ANITI

Communauté d'universités et établissements de Toulouse



Explaining IA-based digital twins for regional weather forecasting

Laurent Risser (CNRS, Institut de Mathématiques de Toulouse, ANITI)

Journée Digital Twins : Usages, technologies et apports pour la supervision et la conduite de systèmes physiques (16/10/26)

ExplEarth chair of ANITI AI cluster



Laure Raynaud
(CNRM/Météo-France)



Laurent Risser
(IMT/CNRS)



Luciano Drozda
(CERFACS)



Maxime Baba
(IMT/CERFACS/UT/CNRS)



Younes Essafouri
(IMT/UT/CNRS)

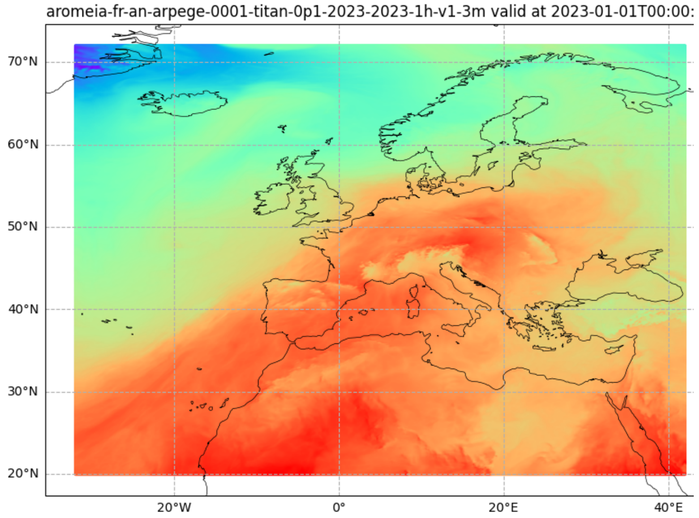


Eliott Lumet
(CNRM/Météo-France)

« IA-cluster » Chair founded for 4 years by « France 2030 » program with two main goals:

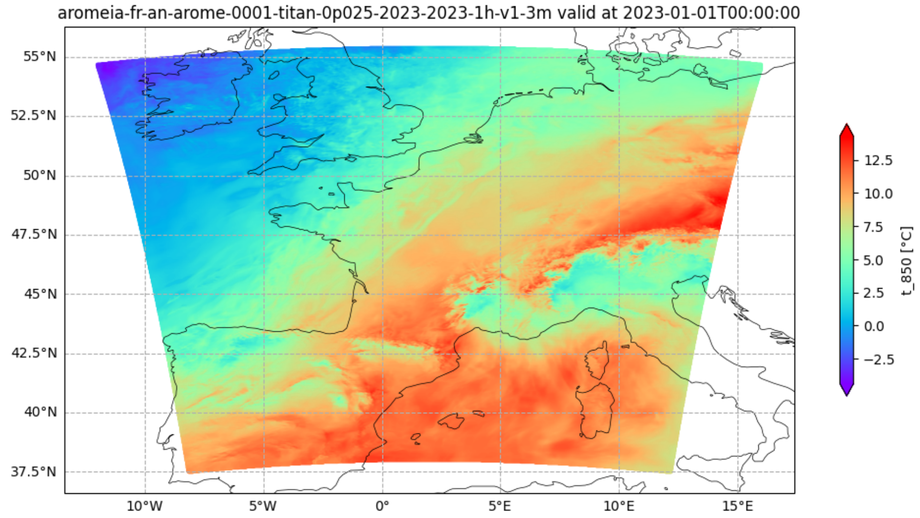
- Developing AI-based models for regional weather forecasting
- Developing explainability methods for AI-based weather forecasting

Example:



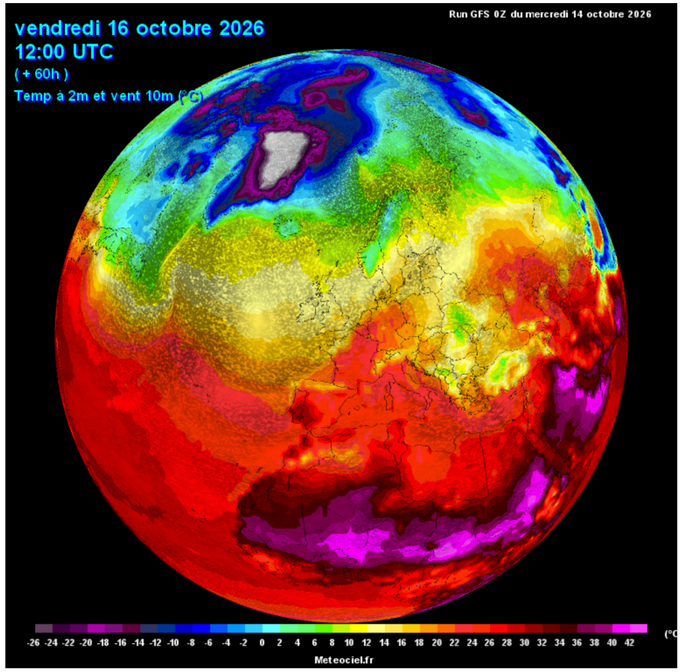
Input: Low resolution global forecasting (5km - Arpege)

Arome downscaling

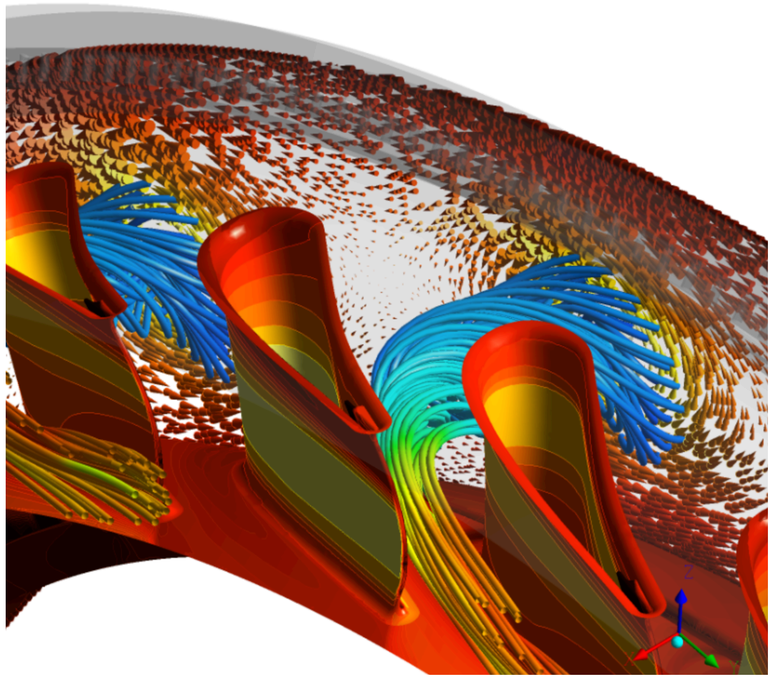


Output: High resolution regional forecasting (1.3km - Arome)

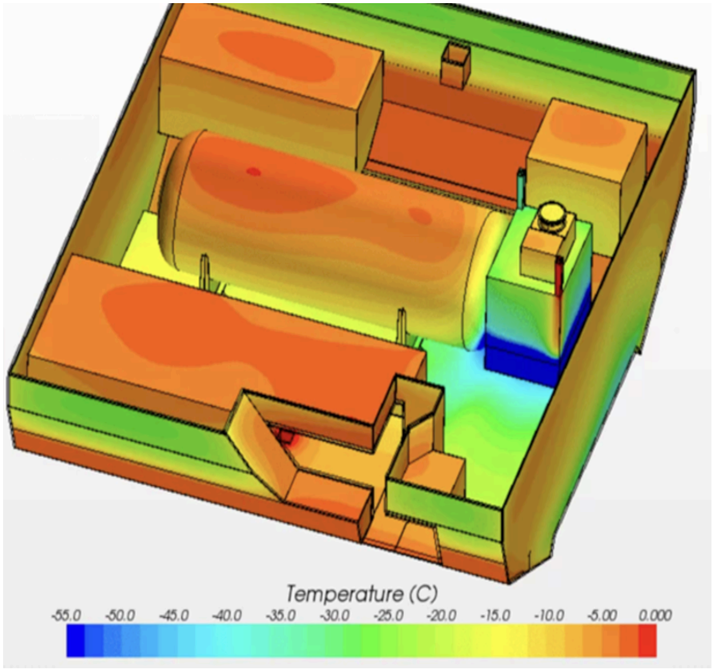
1: Digital twins and simulation



Weather forecasting (Meteociel.fr)



Flux in a turbomachine (FG GLA)



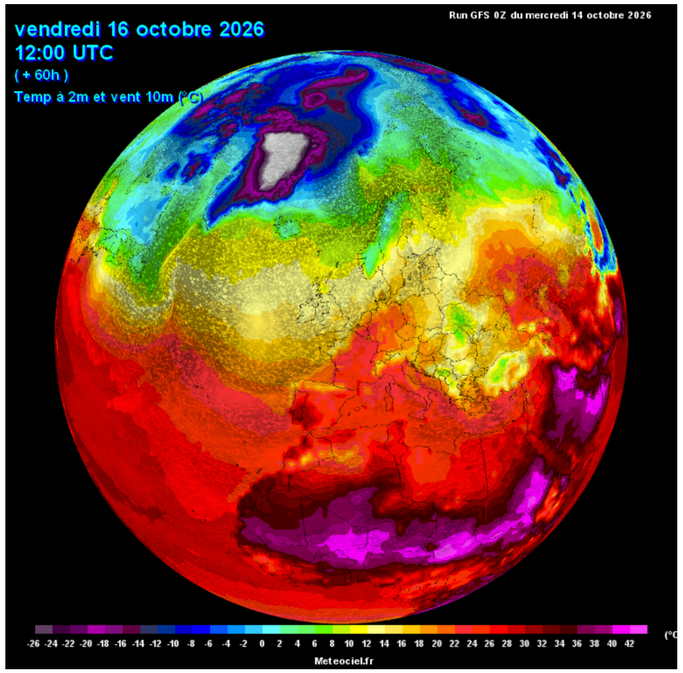
Heat propagation (Creaform)

...

