PROGRAMME OF THE EUROPEAN UNION

Copernicus Europe's eyes on Earth

Atmosphere Monitoring Service atmosphere.copernicus.eu

IMPLEMENTED BY ECMWF

#EUSpace

CAMAERA

ML activities at CAMS and in the CAMAERA project

Enza Di Tomaso (ECMWF) on behalf of

Samuel Remy[1] & CAMAERA team

Nathan Capon[1], Rose-Cloé Meyer[1], Magdalena Anguelova[2], Melanie Ades[3], Jeronimo Escribano[4], Thierry Elias[1] and Antonino Bonanni[3]

[1] HYGEOS, [2] NRL, [3] ECMWF, [4] BSC

Paula Harder & CAMS and AIFS team

Johannes Flemming, Matthew Chantry, Mihai Alexe, Cathy Wing Yi Li, Luke Jones, Baudouin Rault, Gert Mertes, Florian Pinault, Juan Jose Dominguez, Richard Engelen

ECMWF

PROGRAMME OF THE EUROPEAN UNION

CAMAERA is a **H**orizon **E**urope CAMS advancement project which aims to **enhance the quality of the aerosol products** delivered by CAMS through improvements of the modelling capabilities and data assimilation aspects of the CAMS regional and global systems.

ML activities for the IFS-COMPO

Two ongoing activities to implement ML in IFS-COMPO

- Replace selected processes by ML (Horizon Europe CAMAERA project). Focus on:
  - Whitecap fraction and sea-salt aerosol emissions (1/1/2024 – 30/6/2025)
  - Desert dust emissions (1/7/2025 – 31/12/2026)

Process-based ML

Part 1 of this talk

- Replace the whole model + DA by ML (ECMWF – Paula Harder)
  - A first prototype of AIFS-Compo

Full data-driven forecasting

Part 2 of this talk

Part 1

ESTIMATING WHITECAP FRACTION FOR SEA-SALT AEROSOL EMISSIONS IN IFS-COMPO

Current status of sea-salt aerosol emissions in cycle 49R1 IFS-COMPO:

- The whitecap fraction (WF) is estimated by the **Albert** et al. (2016) parameterization:

$$\boldsymbol{WF = a(SST)[WSP + b(SST)]^2}$$

- Sea-salt aerosol emissions are derived using the **Gong** (2003) assumed size distribution

$$(WSP, SST) \xrightarrow{\text{A16}} WF \xrightarrow{\text{Gong}} \textit{Sea-Salt}$$

Our objective : Estimation of whitecap fraction and sea-salt emissions in IFS-COMPO with deep neural networks (DNN) by :

1. Training offline a DNN model to estimate whitecap fraction
2. Integrating this DNN model into IFS-COMPO

INPUT AND TRAINING DATASETS OF THE OFFLINE DNN MODEL

Dataset description

Ground truth : Whitecap fraction (WF) at **10.7** and **37** GHz derived from remote sensing (Anguelova et al 2019)

Time range : 2 years of data with an hourly resolution

Predictors : 8 predictors collected

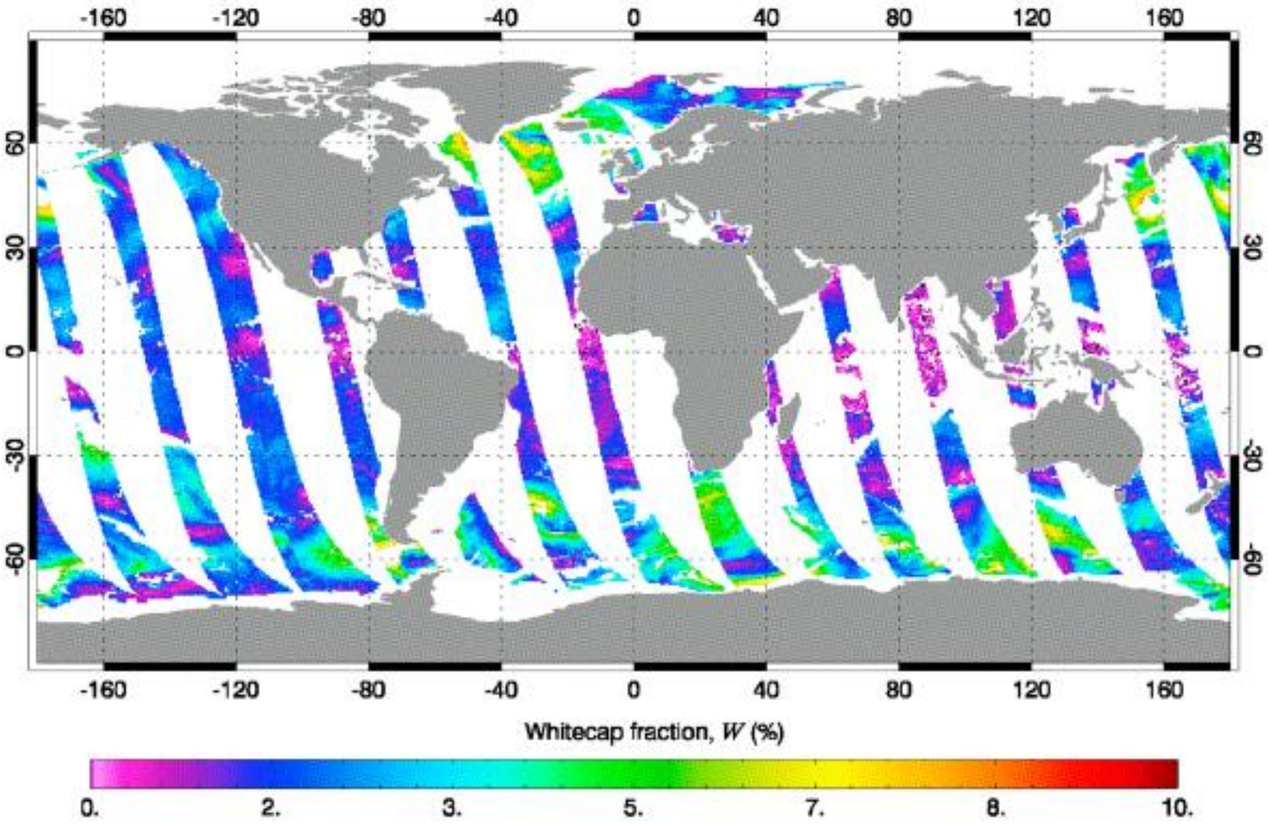
From ERA5 :

