RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

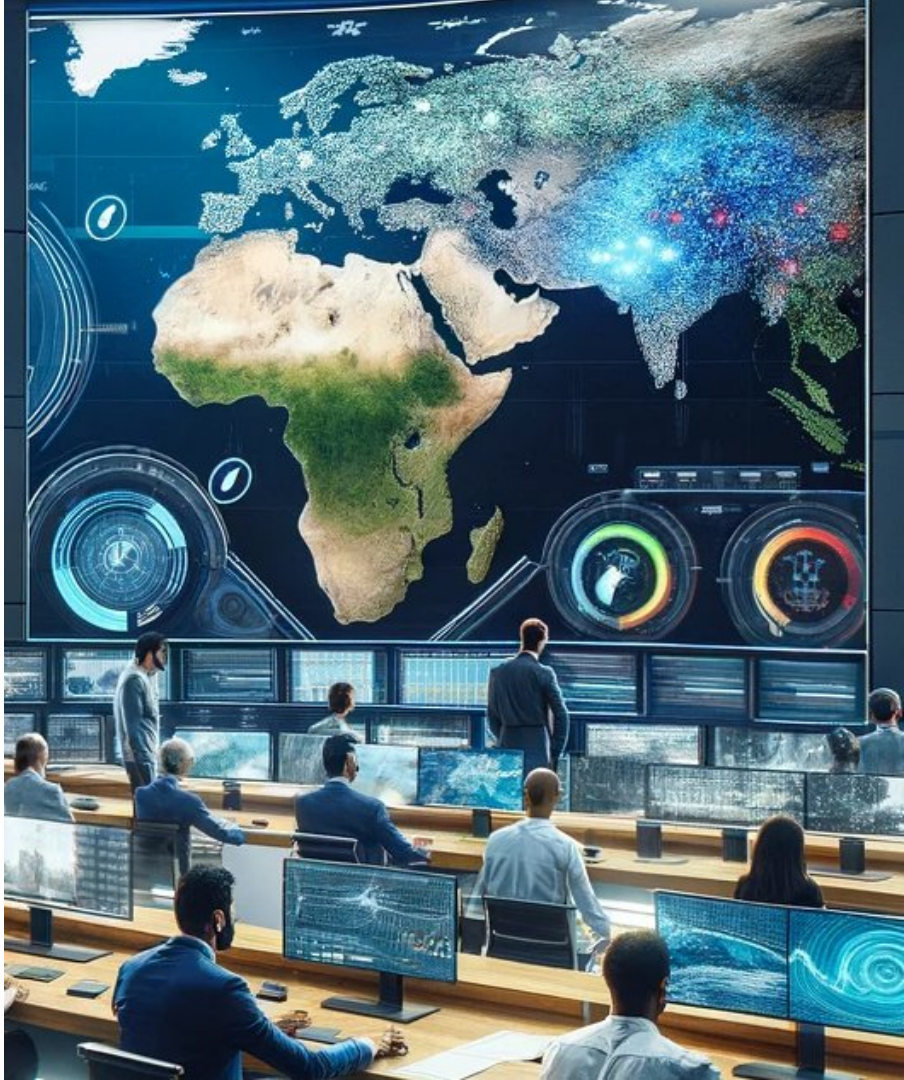
Les activités du DSM / Lab IA

Léa Berthomier, Oscar Dewasmes, Frank Guibert, Bruno Pradel, Théo Tournier

Journées de l'IA - 17 février 2026

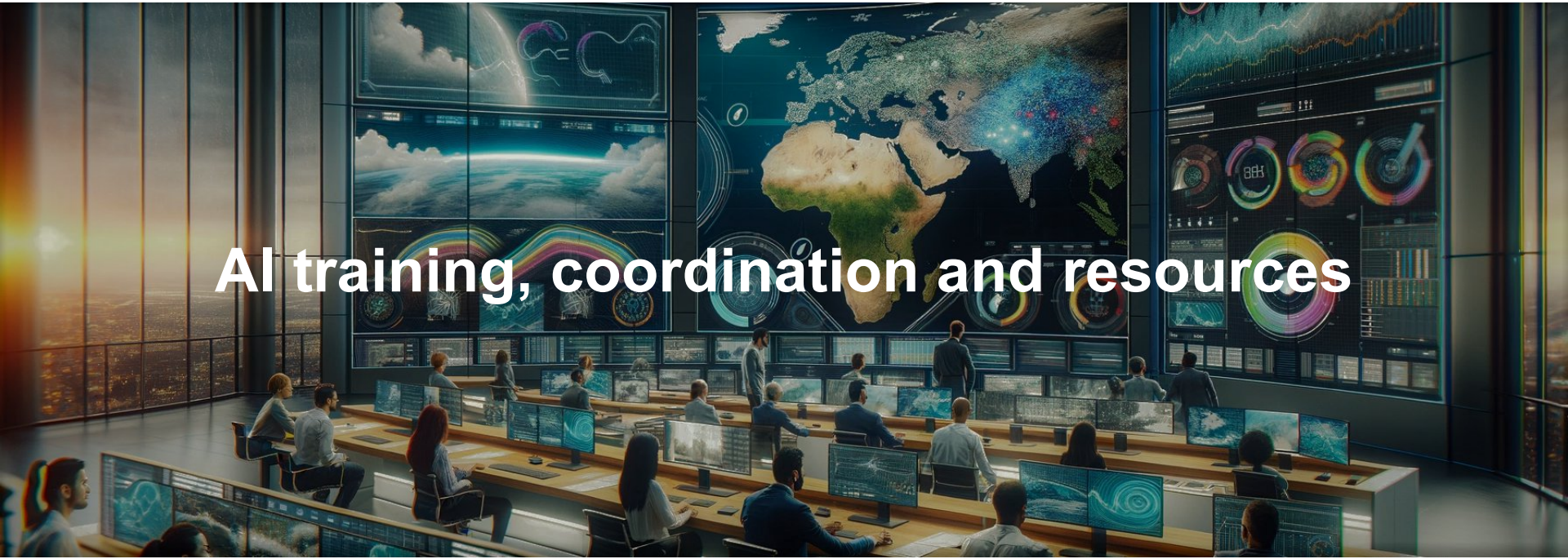
Contents

- AI training, animation and resources
- Generative AI access at MF
- AI Lab: ongoing projects
  - Precipitations nowcasting
  - Marine text report initialization
  - CAMS output statistics
  - Wave model downscaling



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE



AI training, coordination and resources

Club Deep Learning

Monthly discussion forum, both in person and virtually.

Meeting of AI practitioners and enthusiasts, presentations on a voluntary basis, on ongoing work, tutorials, bugs, best practices…

Mailing list subscription: deep_learning@meteo.fr

Machine Learning trainings

1. Continuing Education:
Machine Learning for the Curious (3 days)
Led by Bruno Pradel (DSM/LabIA), 2 sessions per year.
Materials available on GitHub.

2. Online :
FIDLE : Formation d'Introduction au Deep Learning
Made by the **CNRS**, open to all, 20 courses of 2h.
All the courses are available on Youtube.

3. Tips for your project :
You are starting an AI project ?
Contact the DSM/LabIA, to help you to start your project !
AI methodology, best pratcices, peer reviewing, ...
Send a mail to **dsm_labia@meteo.fr**

Initial training

Software engineering (IENM2 - 12h)

- Learn the best practice of software engineering
- python, git, merge request
- Practice with real world project

ENM
École Nationale de la Météorologie

Artifical Intelligence (IENM3 - 40h)

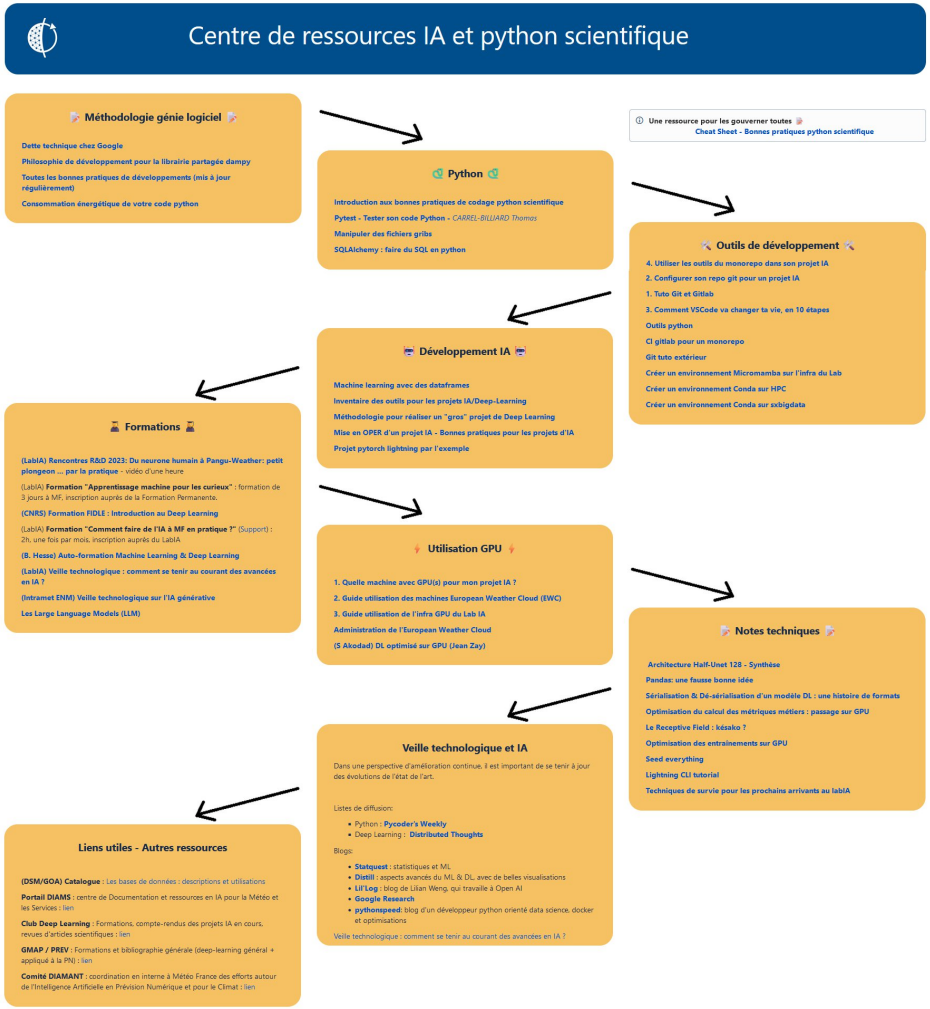
- Theorical basis in AI
- Introduction to political context and to the economical and environmental impacts
- Best practices of software engineering
- Translate a customer need to an AI solution
- Build an AI project
  - Prepare and load data, training, evaluate the neural network
- Presenting and critiquing results

It is possible to add some Météo-France agents to these trainings.