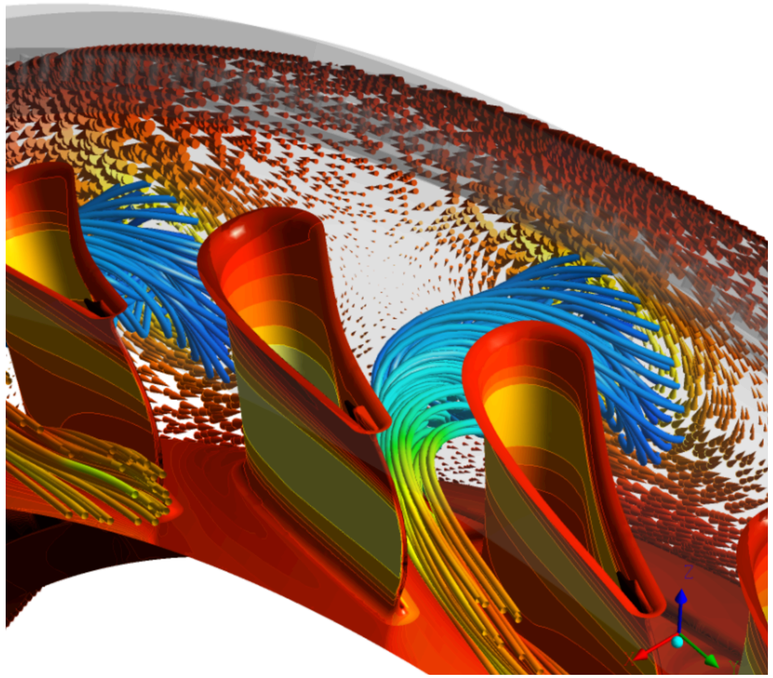
Classic numerical simulation

→ Use of mathematical models on computers to predict the state of a system in the future or in a particular situation

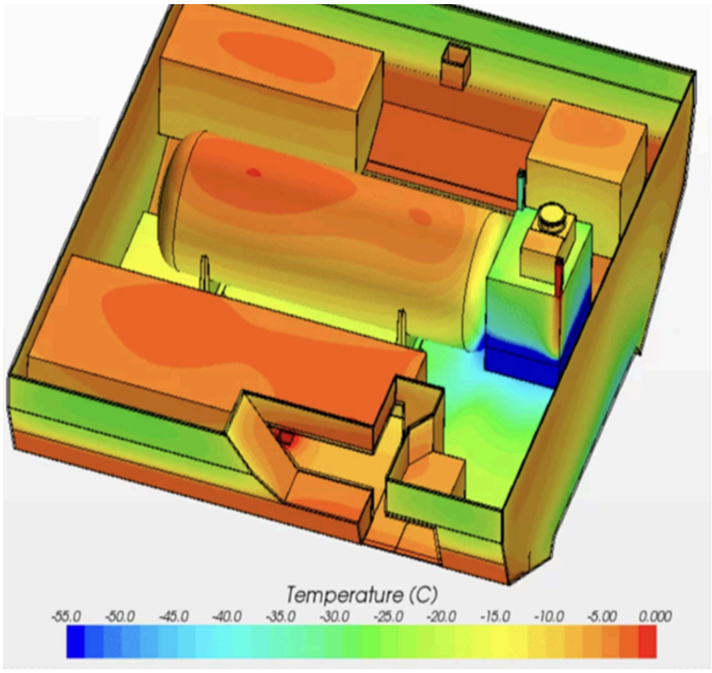
1: Digital twins and simulation



Weather forecasting (Meteociel.fr)



Flux in a turbomachine (FG GLA)



Heat propagation (Creaform)

...

Classic numerical simulation

→ Use of mathematical models on computers to predict the state of a system in the future or in a particular situation

Main benefits

→ Allow to study complex phenomena that are costly or impossible to observe directly

Main limitations

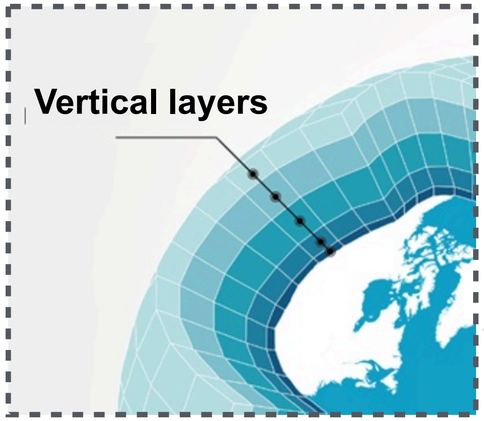
→ Rely on mathematical models (may require in-situ calibration, may be poorly understood or unknown)