- Wind Speed
- Wind Direction
- Sea Surface Temperature
- Mean Wave Period
- Significant Wave Height

From HINDCAST :

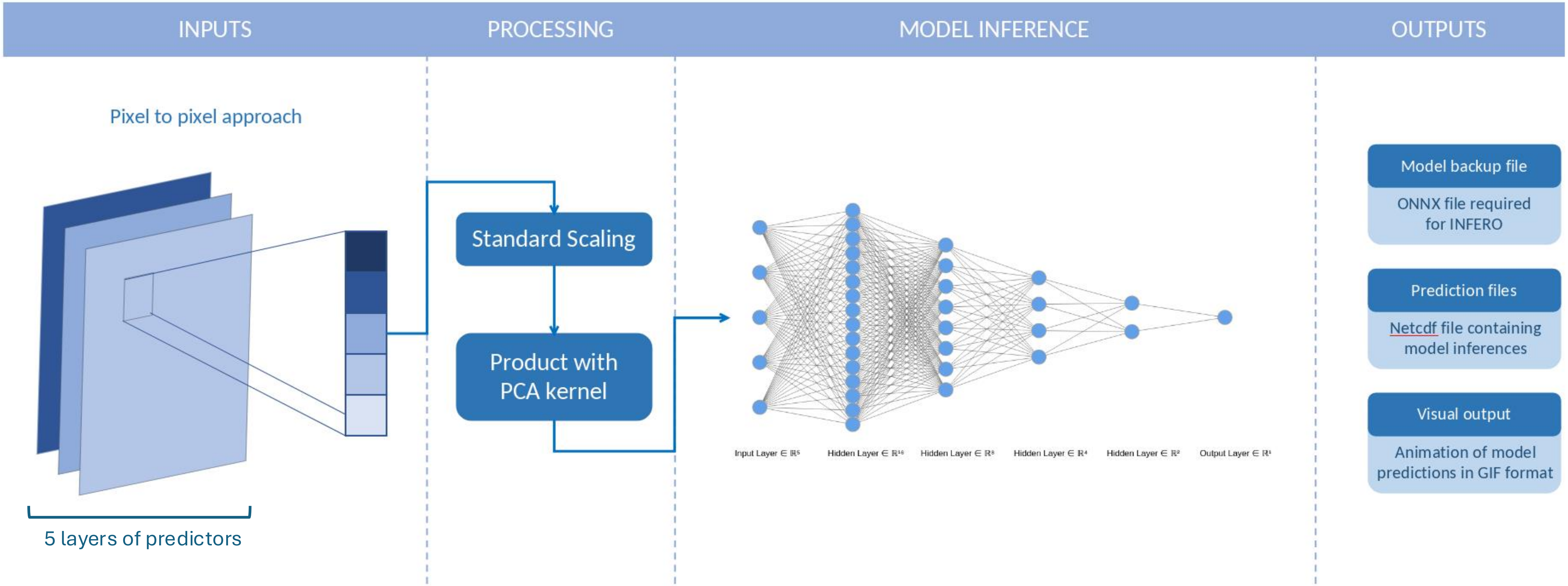
- Total Wave Height
- Significant Wave Height
- Dissipation of turbulent energy from breaking waves

Dimension : around 200 million pixels



Example of daily map of whitecap fraction from Windsat acquisition [Anguelova et al.]