Resource center

AI Lab space on Confluence :

- Methodology in software engineering
- Python
- Development tools
- AI development
- Learnings
- GPU usage
- Technical report
- Technological scanning
- Useful links



[AI Lab ressource center]

GPU ressources

Installation in progress of a new 8 H100 node :

- **modern optimisations** (flash attention, bf16).
- **train last models** (LLM, MLM).
- **+ energy efficient.**
- + RAM memory (80Go) → **train large models** (DGMR).
- Mount */local* disks (14To) to increase data loading.

Extend the */scratch* *(100 -> 114To)* to allow users to **manipulate larger dataset** (90% of usage before the install).

Access to this new node will be via dsm_labia@meteo.fr, only for heavy and optimized projects.

8 H100 node will be available soon !

GPU resources

Multiple access to server with GPUs :

- Bélénos,
- European Weather Cloud,
- LabIA server (*priam*),
- SIRES server (*sxbigdata*),
- Jean Zay

Please contact **LabIA, SIRES** or **DSI/ISI/CC** for further informations.

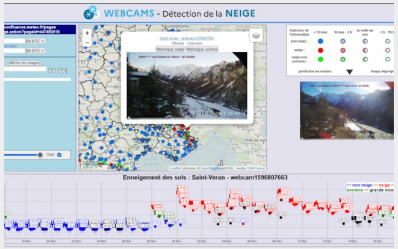
Nom GPU disponible Contact Critères	EWC (1 A100) SIRES ou DSI	*belenos* (3 nœuds 4 V100) DSI/ISI/CC	*priam-sidev (node00[1-5])* (5 nœuds 4 V100) DSM/LabIA	*priam-sidev (node006)* (1 nœud 8 H100) DSM/LabIA	*sxbigdata.cnrm*
Je dois accéder à de gros volumes de données (>10To) sur les bases de données Météo-France.					
Mon projet va durer moins de 6 mois. (ex. stage)					
J'ai besoin de plusieurs GPU en parallèle pour une seule expérience IA.					
J'ai besoin de lancer plusieurs expériences IA en parallèle.					
Je souhaite utiliser Docker pour gérer mon environnement de travail Python.					
Je souhaite utiliser Conda pour gérer mon environnement de travail Python.					
J'ai besoin d'une grande bande passante entre le GPU et les données. (ex. chaque échantillon d'entraînement est volumineux)					
Mon projet, déjà fonctionnel et optimisé, requiert des ressources importantes.					

Confluence's page to help you to choose the best server for your project

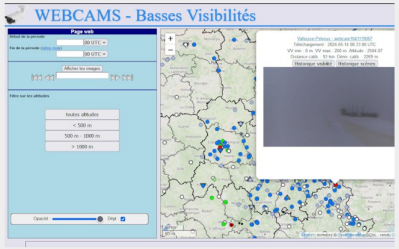
4 nodes with 8 H200 will be available soon on a new server !

Portail des produits IA temps-réel

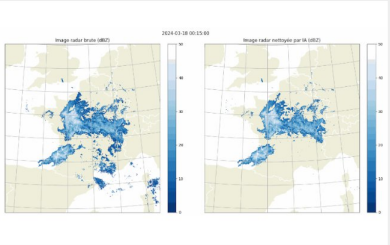
(DSO) Détection de neige sur webcams



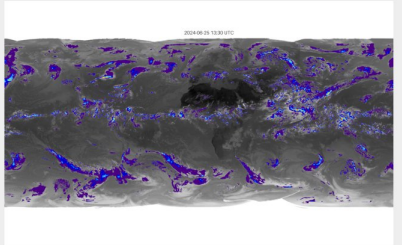
(DSO) Détection de basse visibilité sur webcams



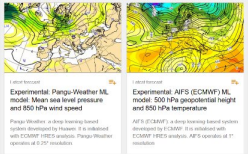
(LablA) Nettoyage d'images radar pour les médias



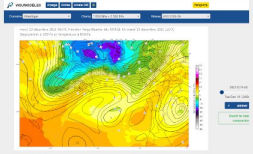
(LablA/CMS) Espresso : détection des pluies par satellite



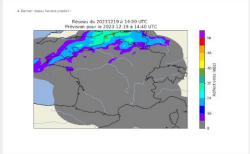
(ECMWF) Emulateurs globaux de PN (AIFS, GraphCast, Pangu...)



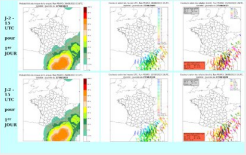
(CNRM) Pangu-Weather initialisé par Arpège



(LablA/DiropPI) Prévision Immédiate de réflectivité

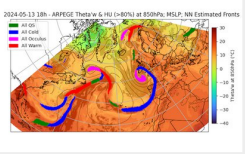


(CNRM) Détection d'échos arqués sur PE Arome

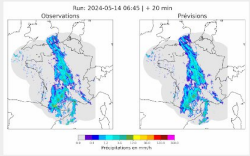


(CNRM) Scénarios de PE de pluie sur le Sud-Est

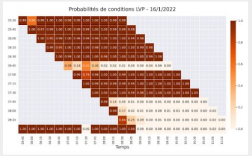
(LablA) Détection de fronts sur Arpège



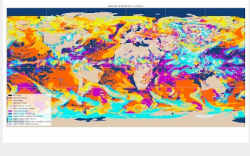
(LablA/ENM) DGMR : Prévision Immédiate de lame d'eau



(LablA) Prévision immédiate de LVP pour Roissy



(LablA) Prévision immédiate de nébulosité



[Portal of AI products]

Please add your project if you do not see it.

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

Generative AI usage at Météo-France

Generative AI access

Usage of Gen. AI is **prohibited by our DSI** !

→ websites with access to Gen. AI are **blocked**.

Generative AI access

Bard GitHub Copilot Gemini

Bing ChatGPT deepseek

The strategy of these private companies is to provide a free service at a loss, in order to

- Build customer loyalty
- Create an **intellectual and systemic dependancy**
- Made customers dependant to a service that will become paid

Generative AI access

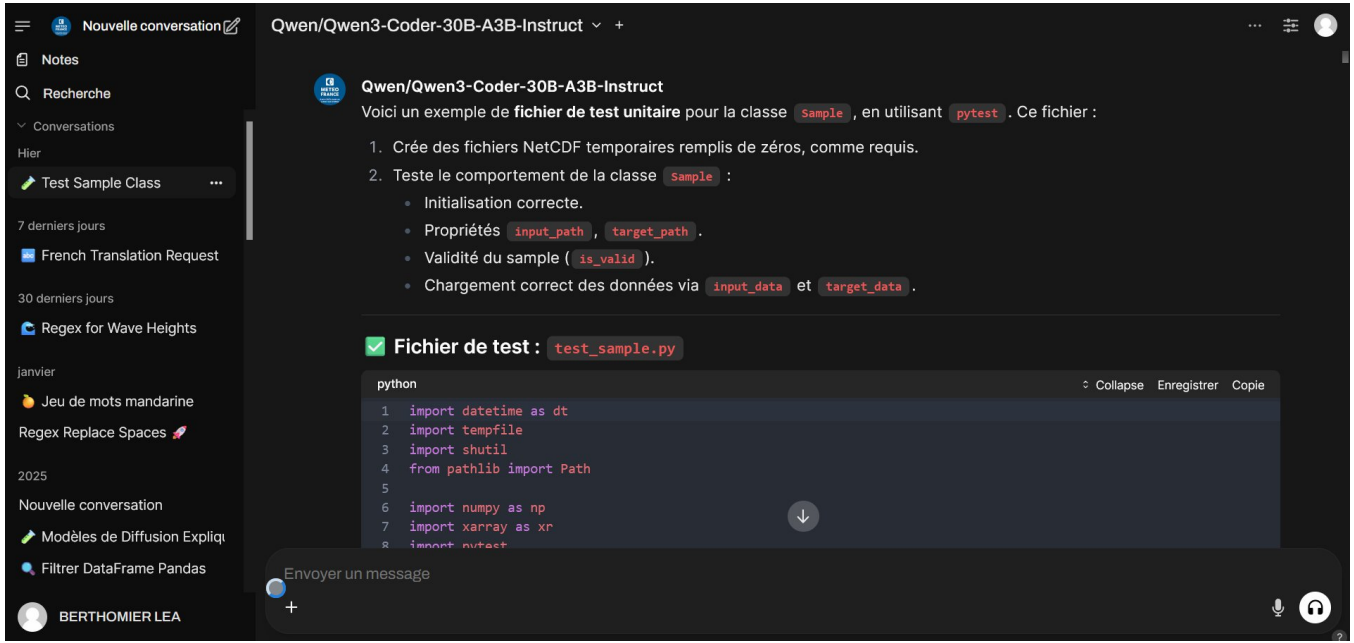
Each request to this private models, your data are send to tier server which be used for the next training or sell to an other company.