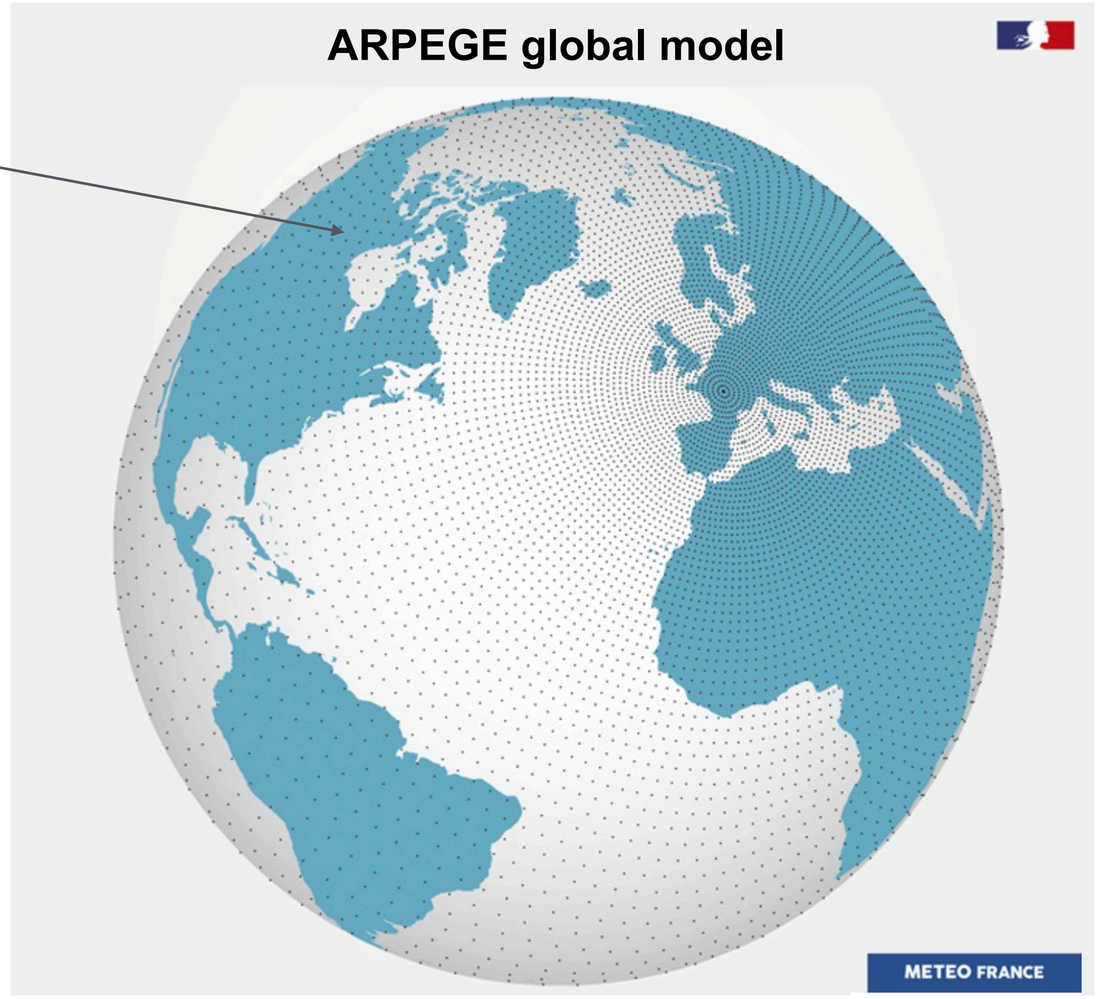
→ Potentially high computational burden



SAFIRE infrastructure

1: Digital twins and simulation



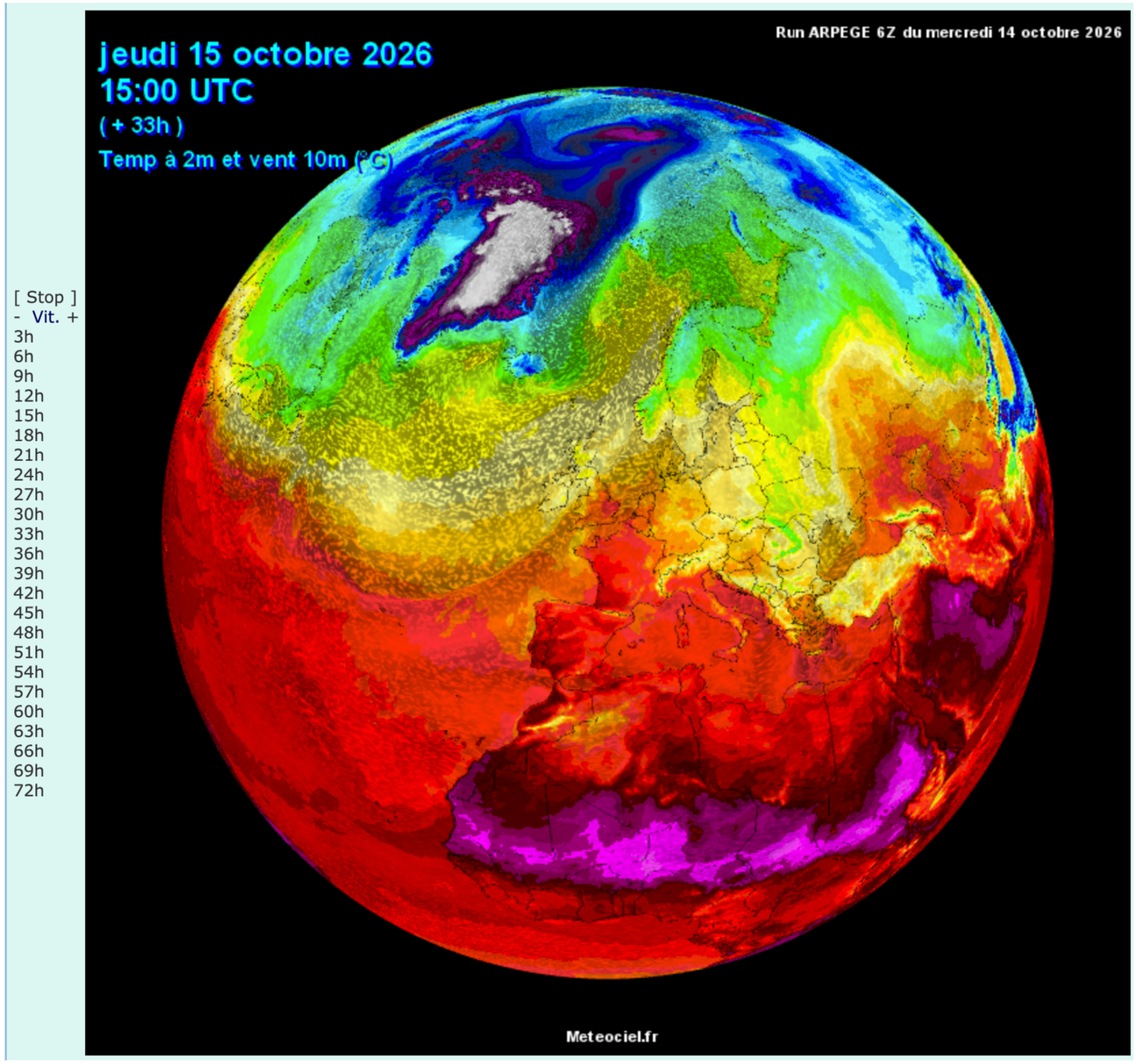


ARPEGE (started in the early 90s)

- Grid steps from 5km to 25km
- Forecasts of up to 4 days
- Output variables at 5 altitudes:
  - Temperature
  - Wind (2 components)
  - Humidity
  - Pressure
- PDE system:
  - Non-linear and chaotic
  - About 10 minutes for a time step
  - About 1h computations for a 72h forecast
  - Massive calculation server

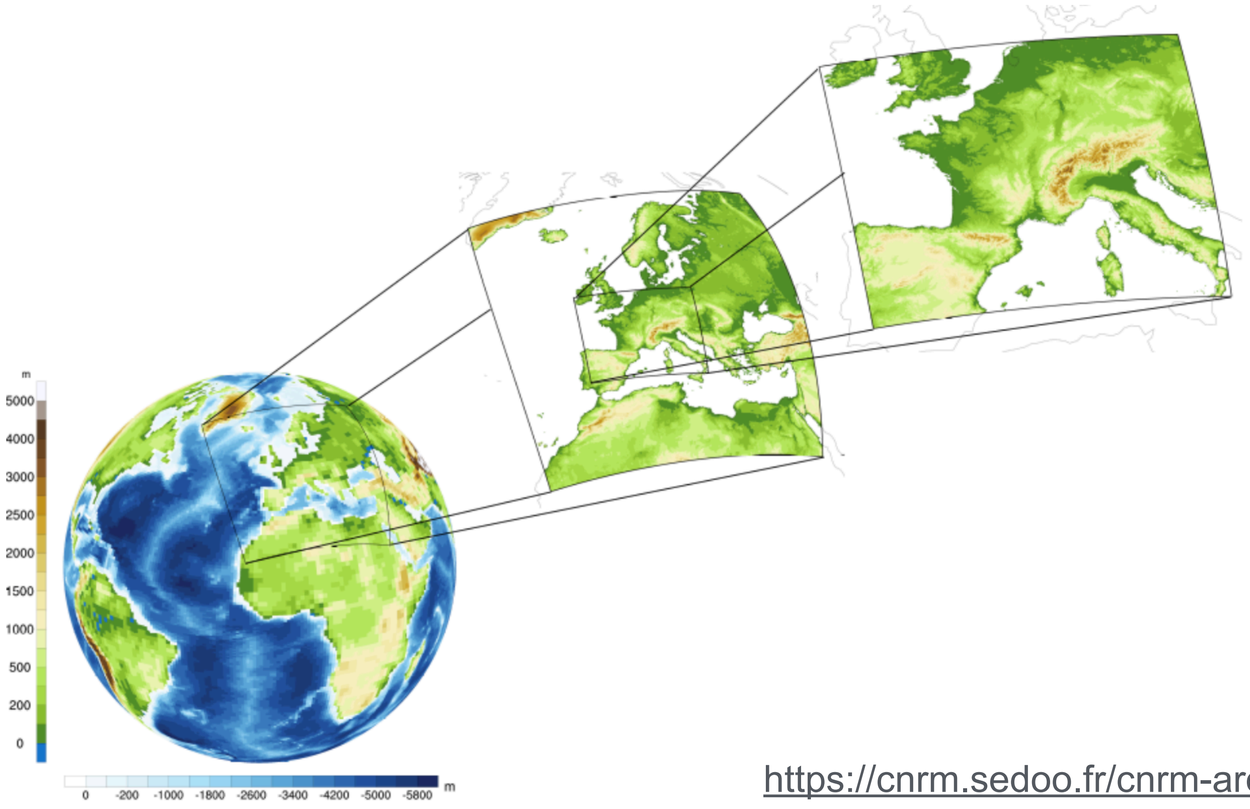
https://meteofrance.com/meteo-a-z/les-modeles-de-prevision-meteo

1: Digital twins and simulation



https://www.meteociel.fr/modeles/gfse_3d.php?arpege=1&lat=50&lon=0&ech=3&zoom=6&mode=5