OFFLINE DNN MODEL ARCHITECTURE

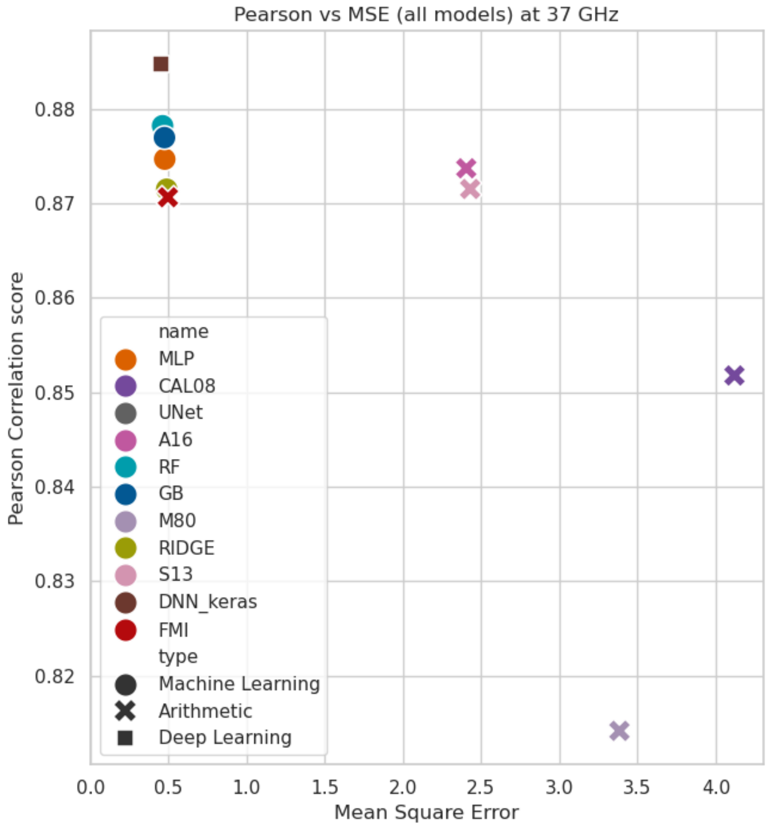


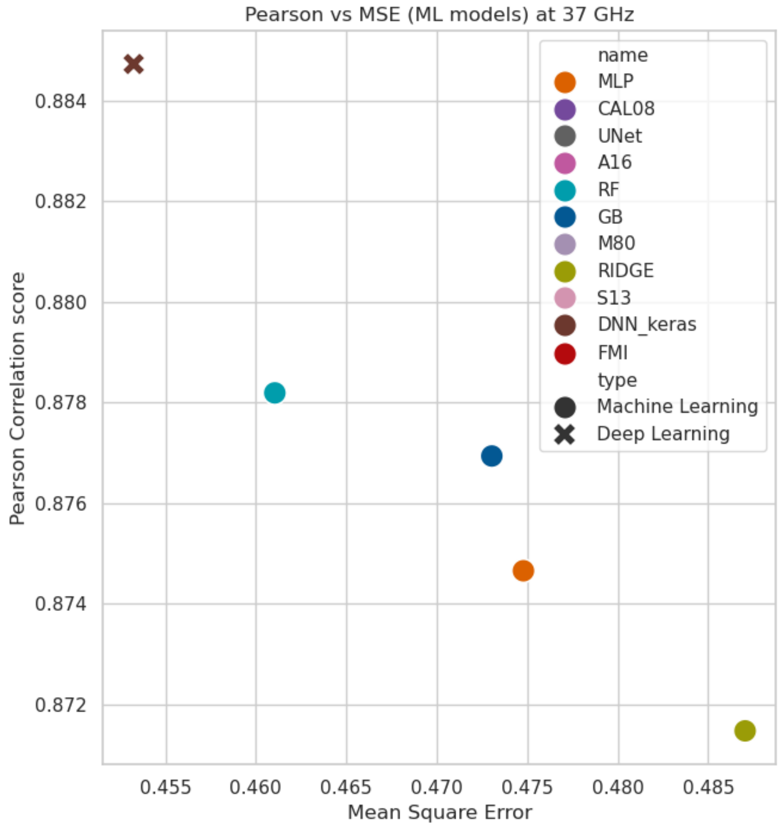
Description of the Deep Neural Network (DNN) model architecture and pre- and post-processing operations

RESULTS OF THE OFFLINE DNN MODEL AND COMPARISON WITH OTHER METHODS

Comparison of simulated and retrieved whitecap fraction – January June 2017:

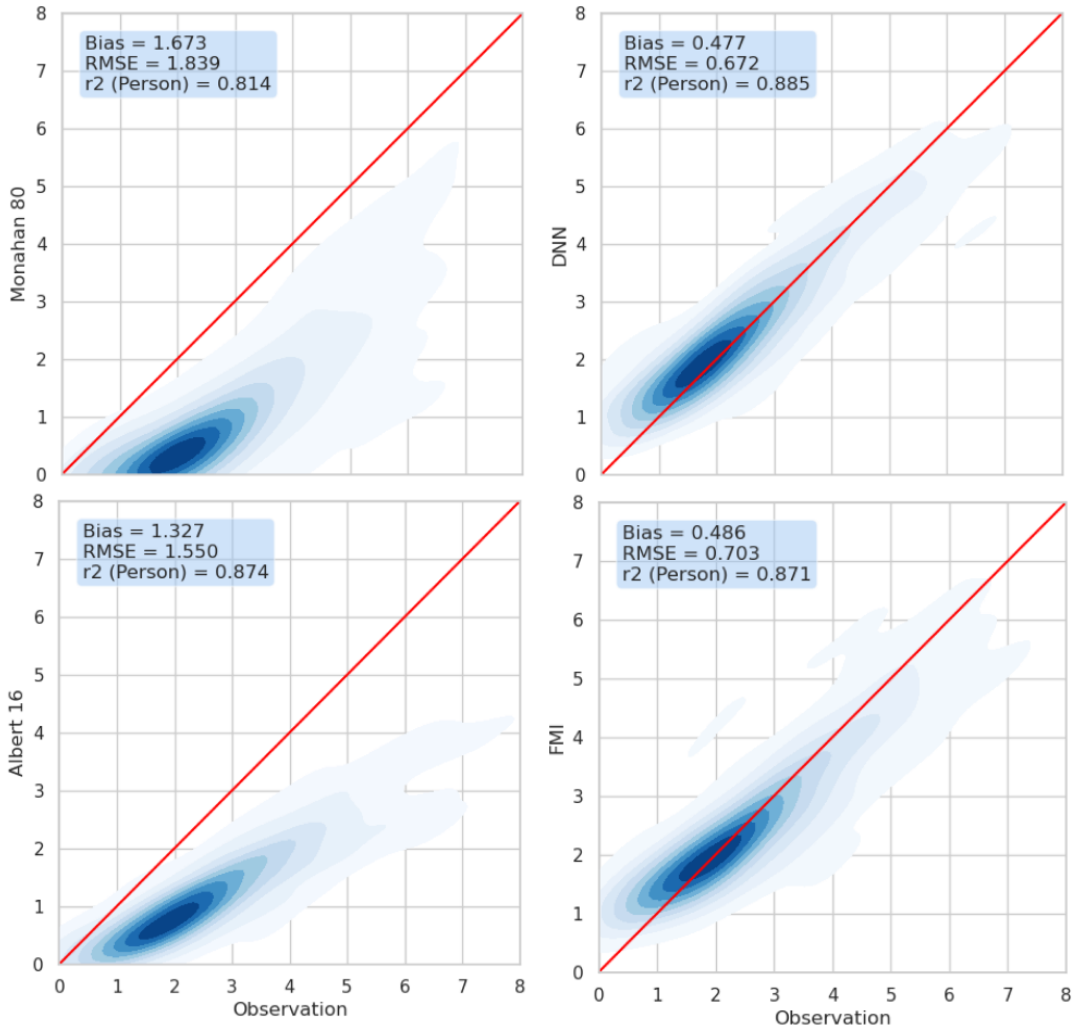
- New « FMI » arithmetic model developed in CAMAERA
- Arithmetics models (Monahan 80, Albert 16) show a **low bias** as compared to our dataset
- DNN **outperforms** arithmetic models
- DNN manages to score better than a simple neural network architecture (MLP)
- No dependency to SST found





RESULTS OF THE OFFLINE DNN MODEL AND COMPARISON WITH OTHER METHODS

DNN outperforms arithmetic models and a simple neural network architecture (MLP)