So you have to be careful to :

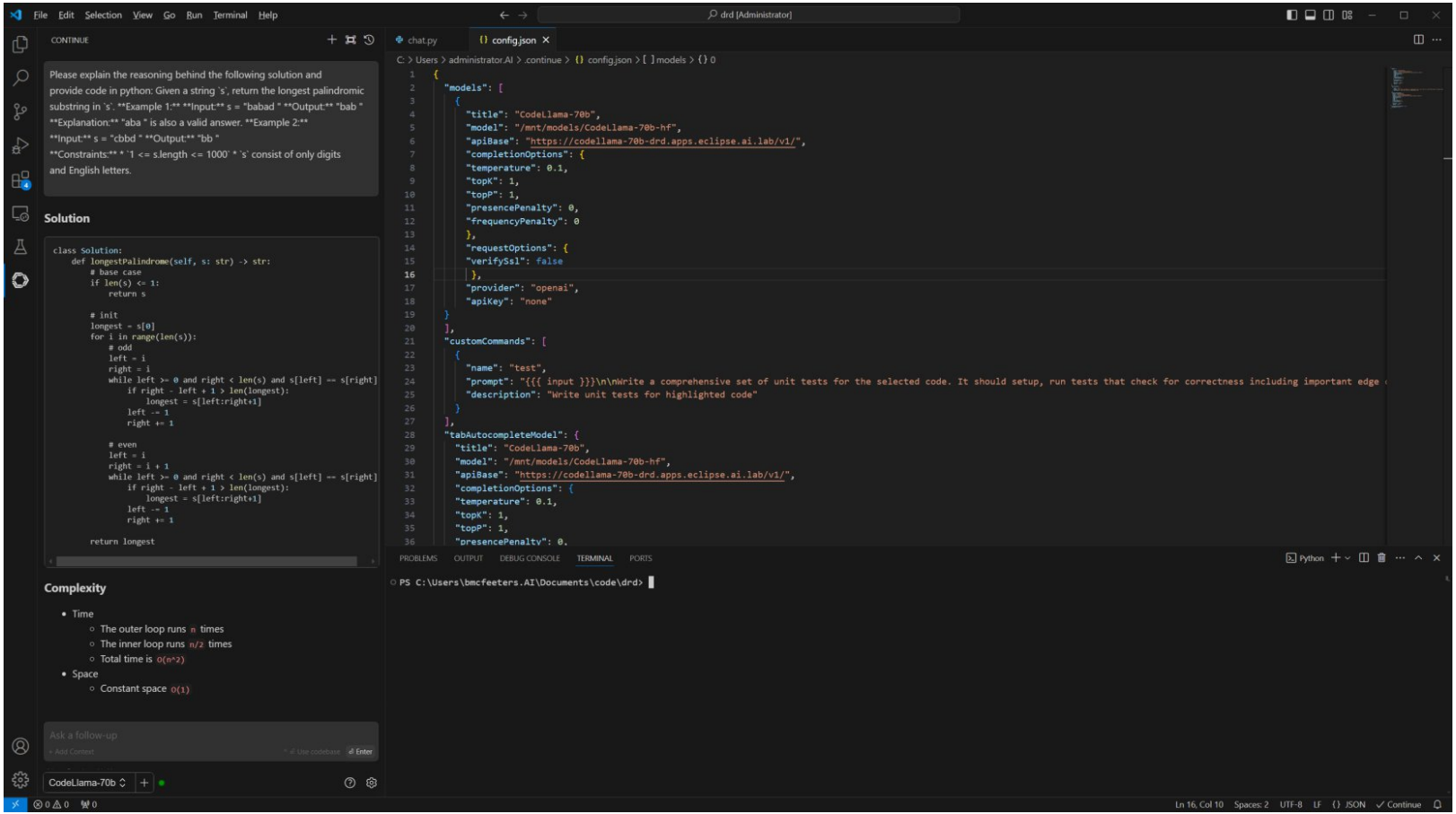
- **Intellectual property leak:** unpublished articles, experimental results, code…
- **Personal data leak:** CVs, personal information…
  → RGPD issue
- **Confidential information leaks:** passwords, information about IT systems, ...
  → Cybersecurity issues: phishing, attacks, ransomware, ...

Generative AI access

Work in progress at DSI/DEV team, in partnership with DINUM, to propose an internal solution, using a sovereign AI, without data leak, to all Météo-France's agents.



Generative AI access




Ongoing projects

Ended and ongoing projects

4

Operational products

Espresso

Radar imagery cleaning

Fog forecasting over Seine river

Cloud cover nowcasting

3

Integration product

Sargassum algae detection

Low visibility nowcasting

Precipitation nowcasting (*w. DIROP/PI*)★

2

Delivered projects

Wind gust transfer funvtion

Waves downscaling ★

3

Current projects

Precipitation nowcasting (sovereign version of DGMR) ★

Marine text bulletin init. ★

MOS CAMS ensemble ★

2

Pending projects

Deepsyg

ChatBot DRIAS

All the projects are made by all the AI Lab members.

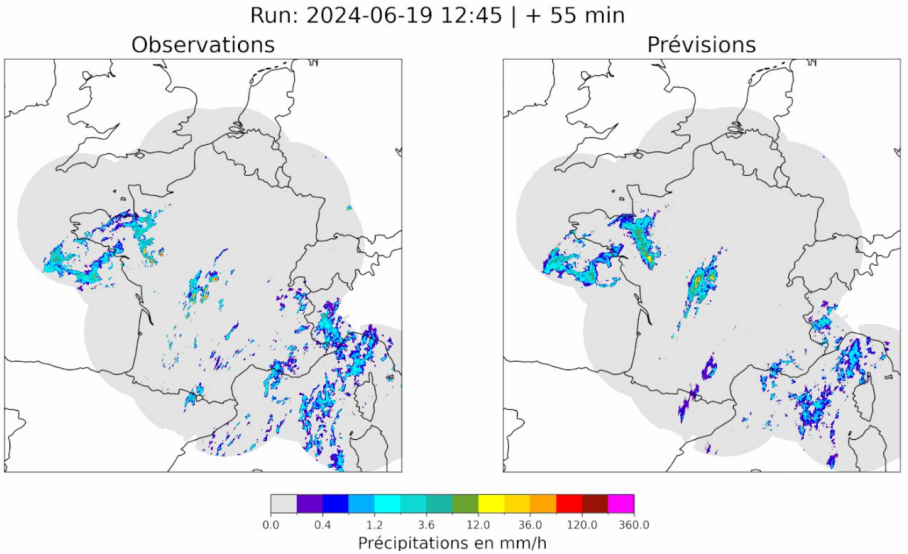
Precipitation Nowcasting (avec DIROP/PI)

Goal : benchmarking 4 deep learning approach for precipation nowcasting

DGMR (Google), **U-Net** (MF), **LDCast** (Meteo-Swiss), **NowcastNet** (Tsingua University)

➡ **DGMR (GAN) works best !**

Problem : uses Google's trained weights, **MF do not owns** weights, commercial usage not allowed



Ongoing work : develop and train internal/proprietary version of DGMR
Training time : 100 days with **4x V100** GPUs
8 days with **8x H100** GPUs

Today at 11a.m. : **DIROP/PI's activities**.

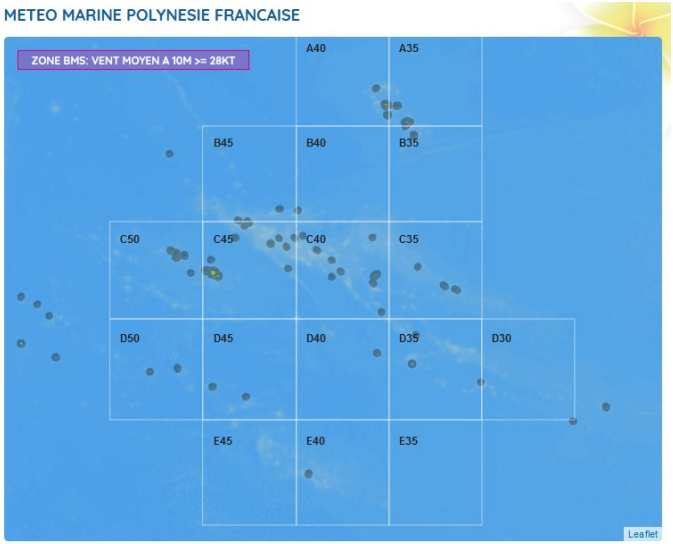
Marine Report
(Bulletin Marine Regulier)
Initialization

Introduction

Goal : initialization of the Bulletin Marine Régulier (BMR)

Applicant : French Polynesia forecasters team

Project state : start on sept. 2024, developpments still in progress.



17 areas are covered by the BMR in French Polynesia.

Description générale de la situation.

05/09/2024 à 12:00UTC.

Régime d'alizé généralisé, faible à modéré sur l'ensemble du Pays.

Zone A35 , Est Marquises / Ua Huka - Hiva Oa

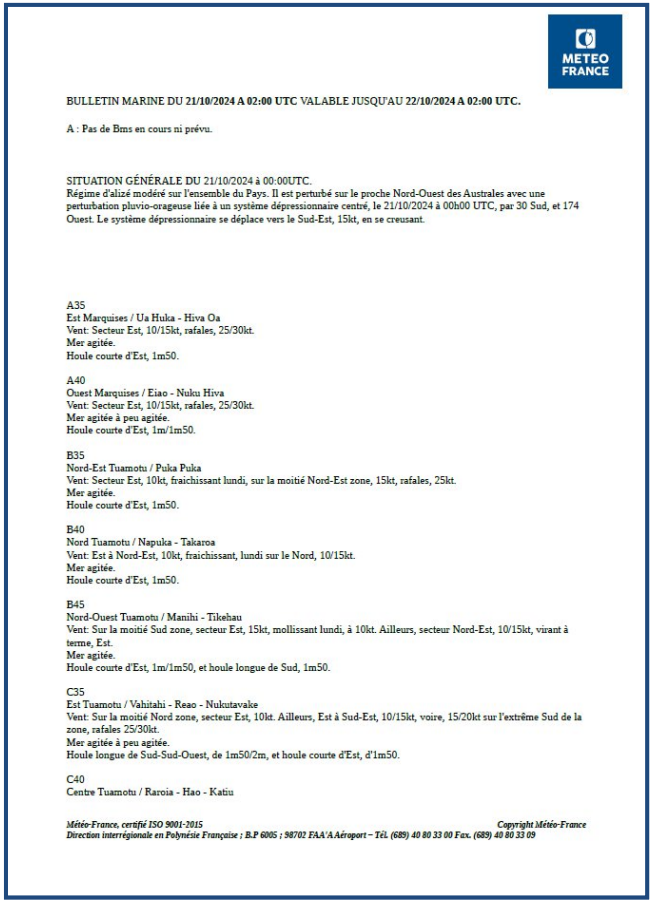
Est, 15kt, rafales 25/30kt.