1: Digital twins and simulation

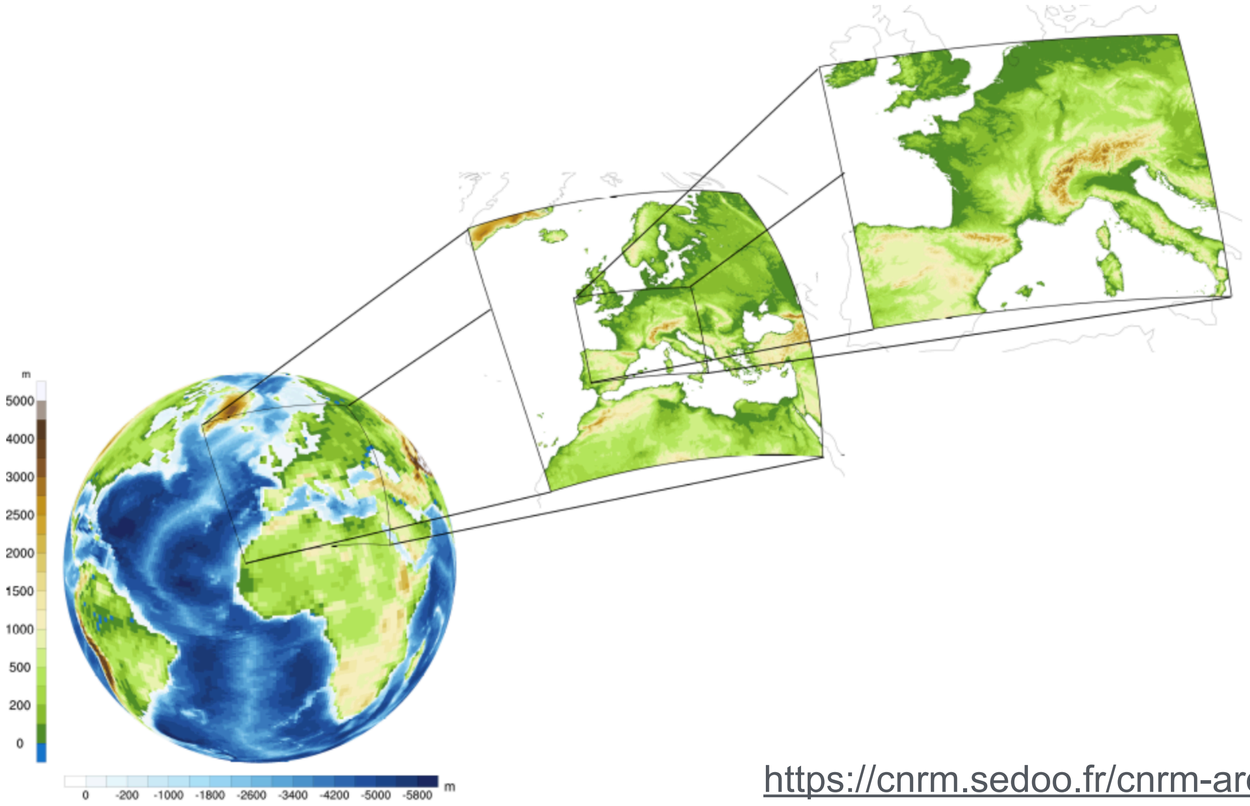


https://cnrm.sedoo.fr/cnrm-arome/

AROME (since 2014)

- Regional model ($\approx 2000 \times 2000$ km)
- 2.5km grid-step
- 60 vertical levels
- Use ARPEGE for boundary conditions
- Fine simulation of convective phenomena:
  - explicit resolution of deep convection
  - detailed surfaces description
  - advanced models for microphysics, clouds and turbulence.
  - …
- Improved weather predictions for:
  - Extreme events
  - Agriculture
  - Energy needs
  - Flight conditions
  - …

1: Digital twins and simulation



https://cnrm.sedoo.fr/cnrm-arome/

AROME (since 2014)

- Regional model ($\approx 2000 \times 2000$ km)
- 2.5km grid-step
- 60 vertical levels
- Use ARPEGE for boundary conditions
- Fine simulation of convective phenomena:
  - explicit resolution of deep convection
  - detailed surfaces description
  - advanced models for microphysics, clouds and turbulence.
  - …
- Improved weather predictions for:
  - Extreme events
  - Agriculture
  - Energy needs
  - Flight conditions
  - …

What next

- Spatial resolution <1km to account for local phenomena (*e.g.* for agriculture in non-flat areas or urban heat islands)
- Accounting for un-modelled phenomena (*e.g.* impact of rivers or large infrastructures, …)
- Lower computational needs / faster predictions

→ Using **data-driven** digital twins appears as a promising solution

2: How AI can be helpful in simulation - example of weather forecasting

GenCast: Diffusion-based ensemble forecasting for medium-range weather

Ilan Price[*,1], Alvaro Sanchez-Gonzalez[*,1], Ferran Alet[1], Timo Ewalds[1], Andrew El-Kadi[2], Jacklynn Stott[1], Shakir Mohamed[1], Peter Battaglia[1], Remi Lam[1] and Matthew Willson[1]
[*]Equal contributions, [1]Google DeepMind, [2]Imperial College, London

GraphCast: Learning skillful medium-range global weather forecasting

Remi Lam[*,1], Alvaro Sanchez-Gonzalez[*,1], Matthew Willson[*,1], Peter Wirnsberger[*,1], Meire Fortunato[*,1], Eaton-Rosen[1], Weihua Hu[1], Alexander Merose[2], Jacklynn Stott[1], Alexander Pritzel[1], Shakir Mohamed[1] and

Pangu-Weather: A 3D High-Resolution System for Fast and Accurate Global Weather Forecast

Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian✉, *Fellow, IEEE*



Since 2022: Data-driven weather forecasting models from *Google*, *Huawei*, *NVIDIA*, *Microsoft*, *ECMWF*, …

2: How AI can be helpful in simulation - example of weather forecasting

On a single Cloud TPU v4 device, GraphCast can generate a 0.25° resolution, 10-day forecast (at 6-hour steps) in under 60 seconds. For comparison, ECMWF's IFS system runs on a 11,664-core cluster, and generates a 0.1° resolution, 10-day forecast (released at 1-hour steps for the first 90 hours, 3-hour steps for hours 93-144, and 6-hour steps from 150-240 hours, in about an hour of compute time [41].

GraphCast: Learning skillful medium-range global weather forecasting

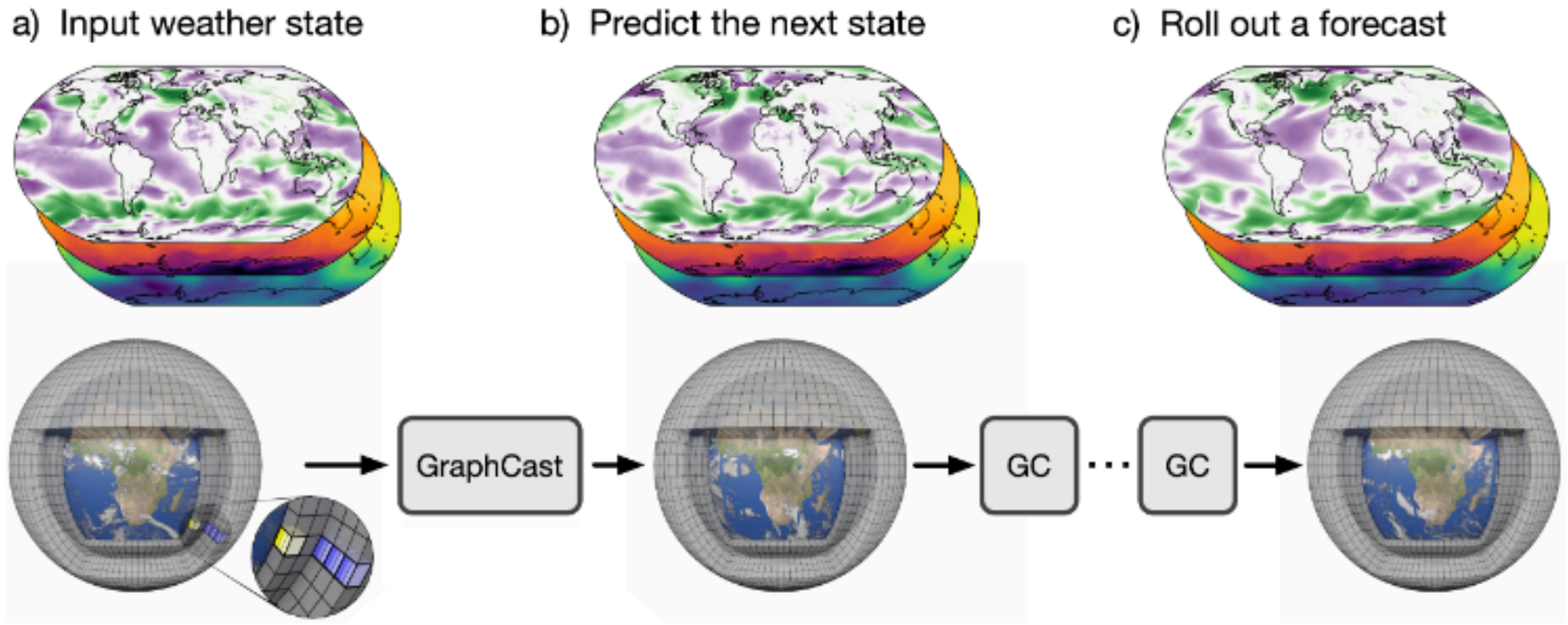
GraphCast's forecast skill and efficiency compared to HRES shows MLWP methods are now competitive with traditional weather forecasting methods. Additionally, GraphCast's performance on severe event forecasting, which it was not directly trained for, demonstrates its robustness and potential for downstream value. We believe this marks a turning point in weather forecasting, …