Simulated (y axis) versus observed (x axis) whitecap fraction (WF) in %

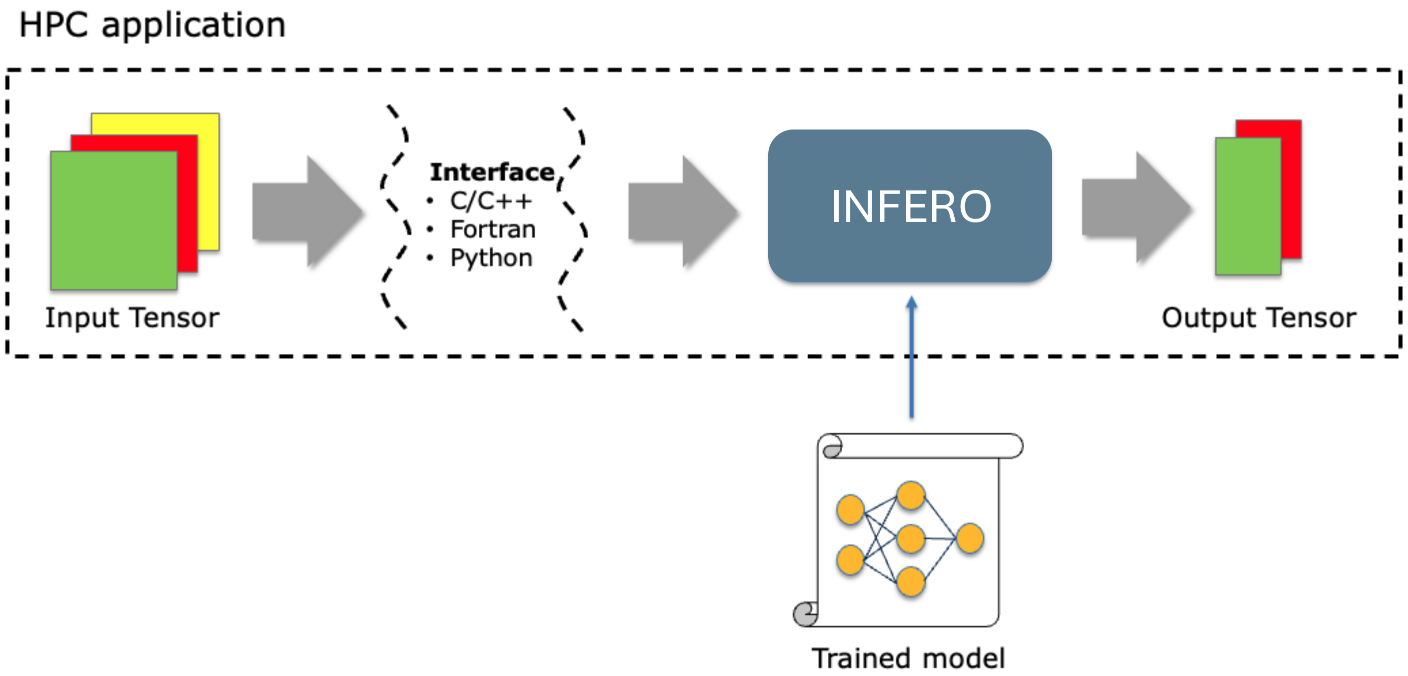
INTEGRATION OF THE DNN INTO IFS-COMPO

Incorporation of an **exported version (ONNX format)** of the DNN to compute whitecap fraction and sea-salt aerosol emissions online in IFS-COMPO.

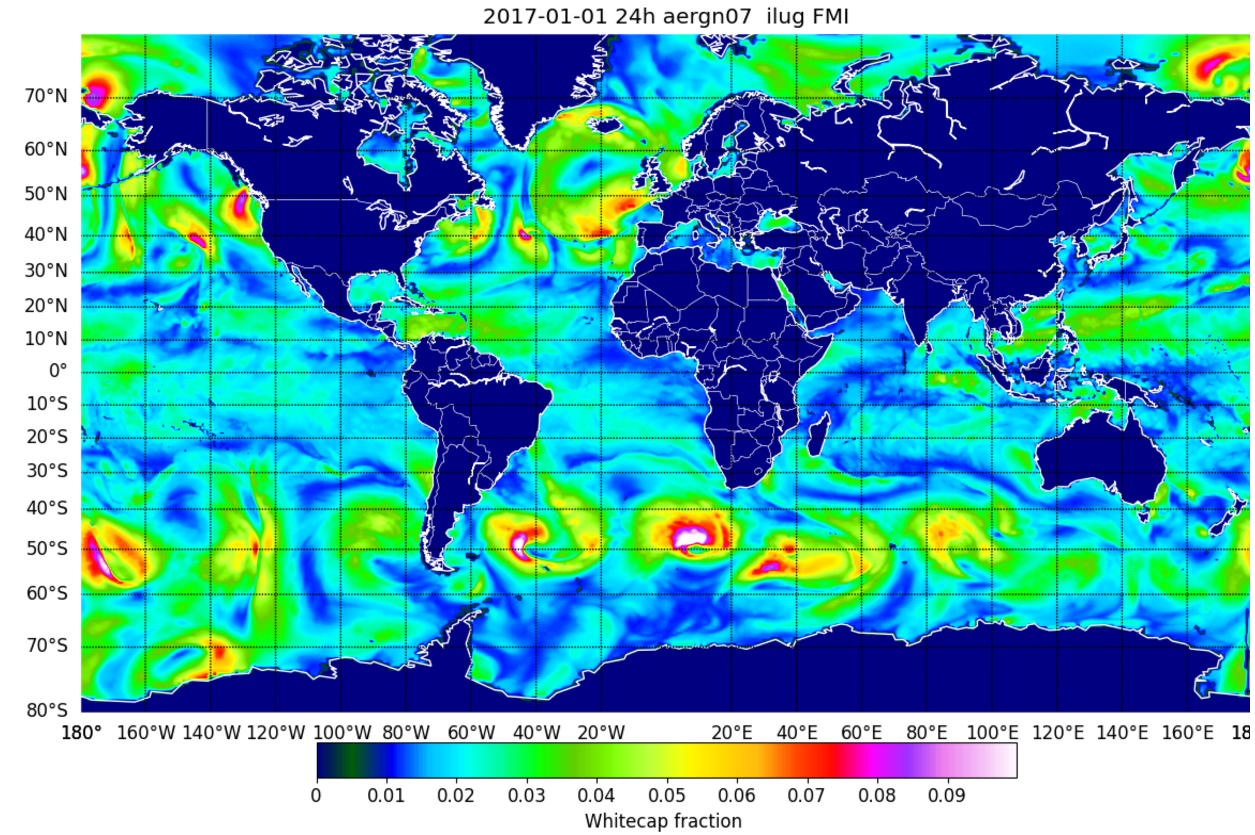
The **INFERO** library has been integrated into IFS-COMPO to interface with Deep Learning models

