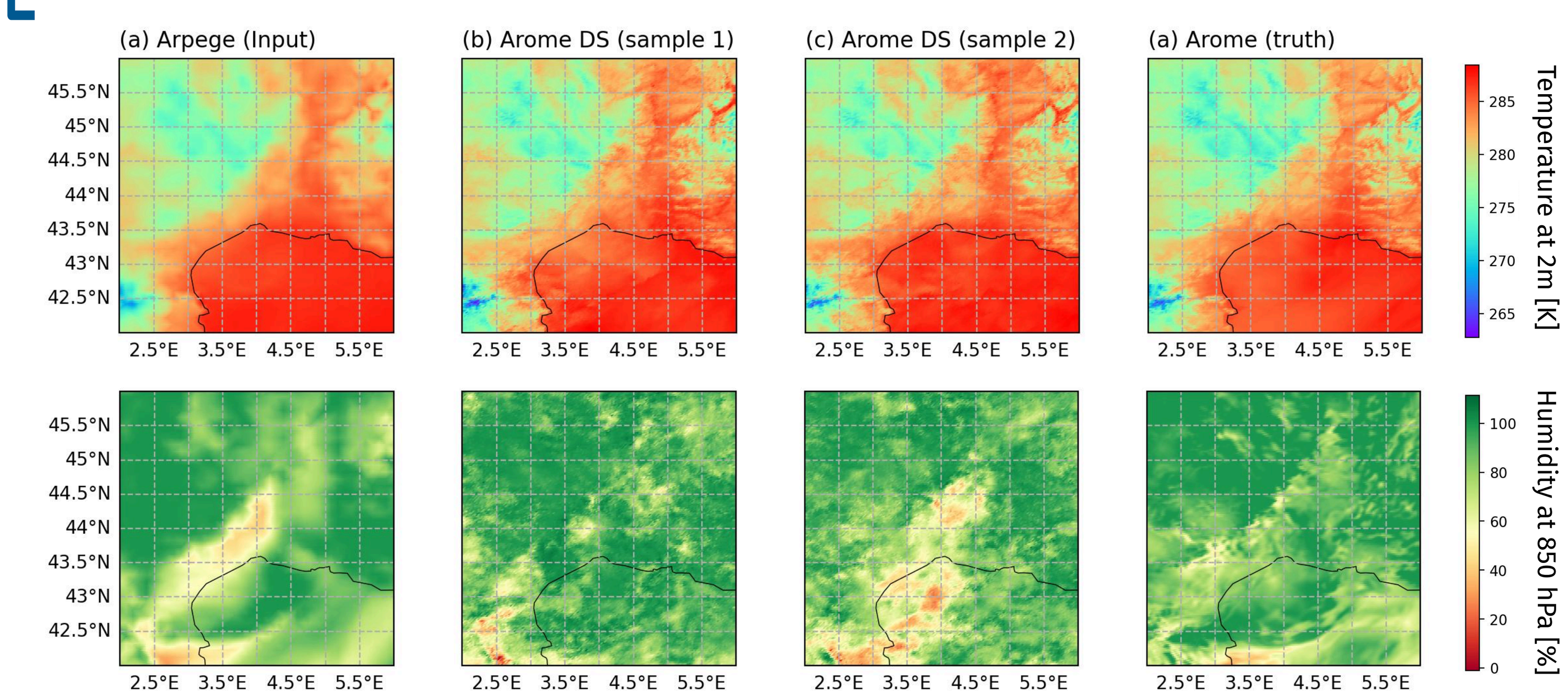
ARPEGE analysis
Resolution: 10 km
Domain: Global cropped over western Europe
Atmospheric variables:
- 2t, 2r, 10u, 10v, msl
- t, r, u, v, z at 5 pressure levels

Model:
The model is a **Generative Diffusion** model made of **GraphTransformers** following a two-encoders-processor-decoder architecture.