Mer agitée.

Houle courte de Sud-Est, 2m.

Example of a BMR for a single area.

Textual data : a long archive of reports



METEO FRANCE

BULLETIN MARINE DU **21/10/2024 A 02:00 UTC** VALABLE JUSQU'AU **22/10/2024 A 02:00 UTC.**

A : Pas de Bms en cours ni prévu.

SITUATION GÉNÉRALE DU 21/10/2024 à 00:00UTC.
Régime d'alizé modéré sur l'ensemble du Pays. Il est perturbé sur le proche Nord-Ouest des Australes avec une perturbation pluvio-orageuse liée à un système dépressionnaire centré, le 21/10/2024 à 00h00 UTC, par 30 Sud, et 174 Ouest. Le système dépressionnaire se déplace vers le Sud-Est, 15kt, en se creusant.

A35
Est Marquises / Ua Huka - Hiva Oa
Vent: Secteur Est, 10/15kt, rafales, 25/30kt.
Mer agitée.
Houle courte d'Est, 1m50.

A40
Ouest Marquises / Eiao - Nuku Hiva
Vent: Secteur Est, 10/15kt, rafales, 25/30kt.
Mer agitée à peu agitée.
Houle courte d'Est, 1m/1m50.

B35
Nord-Est Tuamotu / Puka Puka
Vent: Secteur Est, 10kt, fraichissant lundi, sur la moitié Nord-Est zone, 15kt, rafales, 25kt.
Mer agitée.
Houle courte d'Est, 1m50.

B40
Nord Tuamotu / Napuka - Takaroa
Vent: Est à Nord-Est, 10kt, fraichissant, lundi sur le Nord, 10/15kt.
Mer agitée.
Houle courte d'Est, 1m50.

B45
Nord-Ouest Tuamotu / Manihi - Tikehau
Vent: Sur la moitié Sud zone, secteur Est, 15kt, mollissant lundi, à 10kt. Ailleurs, secteur Nord-Est, 10/15kt, virant à terme, Est.
Mer agitée.
Houle courte d'Est, 1m/1m50, et houle longue de Sud, 1m50.

C35
Est Tuamotu / Vahitahi - Reao - Nukutavake
Vent: Sur la moitié Nord zone, secteur Est, 10kt. Ailleurs, Est à Sud-Est, 10/15kt, voire, 15/20kt sur l'extrême Sud de la zone, rafales 25/30kt.
Mer agitée à peu agitée.
Houle longue de Sud-Sud-Ouest, de 1m50/2m, et houle courte d'Est, d'1m50.

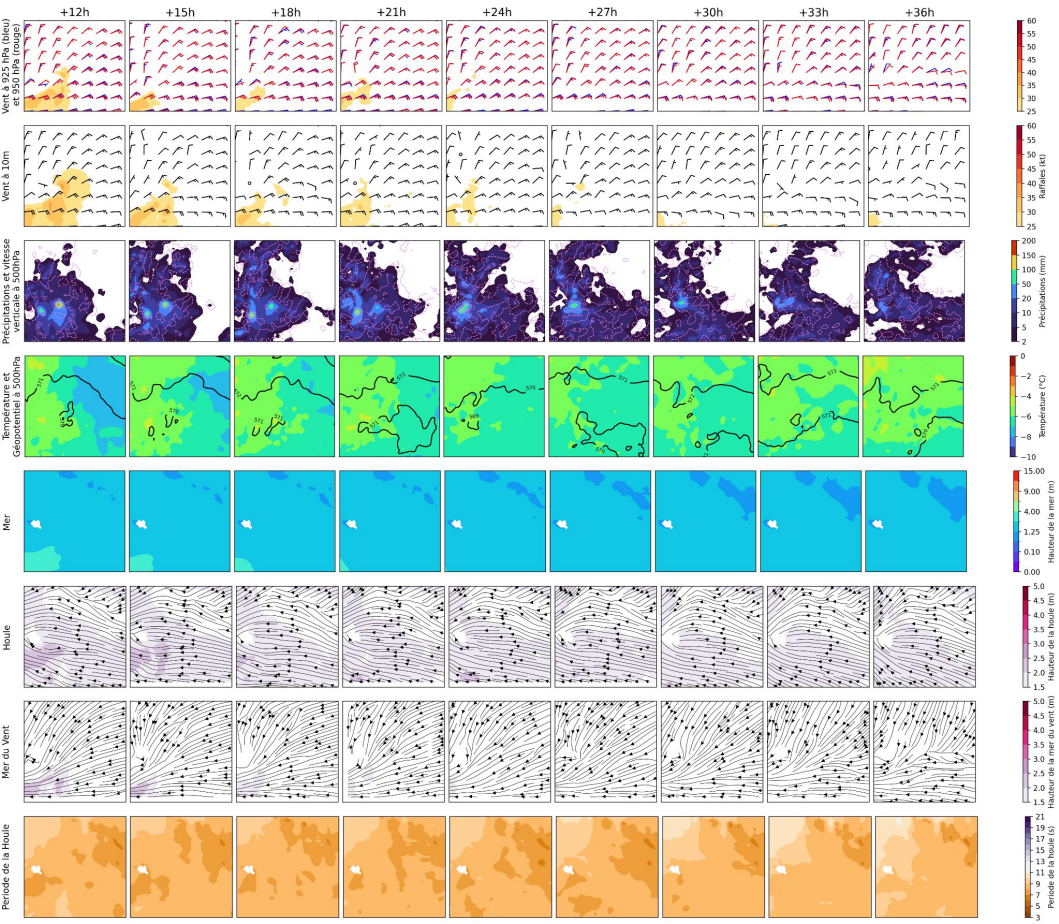
C40
Centre Tuamotu / Raroia - Hao - Katiu

Direction interrégionale en Polynésie Française ; B.P 6005 ; 98702 FAA'A Aéroport – Tél. (689) 40 80 33 00 Fax. (689) 40 80 33 09

- 7-years archive
- 2 report per day per area
- 17 areas in French Polynesia

→ **~ 80 000 training samples**

NWP models



9 leadtimes

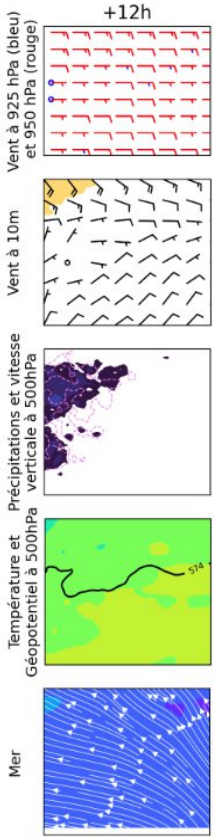
0.1° résolution

23 weather params selected by forecasters :

- **IFS** : wind (multiple levels), wind gusts (10m), precipitation, pressure, temperature, vertical velocity, …
- **MFWAM-IFS** : Hauteur et période de la mer totale et houle primaire et secondaire, …

Methodology

Weather inputs
(11, 23, 51, 51)



+

Some textual inputs (date, hour, ...)

→

Mutlimodal Langage Model

→

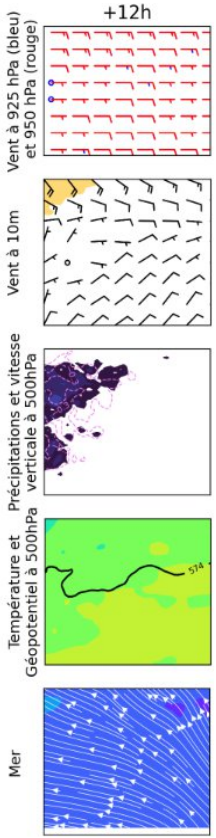
Est, 15kt, rafales 25/30kt.

Mer agitée.

Houle courte de Sud-Est, 2m.

Methodology

Weather inputs
(11, 23, 51, 51)



+

Some textual inputs (date, hour, ...)

+

Last BMR (-12h)

→

Mutlimodal Langage Model

→

Est, 15kt, rafales 25/30kt.

Mer agitée.

Houle courte de Sud-Est, 2m.

First results

Bulletins made by our best model

BMR du Dimanche 2025/01/12 05:00 pour la zone B35.

---------- Ground Truth: ----------

Vent : Nord à Nord-Ouest, 10/15kt, rafales 30/35kt sous grains.
Mer : Mer agitée.
Houle : Houle longue de Nord-Ouest, 1m50.
Temps sensible : Averses ou grains isolés sur le Sud de la zone.

---------- Model output: ----------

Vent : Nord à Nord-Ouest, 15/20kt, rafales 30/35kt sous grains.
Mer : Mer agitée.
Houle : Houle longue de Nord-Ouest, 1m50.
Temps sensible : Grains isolés, sur le Sud de la zone.

First results

Bulletins made by our best model

BMR du Samedi 2025/03/08 16:00 pour la zone A35.