Since 2022: Data-driven weather forecasting models from *Google*, *Huawei*, *NVIDIA*, *Microsoft*, *ECMWF*, …

2: How AI can be helpful in simulation - example of weather forecasting

Example: GraphCast (Lam et al, 2023)

General weather prediction strategy

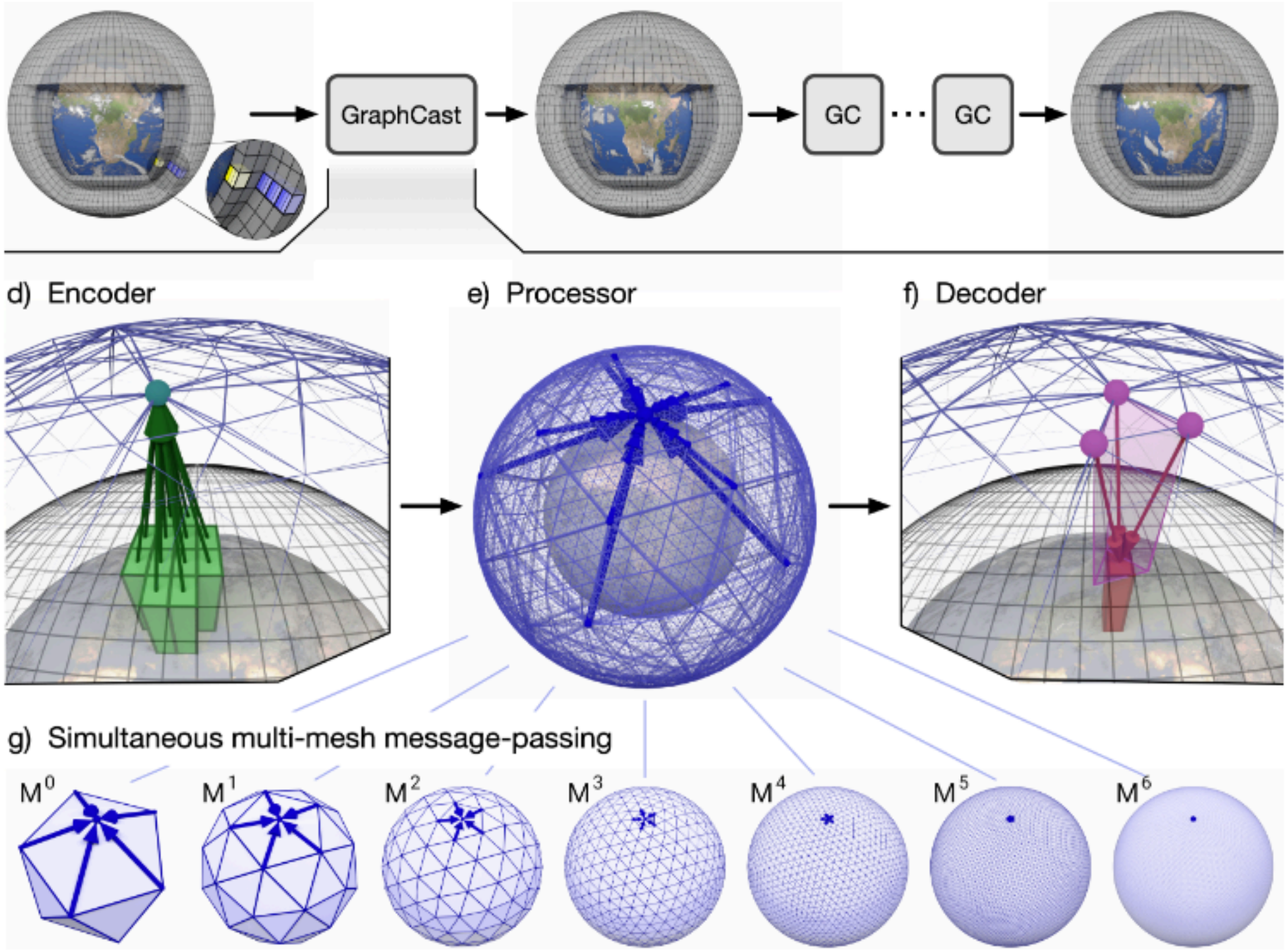


- 0.25° latitude-longitude grid ($721 \times 1440 \approx 10^6$ spatial locations) $\Rightarrow$ Input dimension $> 2.3 \times 10^8$ variables
- 6 atmospheric variables repeated at 37 pressure levels

Open training dataset : ERA5 data → One of the most accurate reconstructions of the past weather on a ~30km global mesh, available since 1940 (Missing : high-resolution)

2: How AI can be helpful in simulation - example of weather forecasting

Example: GraphCast (Lam et al, 2023)



(d) Encoder: maps local regions of the input into nodes of the multi-mesh graph representation

(e) Processor: updates each multi-mesh node using learned message-passing

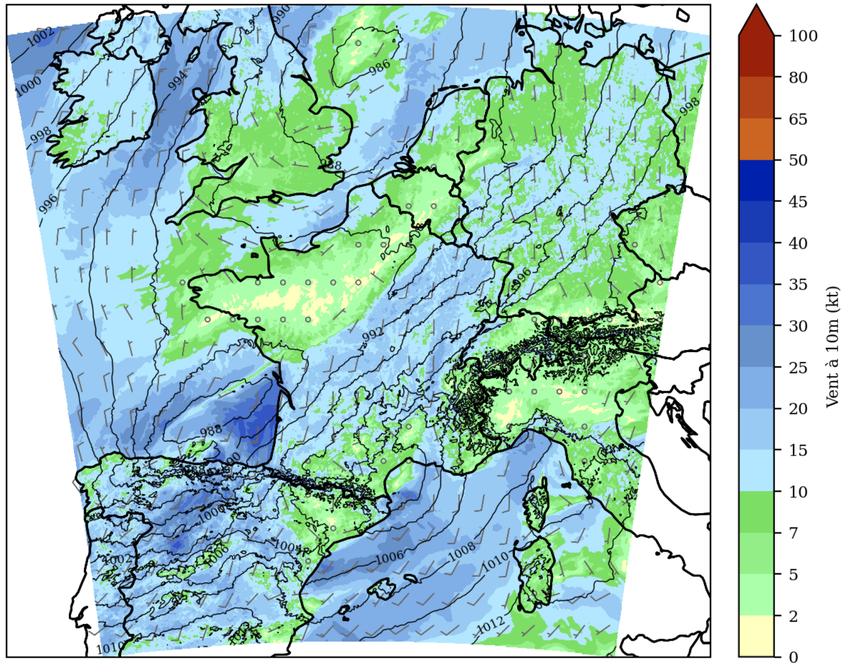
(f) Decoder: maps the processed multi-mesh features back onto the grid representation

2: How AI can be helpful in simulation - example of weather forecasting

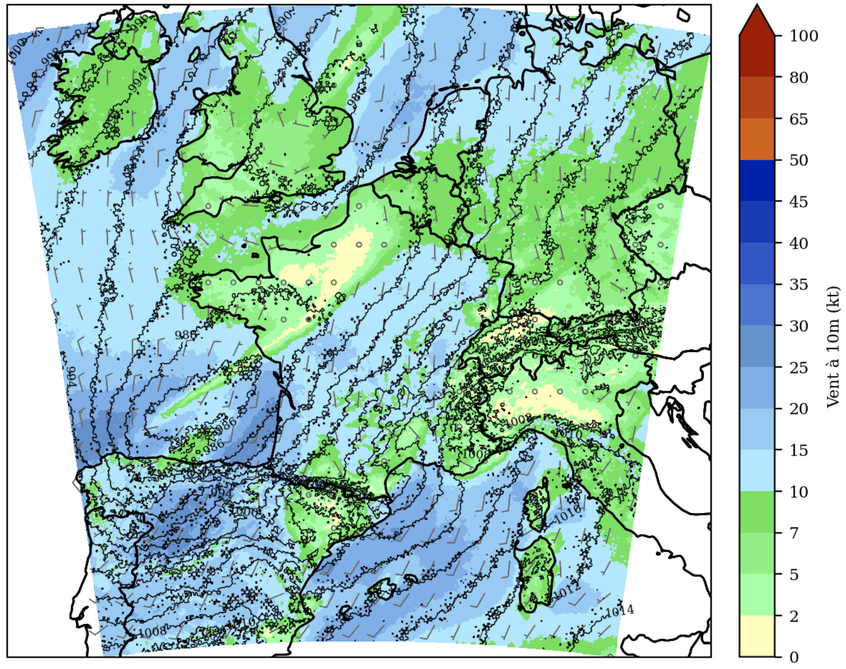
GraphCast-inspired strategy to address high-resolution weather forecasting at CNRM/Météo-France