From Nipen et al. 2024

First results: Downscaled field samples over France South-East



→ Addresses the **fine-scale blurring** typically observed in deterministic AI NWP models
→ Produces **meaningful forecast ensembles**, enabling robust uncertainty quantification

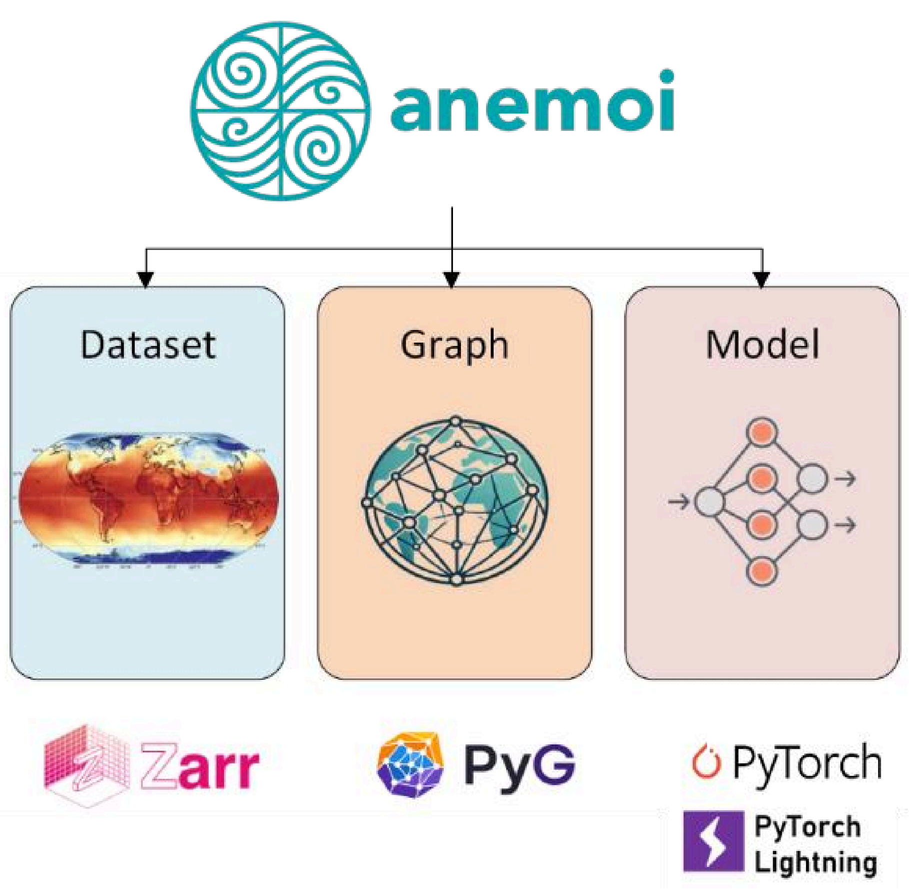
Perspectives:
- Evaluate the ability of the model of learning large-scale transformation (downscaling ≠ super-resolution)
- Investigate the temporal consistency of downscaled samples
- Apply downscaling to predict urban temperature at hectometric scale

Anemoi Framework

Ultimately each project of the chair will rely on **Anemoi framework**, an open source Python code specifically developed for the training and industrialization of AI-NWP models.

Initiated by ECMWF, it is now used by several **meteorological institutes** to develop:

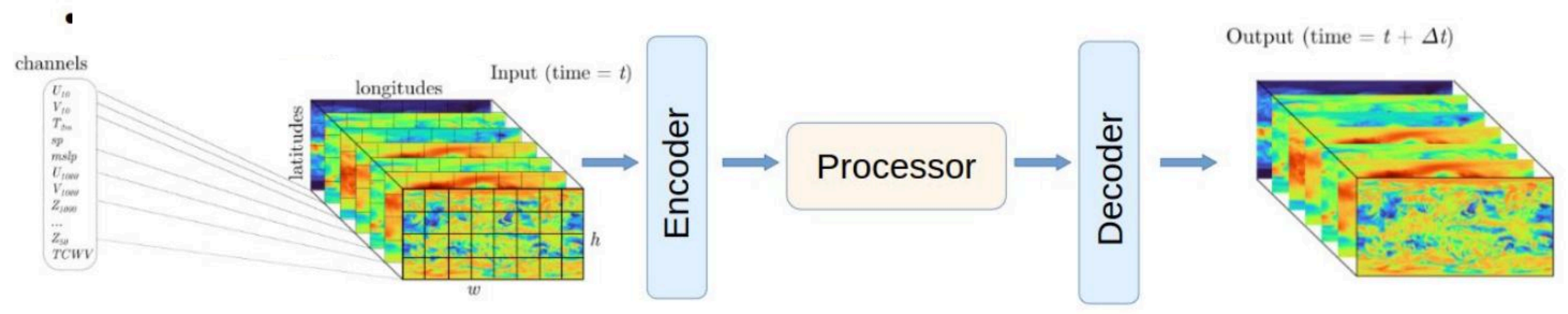
- Global forecast models (e.g. AIFS)
- Regional forecast models
- Coupled atmospheric and ocean models
- Downscaling models
- Ensemble/probabilistic models
- Obs-to-obs model (e.g. GraphDOP)