---------- Ground Truth: ----------

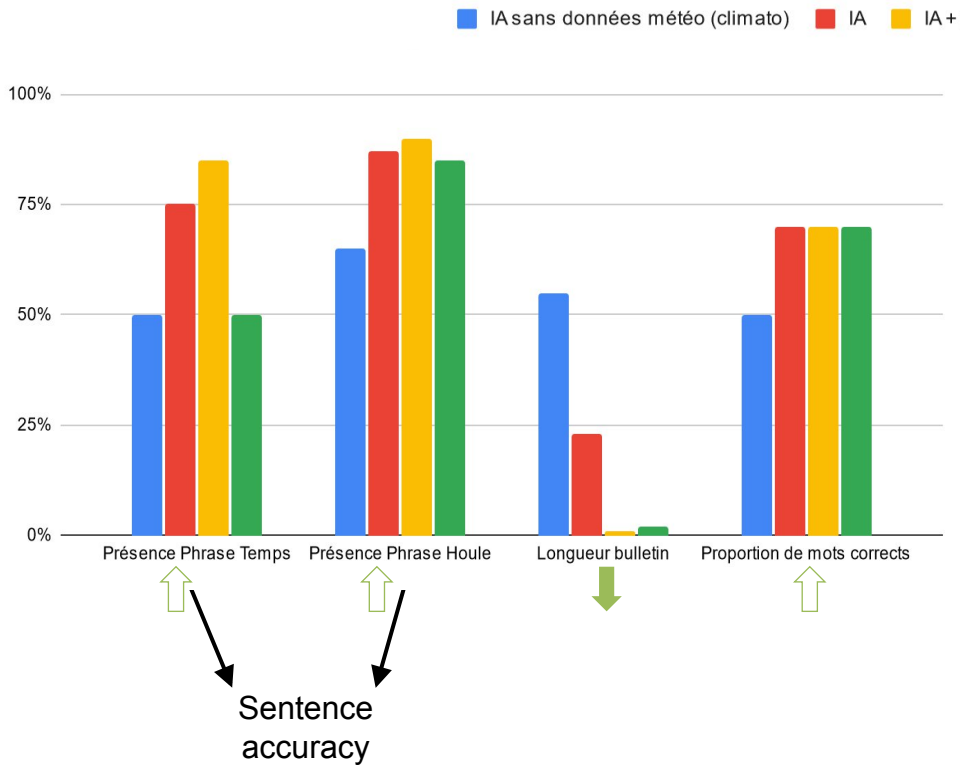
Vent : Est, 10/15kt, mollissant dimanche 05/10kt. Rafales de 25kt, sous grains.
Mer : Mer peu agitée à agitée.
Houle : Houle de Nord, 1m50, sur la zone Nord.
Temps sensible : Averses et grains gagnant progressivement la moitié Nord-Est de la zone.

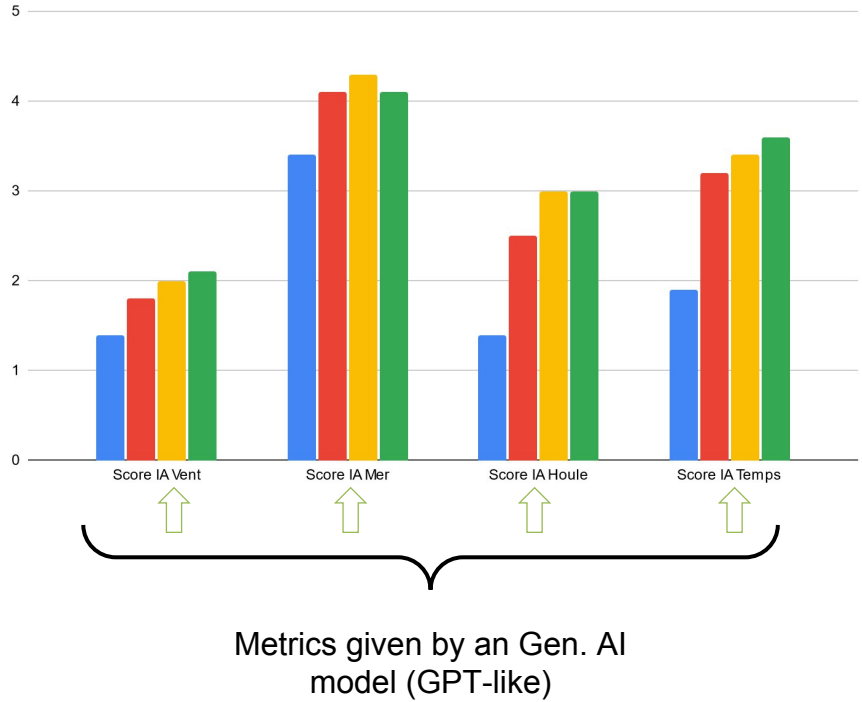
---------- Model output: ----------

Vent : Est, 10/15kt.
Mer : mer peu agitée à agitée.
Houle : Rien à signaler.
Temps sensible : Rien à signaler.

First approach: scores

Scores computed over 320 reports





Forecasters' feedback

Sharing a real-time prototype without a previous bulletin with French Polynesia forecasters:

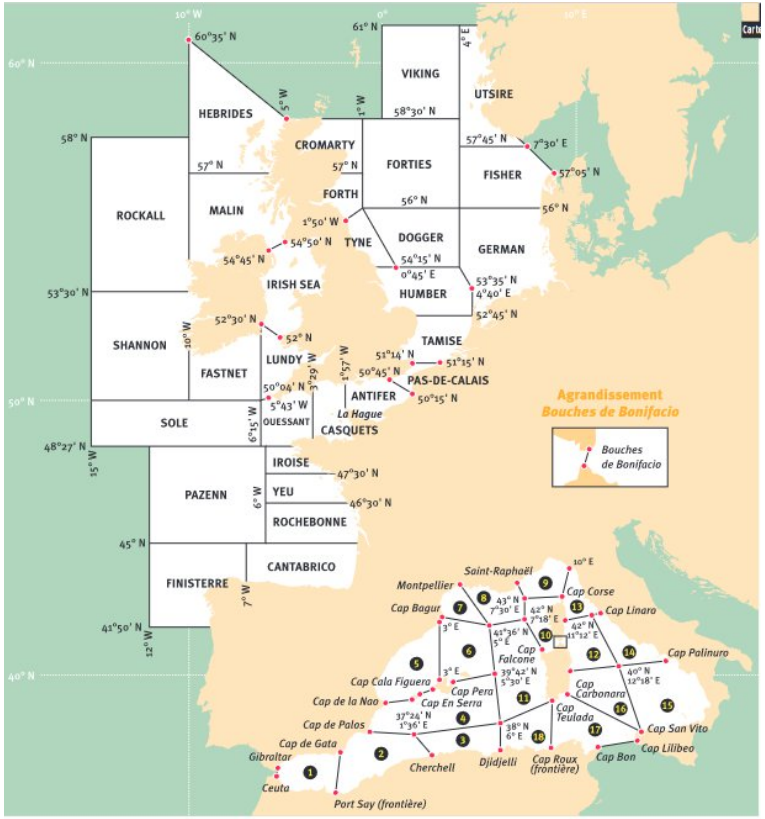
- Remarkable weather setence too often incorrectly mentioned
- Swell sometimes minimized when bathymetry masks swell propagation depending on the area
- The model has not properly learned the order of the days of the week and their timing.

For now, the AI model wrote **realistic reports**, but they do not always match the **weather situation.**
The generated bulletins do **not allow forecasters to save time when writing them.**

Coming soon: sharing of a prototype including the previous bulletin in the input data.

Next steps

- Tweaks to the training method
- Training on Polynesia and Metropolitan data:
  - + 50 zones * 10 years → increases the dataset 10 folds



RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

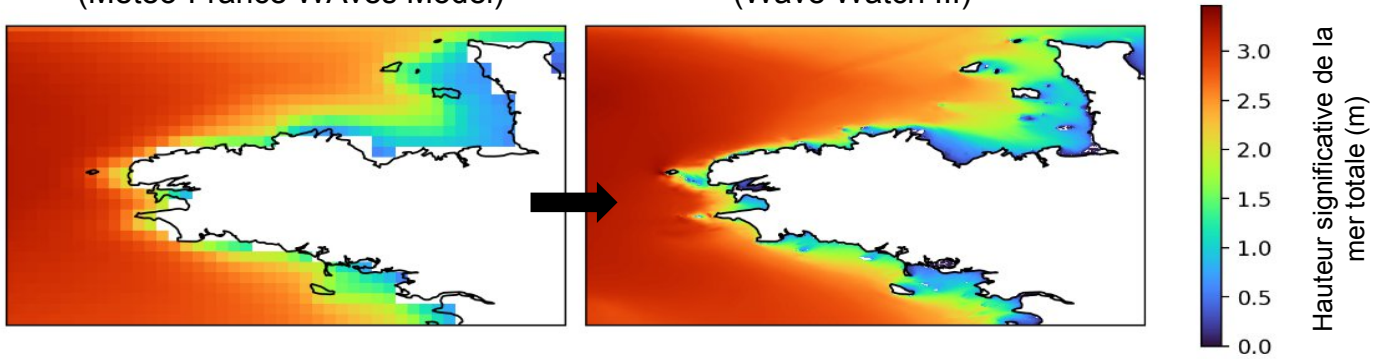
Downscaling Wave models

by Mathilde Ferreira

Wave models

MFWAM
(Météo-France WAves Model)

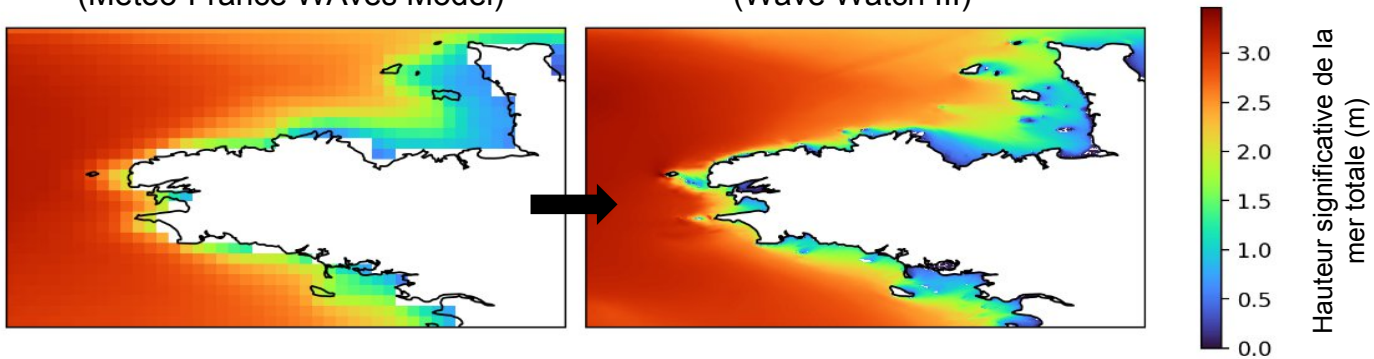
WW3
(Wave Watch III)



Introduction

MFWAM
(Météo-France WAves Model)

WW3
(Wave Watch III)



Model	Domain	Resolution	# grid points	# Leadtimes	Computing time	# CPUs	hour.CPU
MFWAM	Global	0.1°	2 800 000	66 (+102h)	45 min	1964	1473
WW3	France	200 m	180 000	56 (+72h)	45 min	320	240

➡ WW3 is **more expensive** than MFWAM

Goal: to perform a downscaling from MFWAM to WW3 to develop an **ensemble version** with 35 members, run 2 times a day at very low cost.