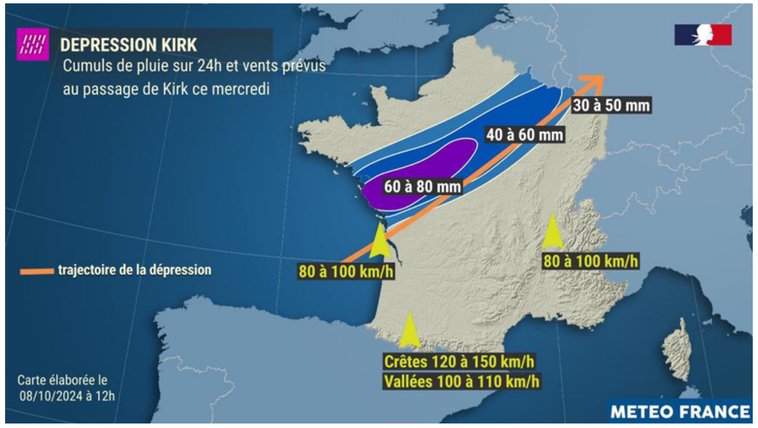
- First data-driven model for regional weather forecasting over France set up at CNRM
- Storm Kirk (09/10/2024) forecasted below

Operational physics-based model



Experimental AI-based model



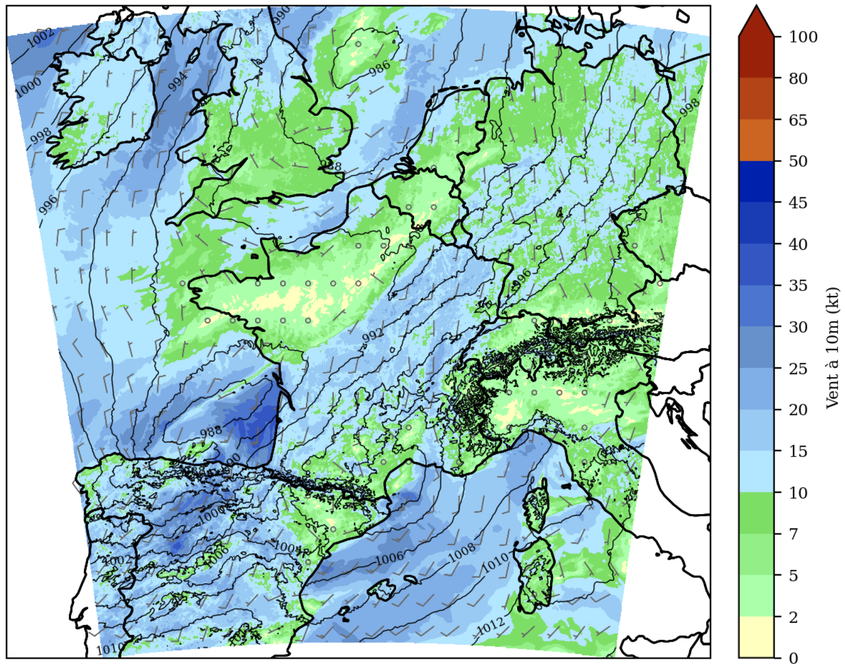


2: How AI can be helpful in simulation - example of weather forecasting

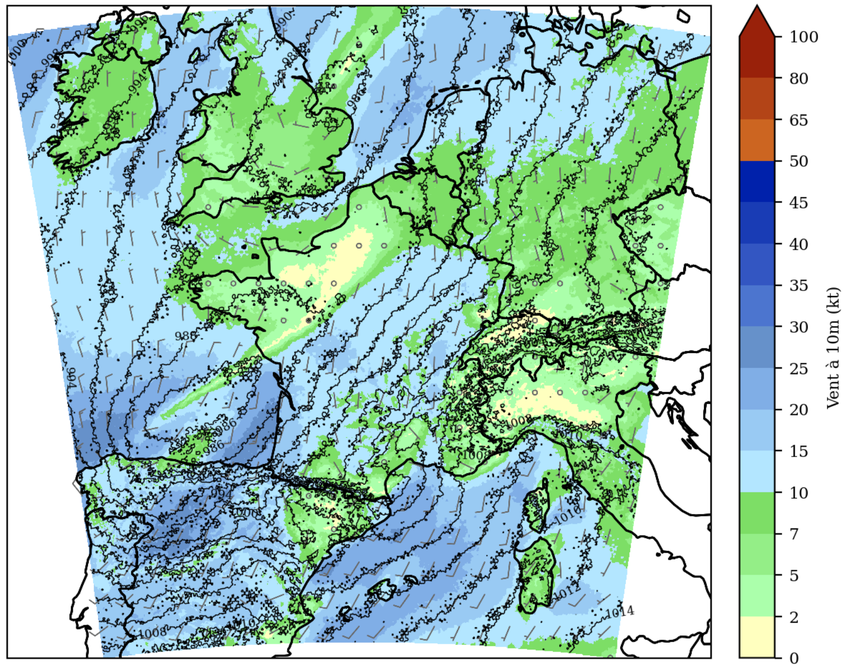
GraphCast-inspired strategy to address high-resolution weather forecasting at CNRM/Météo-France

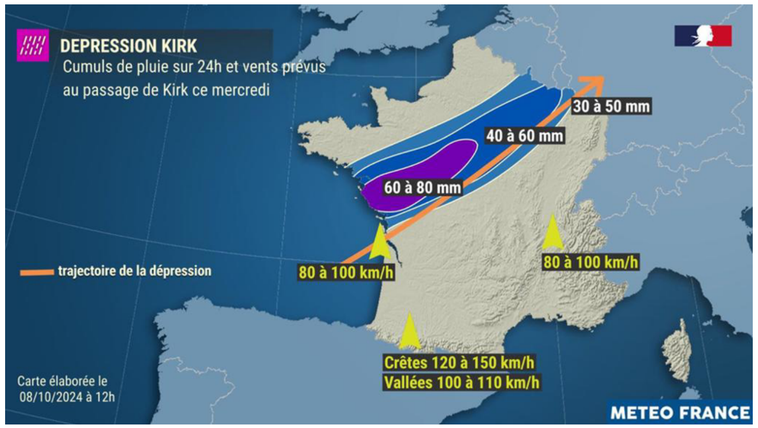
- First data-driven model for regional weather forecasting over France set up at CNRM
- Storm Kirk (09/10/2024) forecasted below

Operational physics-based model



Experimental AI-based model





- **+** : Real-time demonstrator of AI-AROME model since February 2026 (internal use for now)
- **-** : Some rare or extreme events are poorly predicted by the AI-based model
- **-** : It remains a black-box model… should we take high impact decisions based on it?

Scientific fields:

- Biases in machine learning
- Explainable AI

3: Explaining the predictions for trustworthy IA

Hidden Biases and Neural-Networks: The « Huskies vs Wolves » example (Ribeiro et al, 2016)

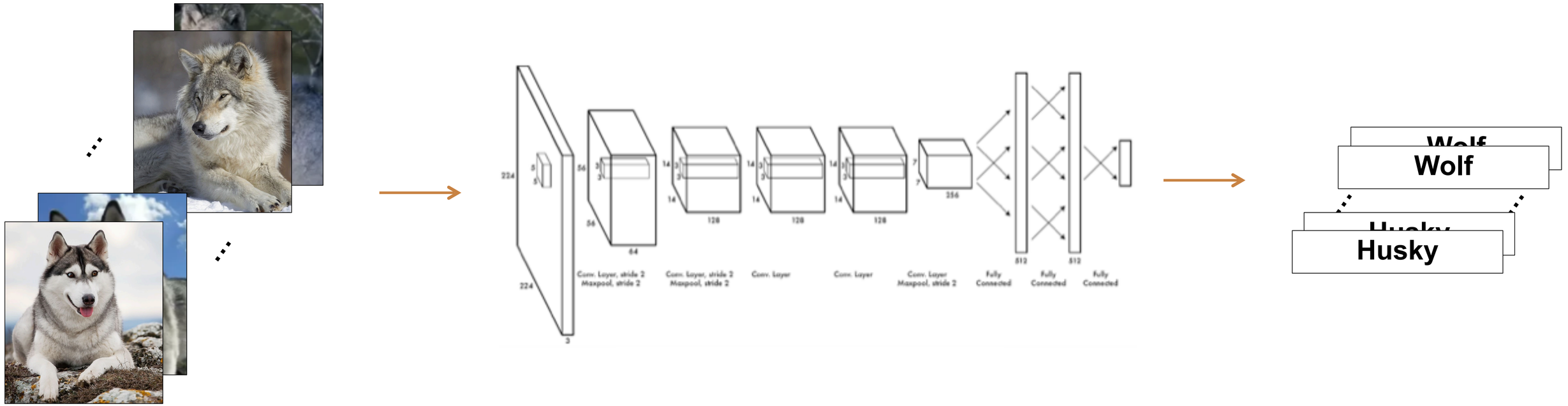


Neural-network goal: Detect if a wolf or husky is represented in a photo (in case we have a doubt)