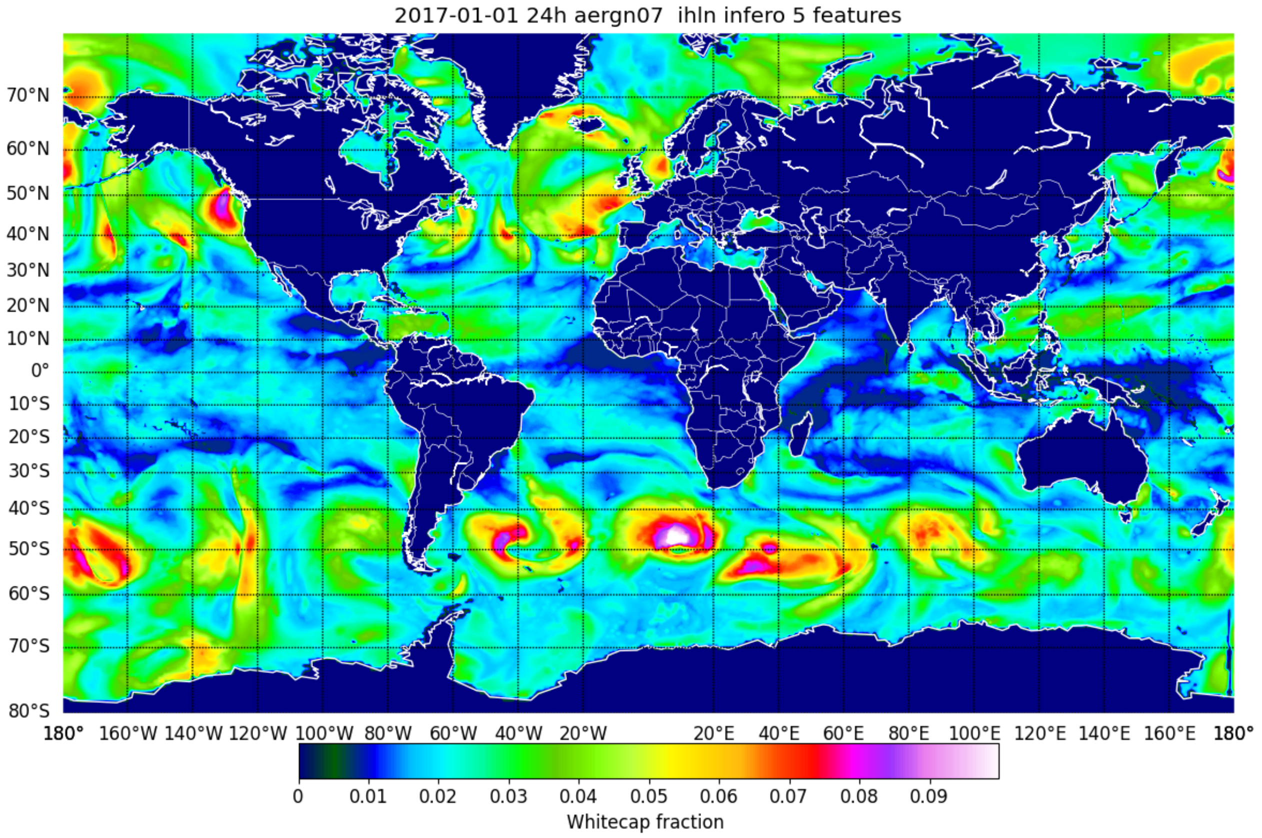
Interest : runs a learning model in ONNX format from a Fortran script