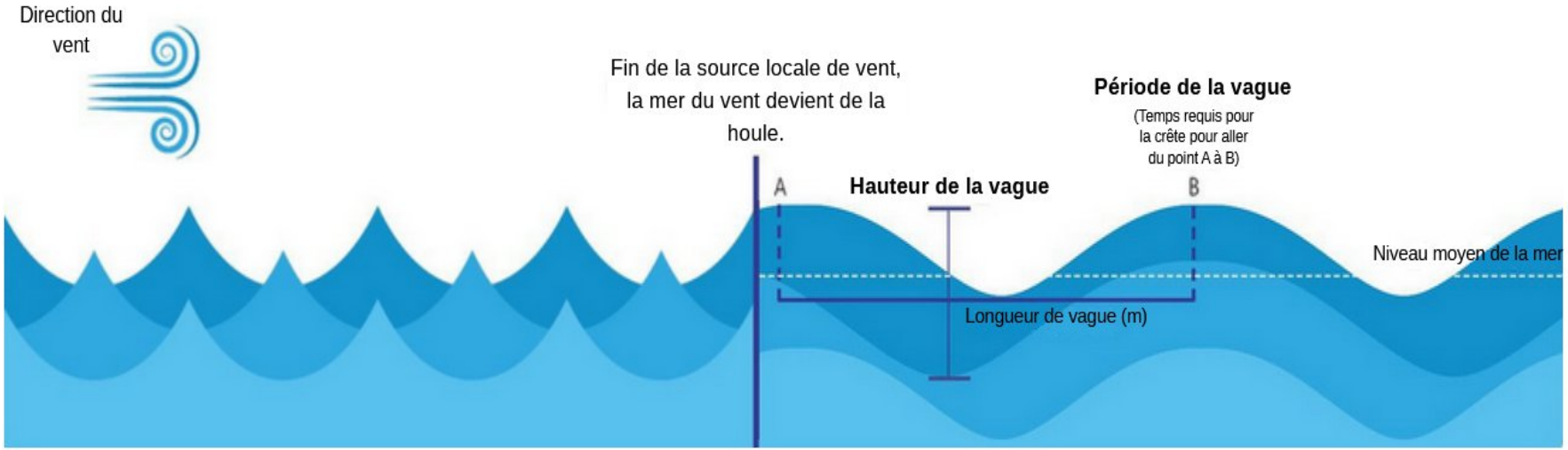
Wave parameters

MER DU VENT

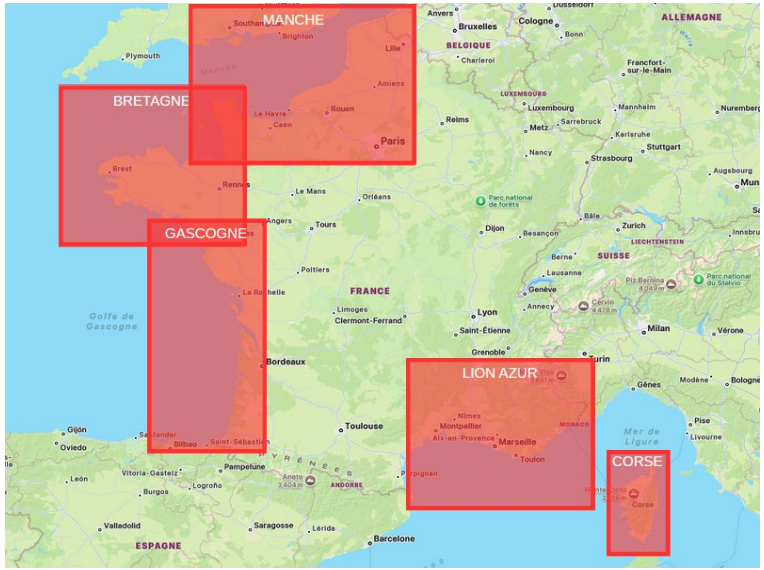
La mer du vent est générée immédiatement par le vent local. Elle n'est pas autosuffisante et s'arrête lorsque le vent s'arrête.

HOULE

La houle est autosuffisante et générée par une énergie sous la surface de l'océan, plus besoin de vent local.



Project perimeter



Goal : Train a ML model for each **zone**.
5 zones : Bretagne, Corse, Gascogne, Manche, Golfe du Lion.
→ The Bretagne zone is the first one to be studied.

Training periode spreads over the year 2023 :
- 9 months for training
- 3 months for testing and validation

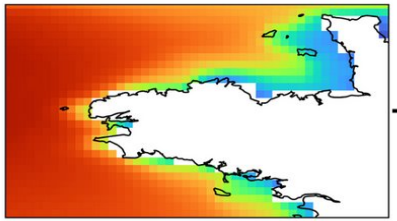
Goal : downscaling from **10km to 200m** (x50)

State of the art: x4 à x10

Strategy: Start by training a 2x model then increase downscaling ratio for subsequent models.

Method: classic

Prévisions
MFWAM 0.1°



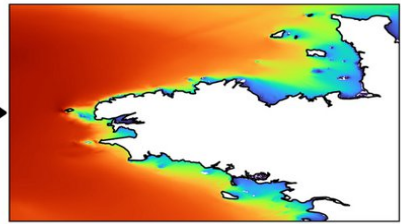
9 paramètres de vagues
(directions en sin/cos)

Interpolation vers 0.002°

Normaliser

Réseau de neurones

Analyses
WW3 0.002°

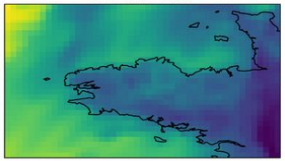


9 paramètres de vagues
(directions en sin/cos)

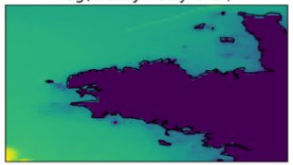
Normaliser

Forcages

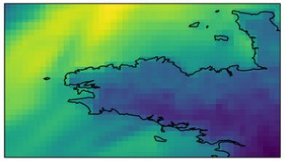
ARPEGE U10



log(Bathymetry + 1)