Regional NWP model emulation

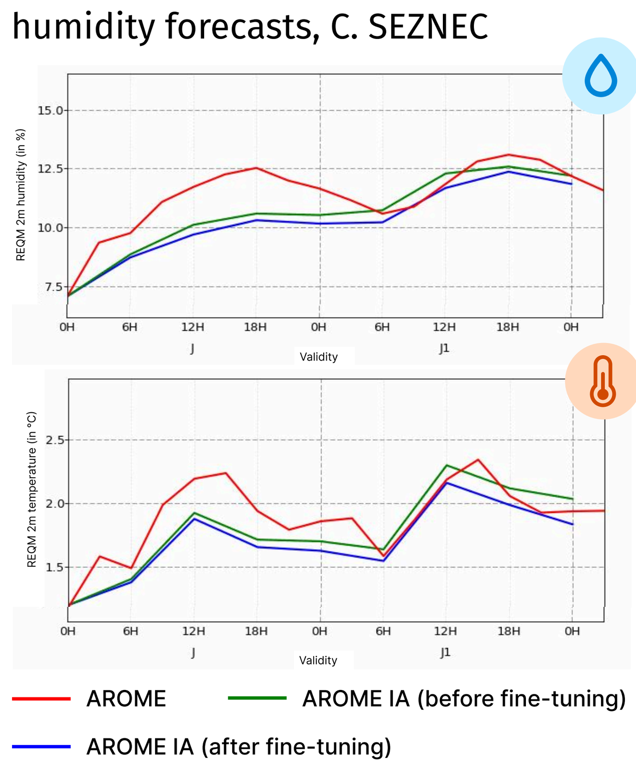
Aim: develop the first **AI-based emulator** of Météo-France's regional model
Conversely to the downscaling task, this involves learning the dynamics of the physical system, i.e. its evolution in time.

Model:
The envisioned model is made of **GraphTransformers** operating on a stretched global grid, with specific kilo-metric refinement over Arome's domain. The ECMWF's AI NWP global model (Lang et al. 2024) is **fine-tuned** on TITAN data.

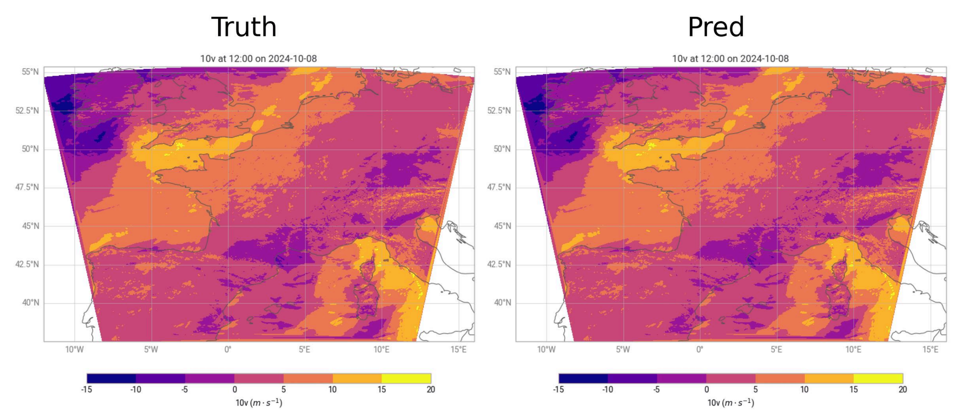


Results:

(a) **RMSE performances from June 3rd 2024 to June 30th, 2024:** comparison of AROME and AROME IA for t2m and humidity forecasts, C. SEZNEC



(b) **Meridional wind velocity at 10m(10v) forecast at 12:00 October 8th:** Comparison of AROME test field and AROME IA forecast, B. DEVILLERS



→ Produces **spatially- and temporally- consistent** forecasts
→ Matches and even **outperforms** AROME accuracy for some variables/lead times

Perspectives:
- Reduce fine-tuning time and cost using low-rank adaptation (LoRA)
- Compare the downscaling and emulation approaches
- Develop an ensemble-based alternative for uncertainty quantification

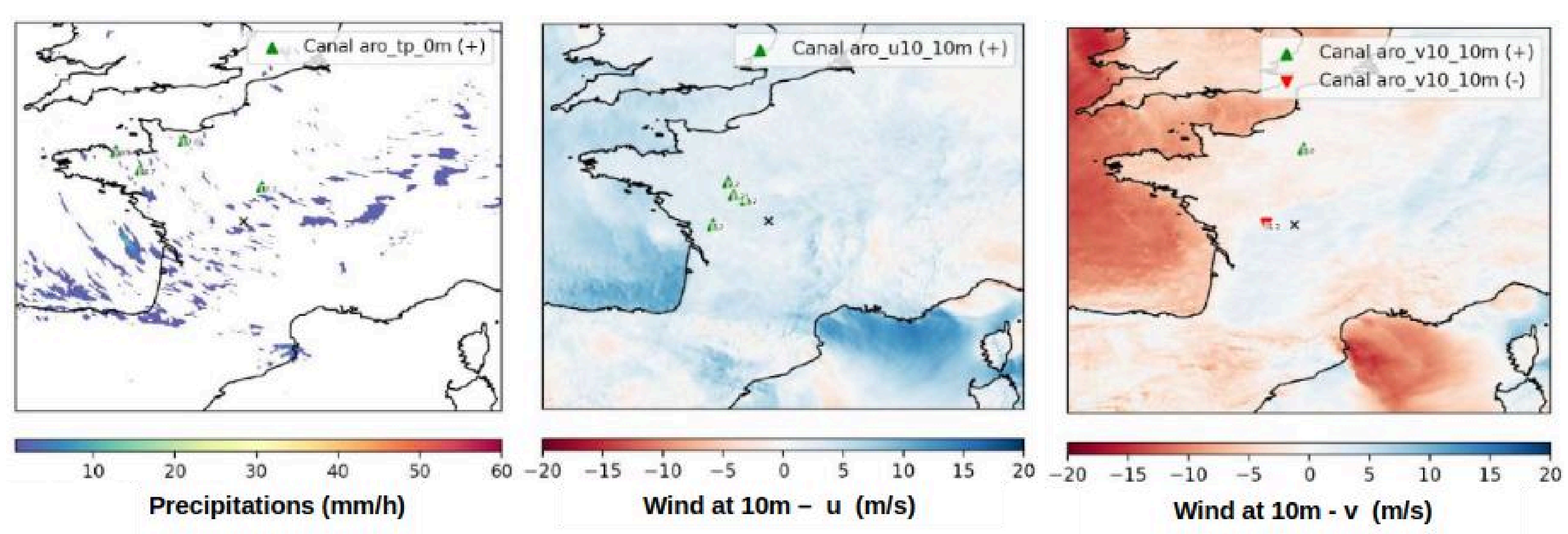
Explanation of AI-based weather predictions

Aim: Apply standard **explainable AI** (XAI) methods to AI-based NWP models, adapting them to the specificities of meteorological data, in order to better understand AI forecasts

Model:
The XAI method is applied retrospectively to a UNETR++ regional forecasting model trained on TITAN data.

Challenges:
- Very high dimensionality of the input data
- Standard XAI models designed for RGB images but not for physically-relevant data
- Interpretability of "technical" explanations to different categories of users

First results: Detection of the most impactful area/channels for local prediction



Perspectives:
- Apply the same method to more recent graph AI-NWP models developed in Anemoi
- Develop strategies for specific inputs to improve forecasts
- Detect AI-NWP model vulnerabilities