Representation of the incorporation of our model into IFS



EXAMPLE OF IFS-COMPO SIMULATED WHITECAP FRACTION



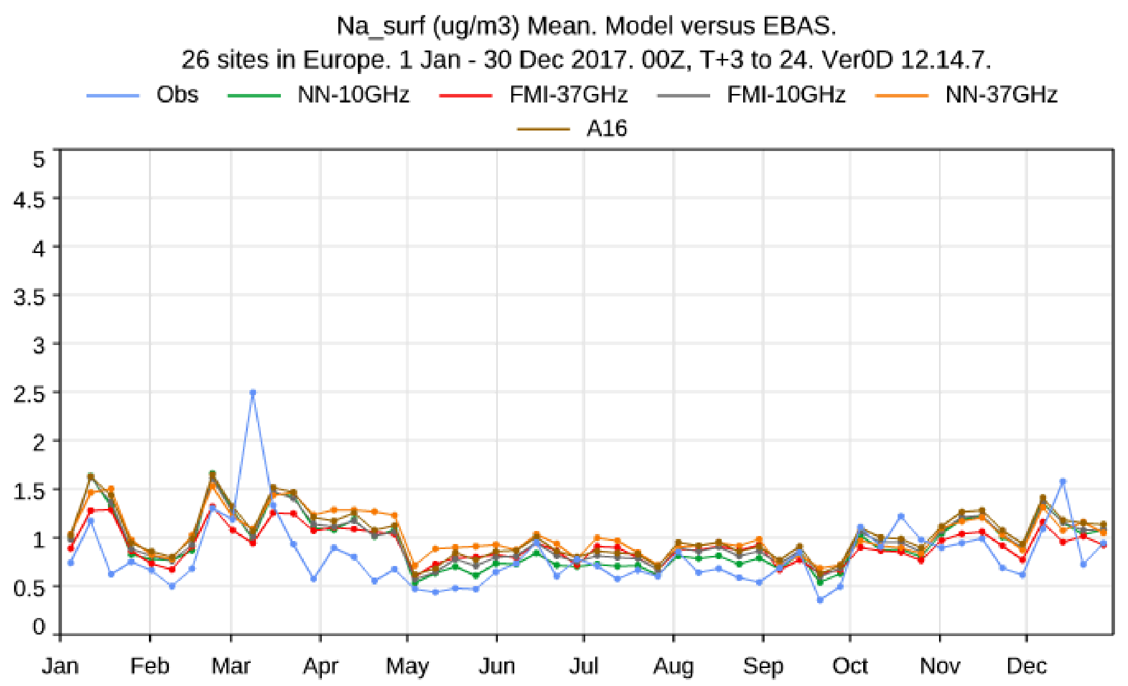


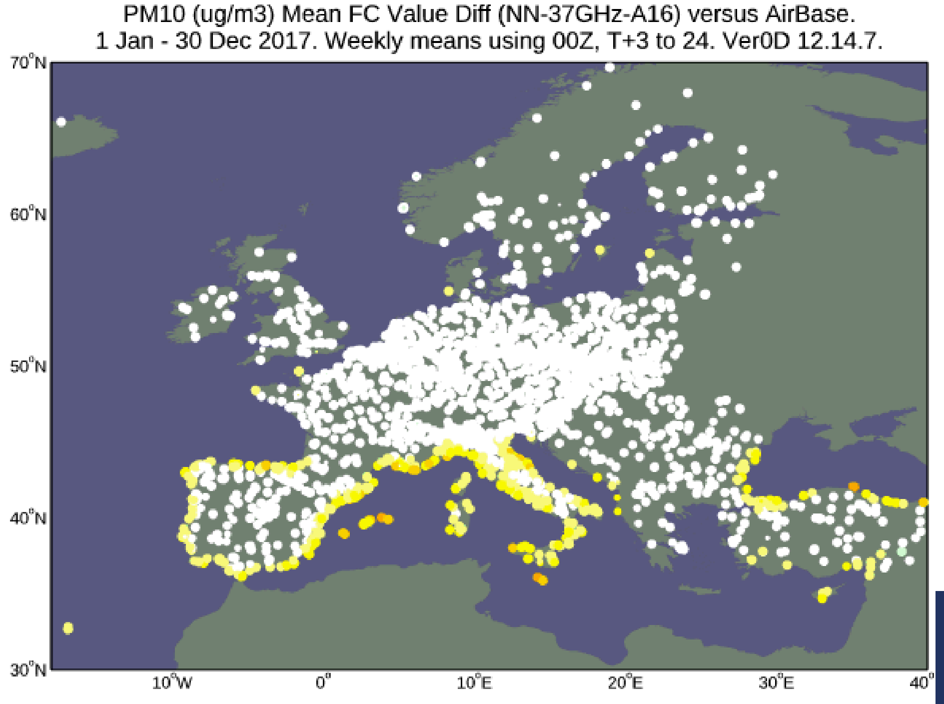
Simulated whitecap fraction by IFS-COMPO on 1/1/2017 0UTC, using the newly developed FMI scheme (left), and with deep learning model (using 5 predictors) enabled through the INFERO library (right).

EVALUATION

Forecast only simulations have been carried out for 2017, for A16 (operational setup), FMI and Neural network (NN). In general,

- The impact on simulated AOD is very small
- Some impact on coastal PM10
- Some impact on simulated surface concentration of Na/Cl over Europe, but no clear improvement





WHAT IS NEXT?

- Use of the same methodology for desert dust emissions
- Reference dataset : best estimate of daily IFS-COMPO dust emissions through offline inversion with an ensemble approach.

Assimilated DOD
VIIRS SNPP Deep Blue

IFS-COMPO Ensemble

20 members with perturbed emissions at the time

100 *active* members

LETKF (smoother)

Control: dust emission
- 3 days temporal resolution
- Model horizontal resolution

Observations: Dust AOD or AOD 550nm

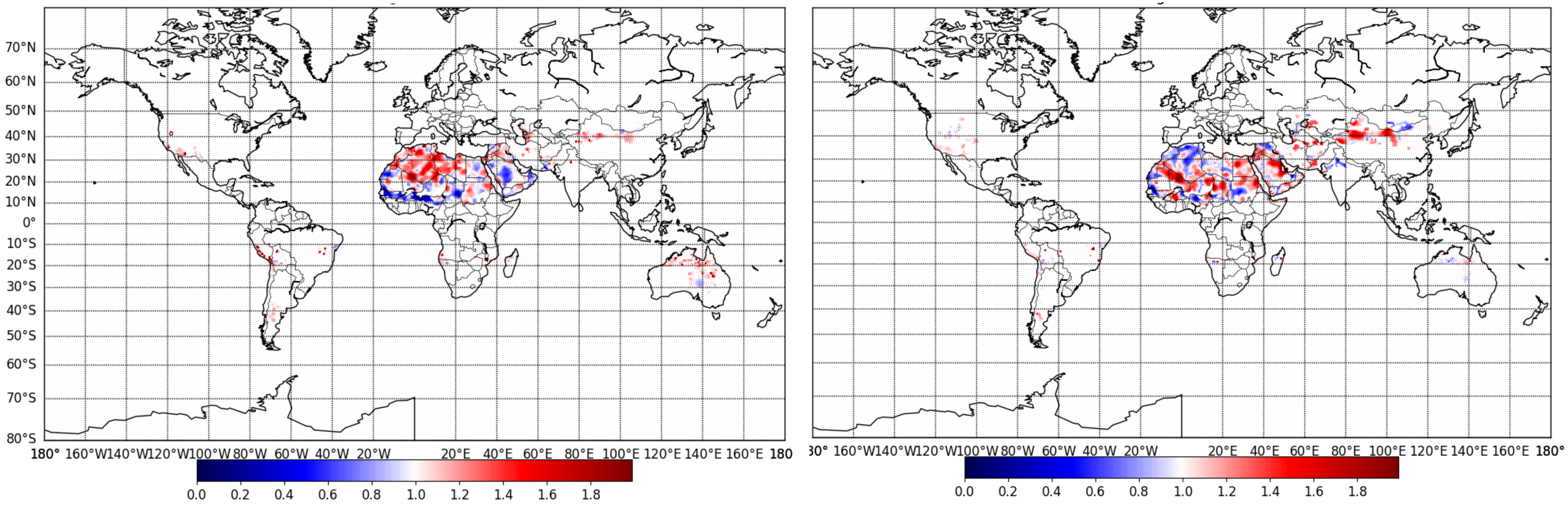
Analysis emissions (correction factors)

IFS

Analysis driven AOD, Dust AOD, PM, etc.