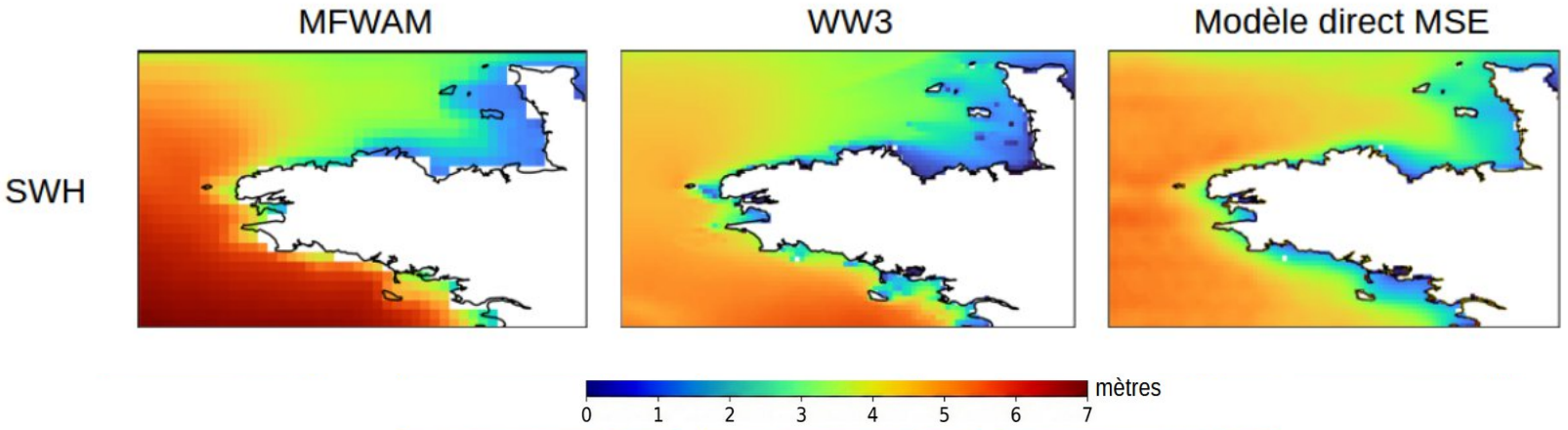
ARPEGE V10



Land / Sea mask

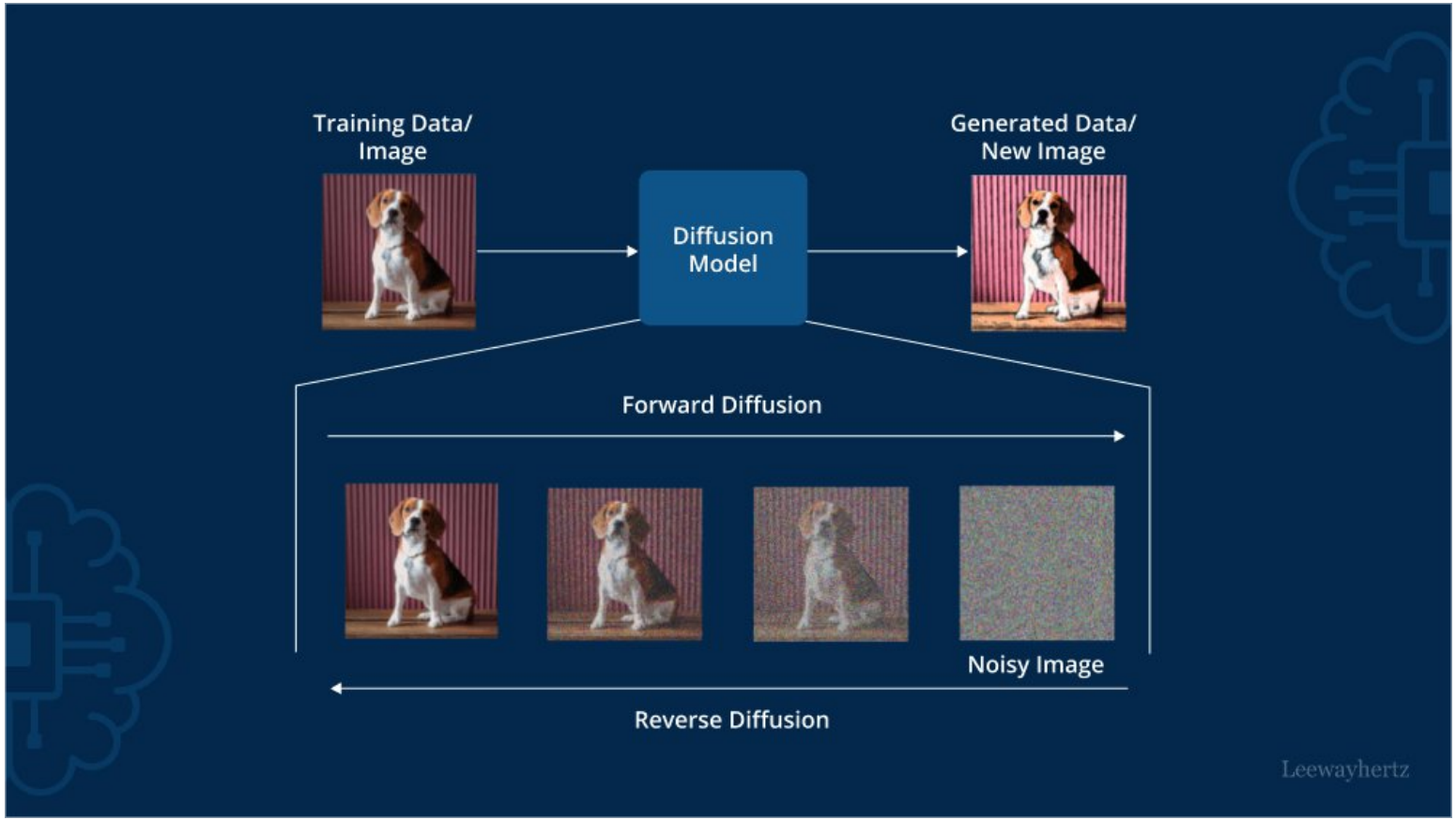
(Concaténation de canaux en input)

Classical method results



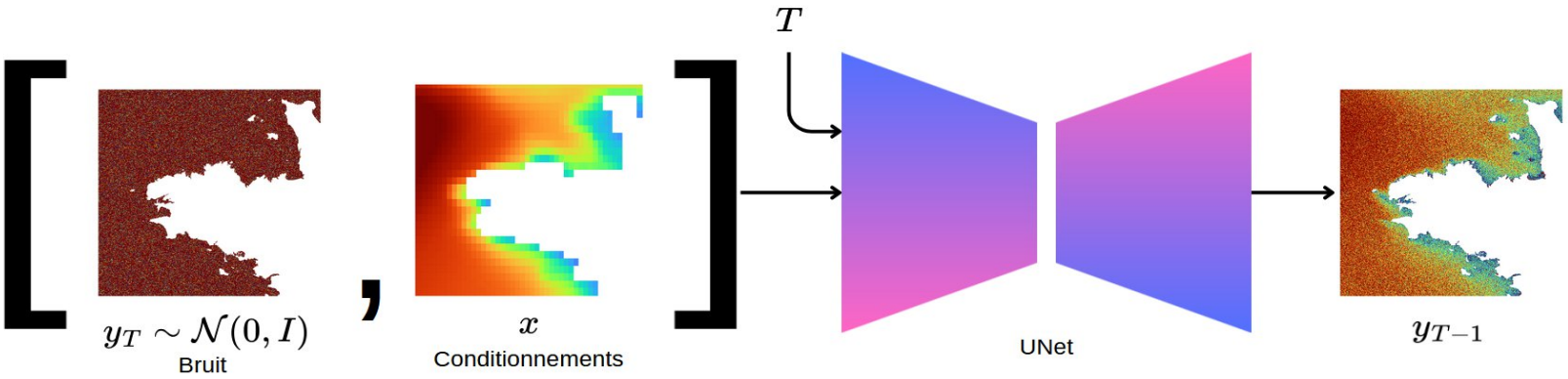
→ L'approche directe donne des résultats très **lissés**

Method : Diffusion

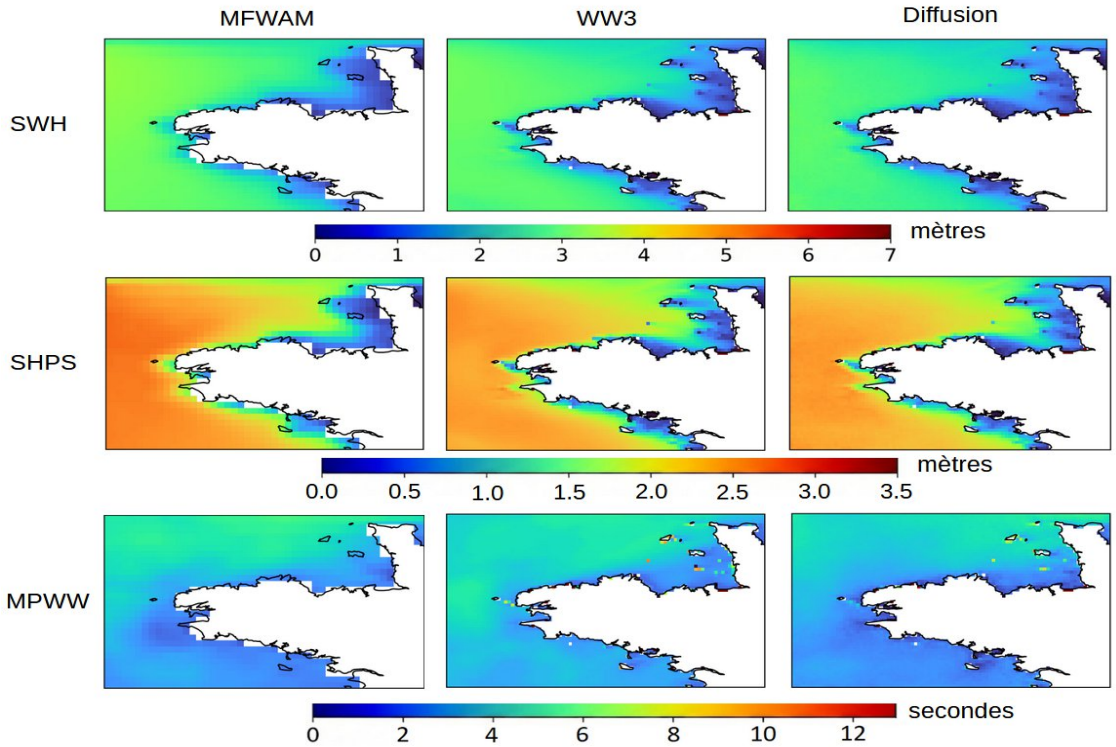


Méthodologie : Approche par diffusion

Inference : From a fully noise image, successive de-noising steps builds up the final image.