3: Explaining the predictions for trustworthy IA

Hidden Biases and Neural-Networks: The « Huskies vs Wolves » example (Ribeiro et al, 2016)

Step 1: Train the neural network using a series of examples

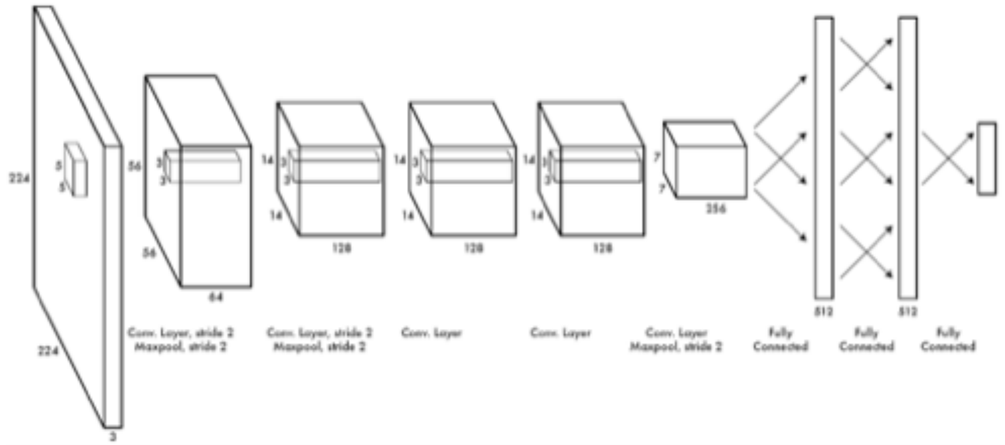


3: Explaining the predictions for trustworthy IA

Hidden Biases and Neural-Networks: The « Huskies vs Wolves » example (Ribeiro et al, 2016)

Step 2: Predictions





Wolf

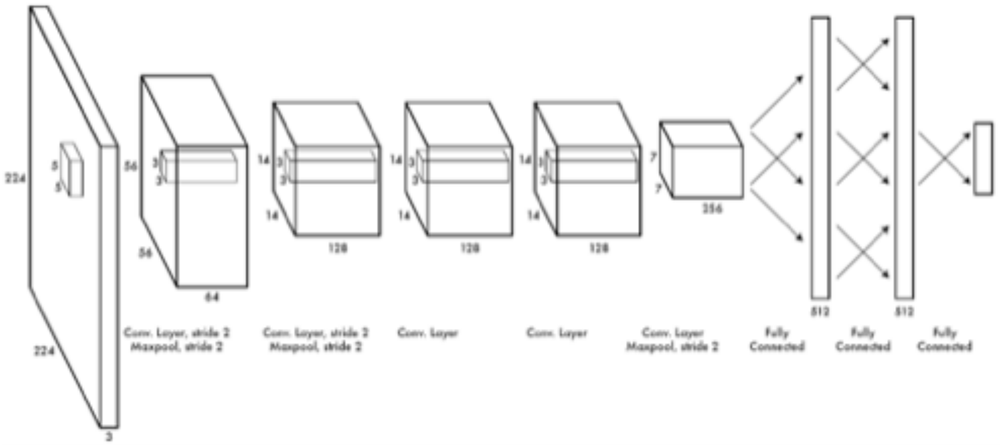
95% accuracy on images that have not been used for learning!

3: Explaining the predictions for trustworthy IA

Hidden Biases and Neural-Networks: The « Huskies vs Wolves » example (Ribeiro et al, 2016)

Step 2: Predictions





Wolf

It's actually a husky!

Why?

⇢ The images used to learn the neural network often represent wolves in the snow, which is not the case for huskies

⇢ It is easier for a neural network to detect snow than the shape of the animal



3: Explaining the predictions for trustworthy IA

Hidden Biases and Neural-Networks: The « Huskies vs Wolves » example (Ribeiro et al, 2016)

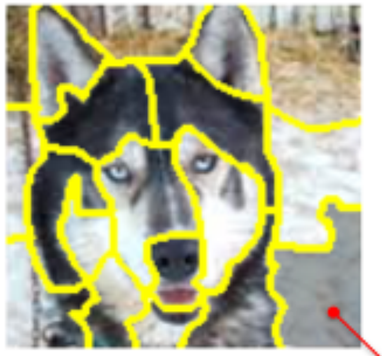
→ How to discover unknown biases: Explanations!



Husky* predicted as a *Wolf



(a) Segment the image into several regions



(b) Impact of changes in each region on the prediction



(c) Representation of the most impactful regions

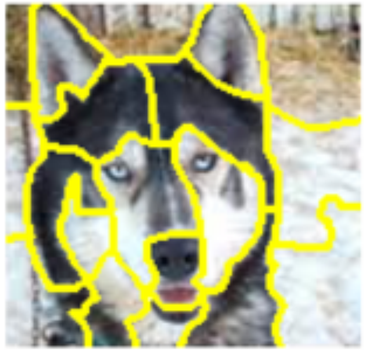
3: Explaining the predictions for trustworthy IA

Hidden Biases and Neural-Networks: The « Huskies vs Wolves » example (Ribeiro et al, 2016)

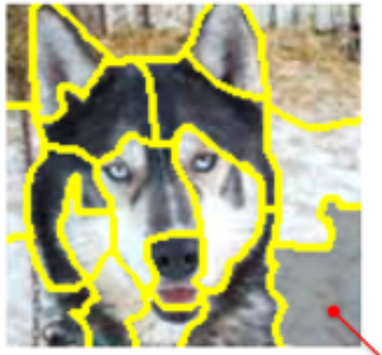
→ How to discover unknown biases: Explanations!



Husky* predicted as a *Wolf



(a) Segment the image into several regions



(b) Impact of changes in each region on the prediction



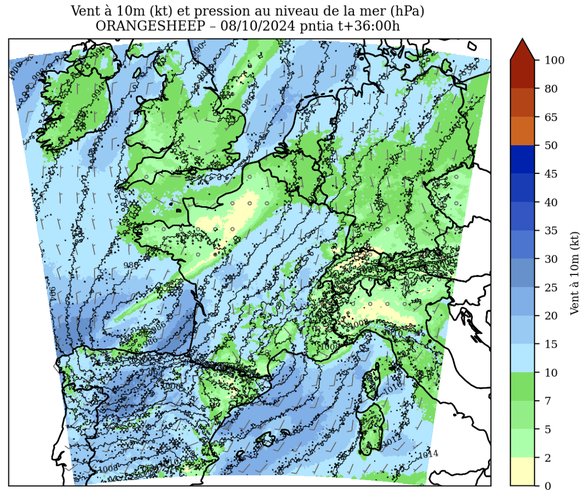
(c) Representation of the most impactful regions

Application to weather forecasting

- Understand and/or evaluate AI-based models robustness for rare or extreme events
- Explain the predictions to make them trustworthy