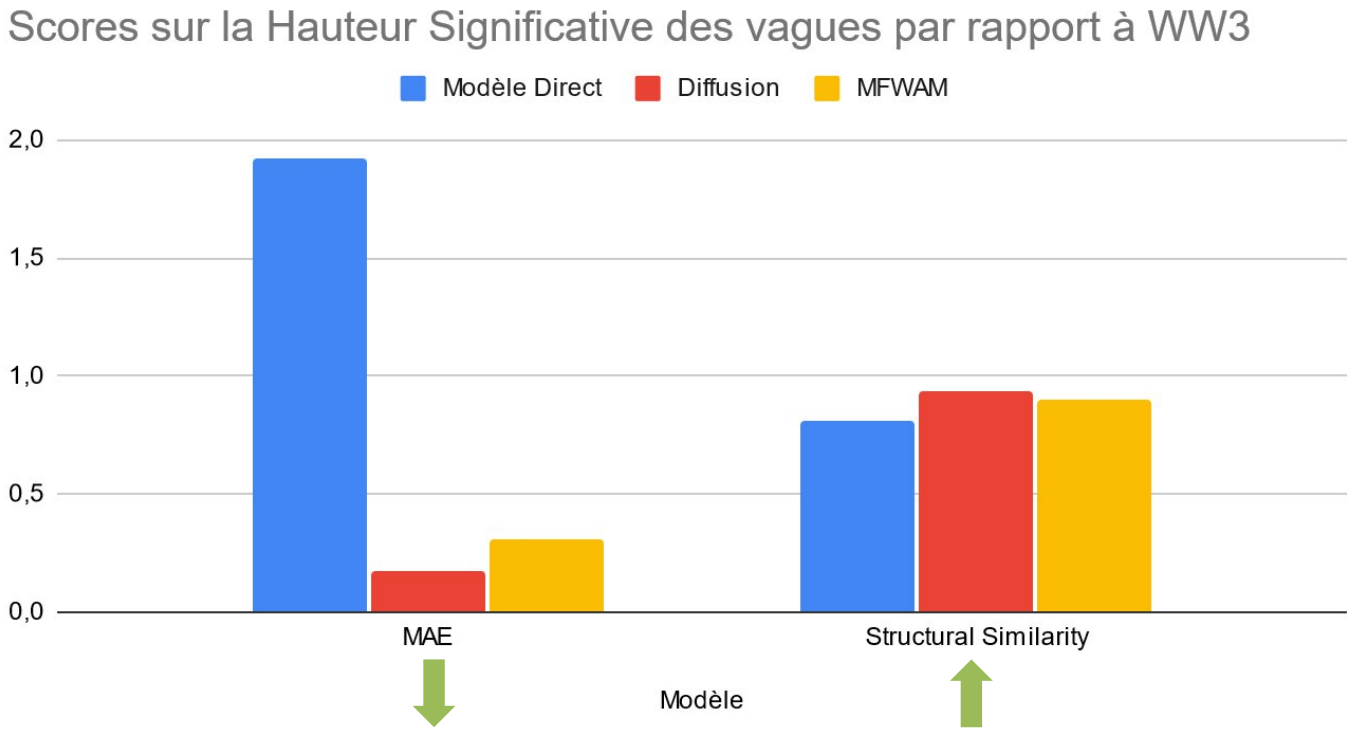


Diffusion results



→ The diffusion model is much more detailed.

Résultats quantitatifs



Comparaison des temps de calcul

Required compute time for forcasts up to +102h

Model	Emprise géographique	Temps de calcul	Nombre de CPU	Heures.CPU
Approche classique	Bretagne, res. 5km	20 sec	8	0.05
Diffusion	Bretagne, res. 5km	8 min	8	1
WW3	Métropole, res. 200m	45 min	320	240

Conclusion

→ **The 5km diffusion model has been validated**

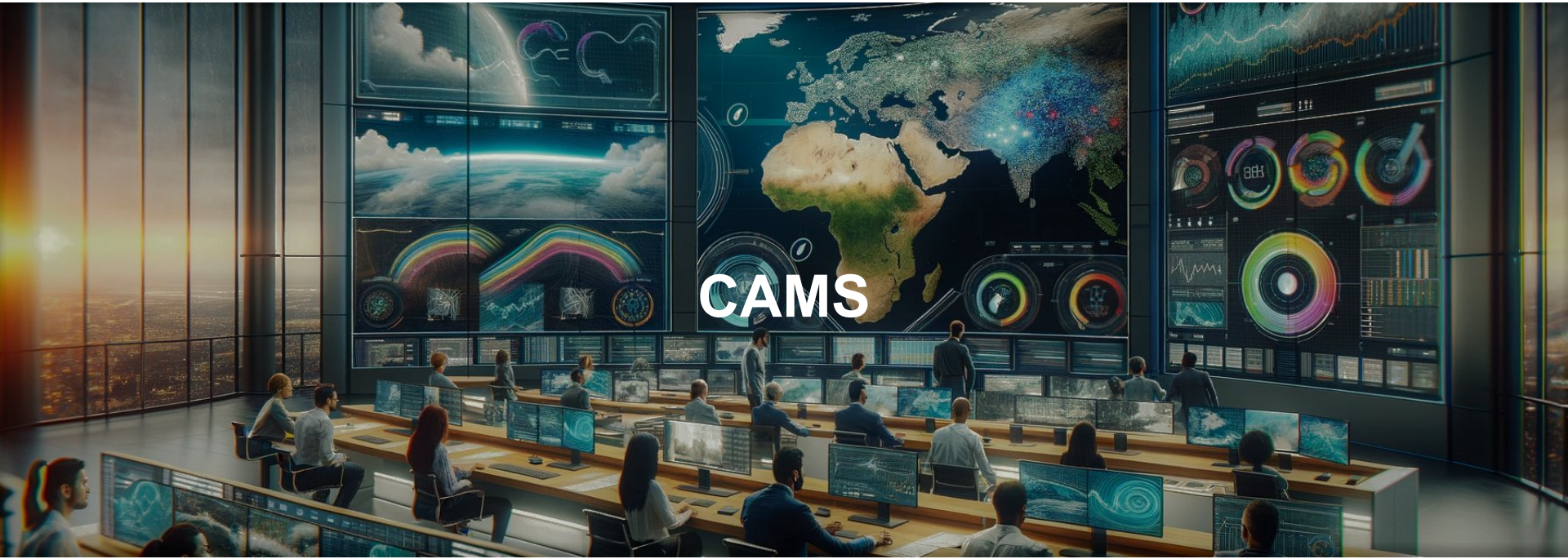
Perspectives :

- Add sea currents as inputs of models
- Validate model's performances on extreme storms
- Increase downscaling factor up to x10 to reach **1km resolution**.

meteofrance/wave-downscaling

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE



Context Goal

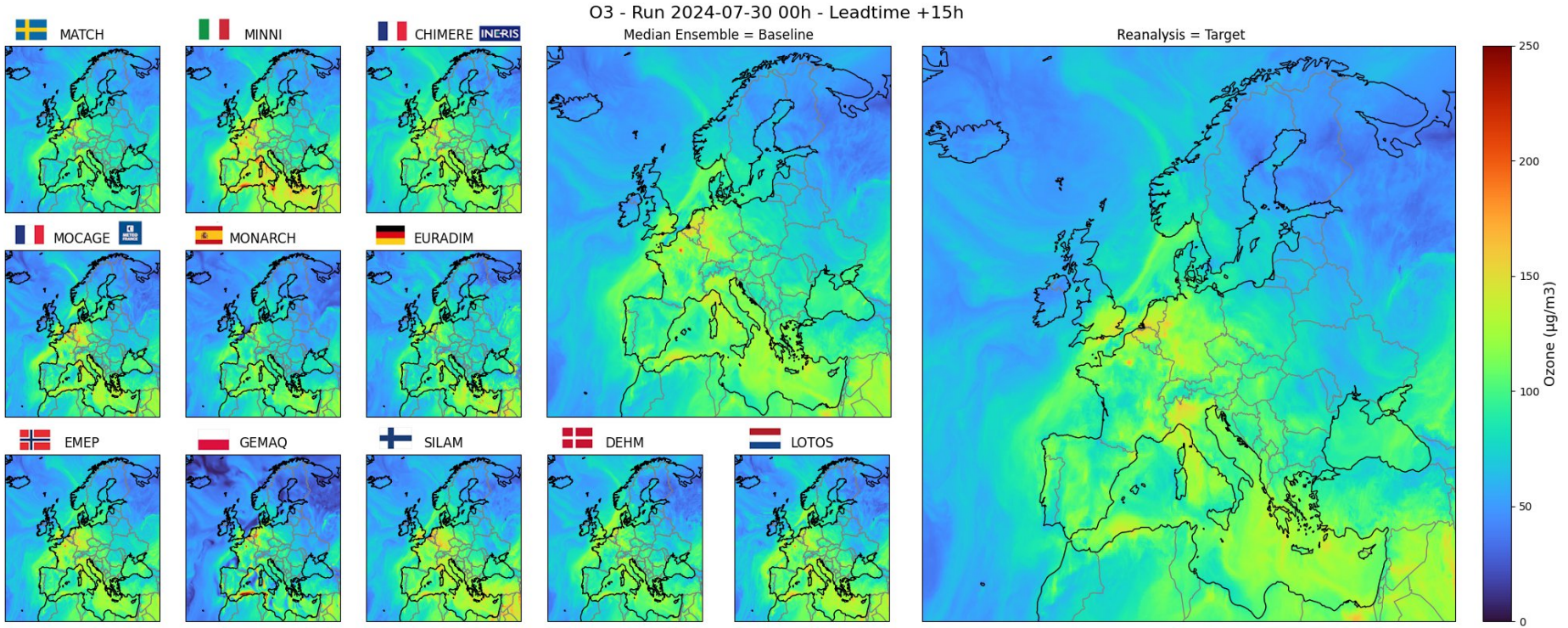
Context : Météo-France participate in european “Copernicus Atmosphere Monitoring Service” (CAMS) for air quality forecasts.

Today:, 11 national center provide 11 regional air quality forecast models and the ensemble forecast is the median values of the 11 models.

Goal : Use deep learning techniques combining all 11 forecasts to improve on current ensemble median model.

Collaboration : DSM/CS/ENV & CNRM/GMGEC/PLASMA & INERIS.

Data



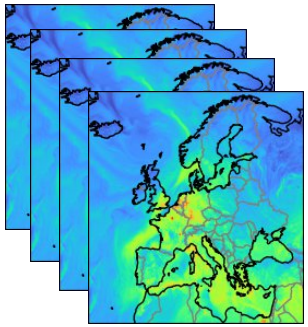
- Forecast models : 11
- Vertical levels : 10
- Species : 22

- Lead times : up to +4 days
- Data history: 3 years

First experiments

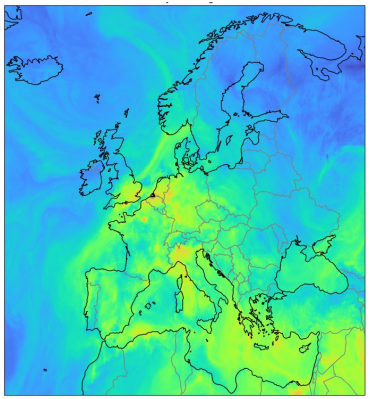
11 forecasts
Unique:
– species : Ozone
– lead time : +15h
– vertical level : surface

11 forecasts



Neural network

Reanalysis



GitHub

meteofrance/cams-dl-ensemble

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

METEO FRANCE

We'll be presenting MFAI, our Python library,
tomorrow at 4.30pm

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

Questions ?

dsm_labia@meteo.fr