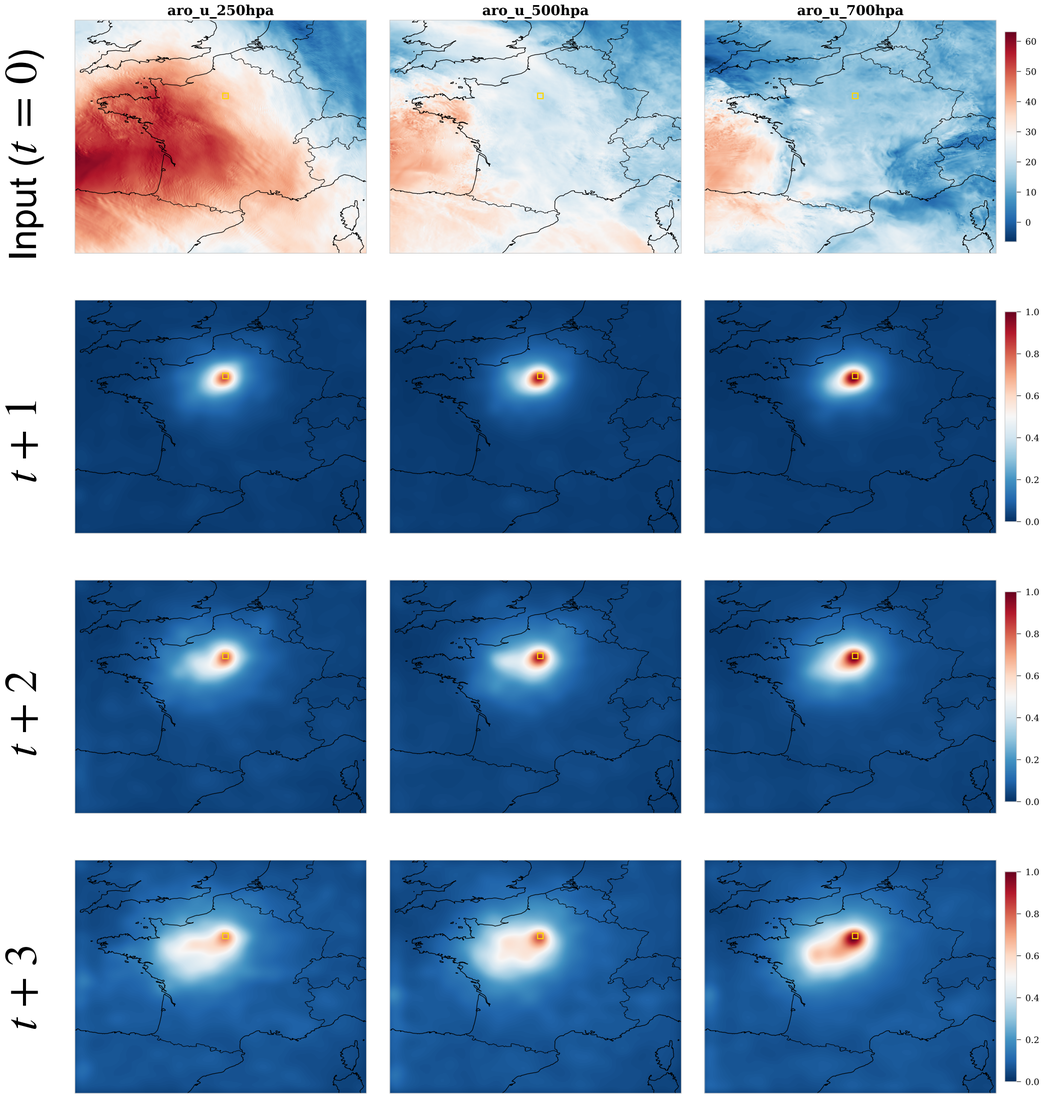
Scientific questions

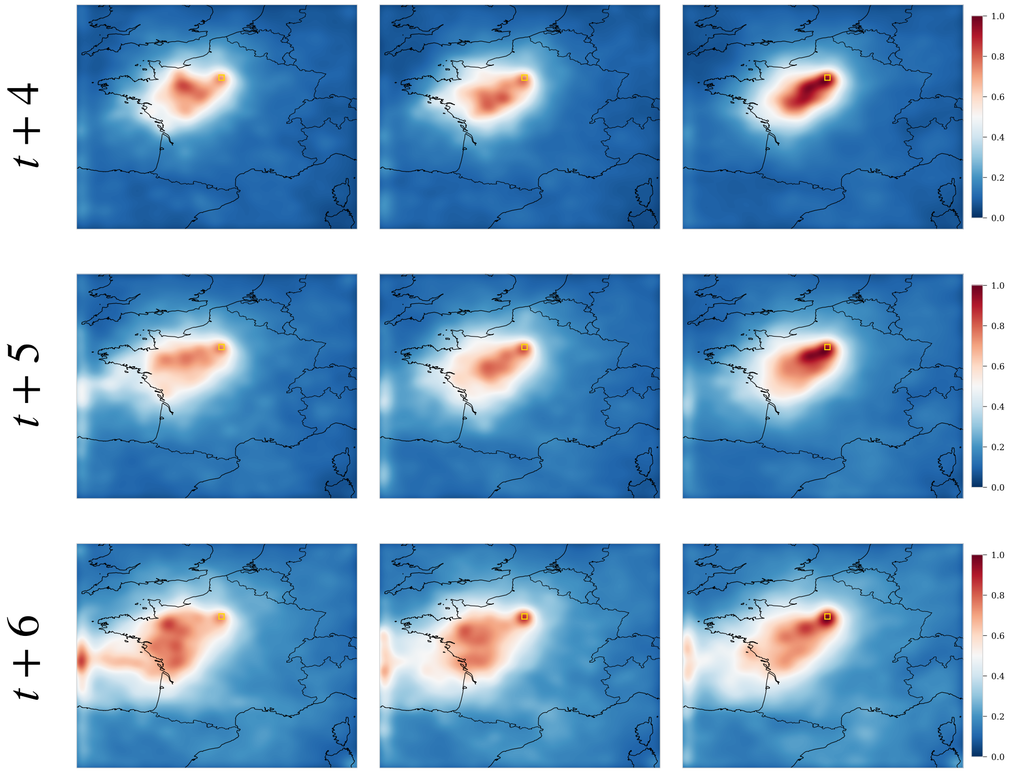
- How to extend existing tools (for images, tabular data, …) to physics-aware and high-dimensional data?
- How to make the explanations understandable for meteorologists, public decision makers, …



3: Explaining the predictions for trustworthy IA

PhD thesis of Younes Essafouri on Explainable AI for regional weather forecasting





Auto-regressive attribution maps during Storm Ciarán (11/2023)

→ Strategy accepted for NeurIPS 2026

→ https://younesessafouri.github.io/WassersteinGrad/

Essafouri Y., Raynaud L., Drozda L, Risser L.: Explanation of Dynamic Physical Field Predictions using WassersteinGrad: Application to Autoregressive Weather Forecasting. Proceedings of the *Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2026.

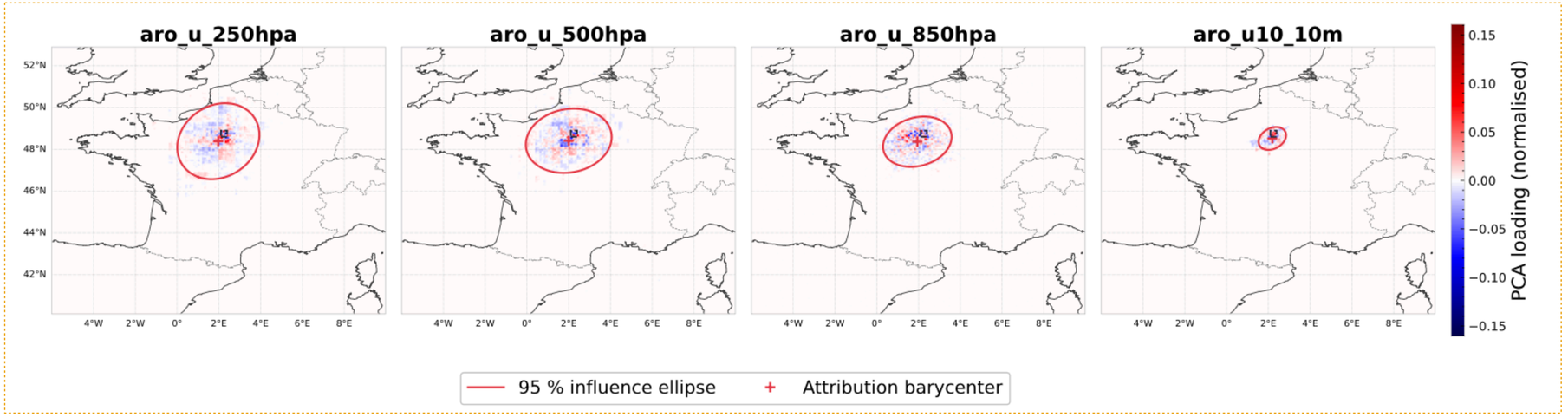
3: Explaining the predictions for trustworthy IA

PhD thesis of Younes Essafouri on Explainable AI for regional weather forecasting

Extraction of quantified features across input variables

- Strategy presented at EGU 2026
- Ongoing package development

Variable	Centre (Lon, Lat)	Area (Spread)	Main Axis	Anisotropy
	Paris: (2.35°, 48.86°)			
Zonal Wind 250 hPa	(1.87°, 48.51°)	0.37	37.3°	1.13
Zonal Wind 500 hPa	(1.90°, 48.53°)	0.27	30.0°	1.21
Zonal Wind 850 hPa	(1.98°, 48.54°)	0.15	19.6°	1.37
Surface Wind 10m	(2.04°, 48.55°)	0.10	24.7°	1.35



Essafouri Y., Seznec C., Drozda L., Raynaud L. and Risser L.: A Framework for Explainable AI in Weather Forecasting Diagnosing Deep Learning Models via Gradient-Based Attributions. Proceedings of the EGU General Assembly 2026

Conclusions

Neural-based digital-twins appear as an interesting option:

- to simulate poorly understood phenomena
- to strongly accelerate numerical simulations

It now appears as essential to (at least partially) understand how neural-based decisions are made:

- to make them trustworthy in high-impact contexts
- to detect unknown biases
- to simply understand what has been captured in the data

To go further on the use of AI-systems in an industrial context:

- European « AI act » regulation: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689
- ISO standards on AI: https://www.iso.org/sectors/it-technologies/ai
- Institut National pour l'évaluation et la sécurité de l'IA (INESIA) : https://www.sgdsn.gouv.fr/files/files/Nos_missions/Feuille%20de%20route%20INESIA%202026-2027.pdf