WHAT IS NEXT?

- Use of the same methodology for desert dust emissions
- Reference dataset : best estimate of daily IFS-COMPO dust emissions through offline inversion with an ensemble approach



Ratio of mean monthly dust emissions from best estimate over those from the control run; January (left) and May (right) 2017. Red indicate larger emissions from best estimate, blue, lower emissions.

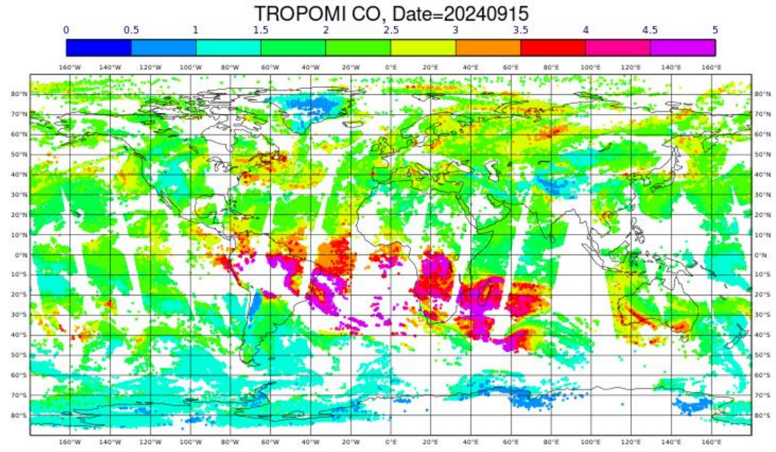
Part 2

Goal: Forecasting Atmospheric Composition with AI

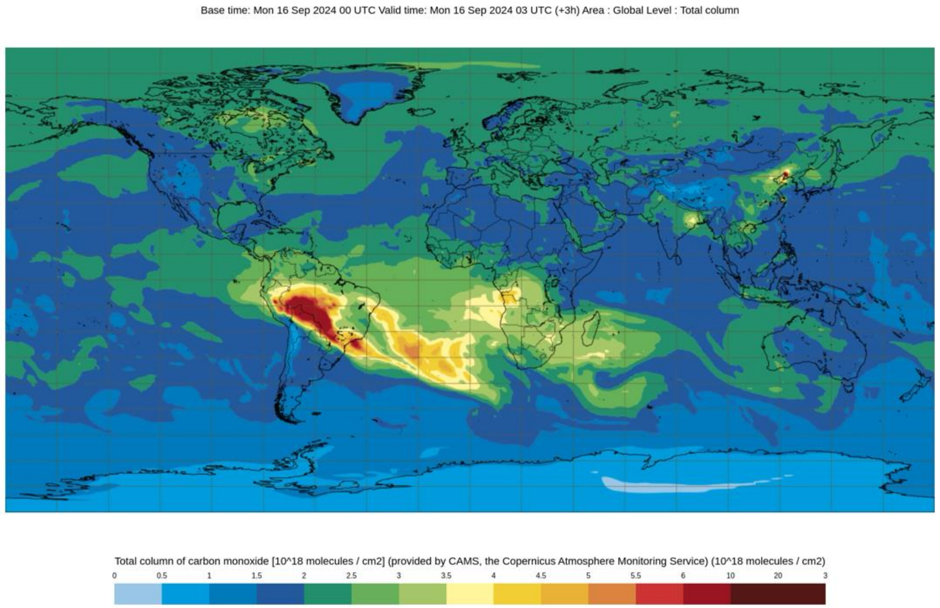
Building an AI version of IFS-Compo

40 km resolution

5-day AC and NWP forecast at 0 and 12 UTC



Carbon Monoxide forecasts

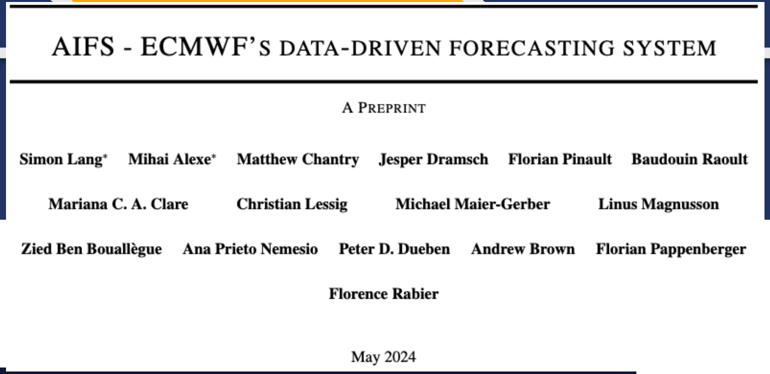


Satellite observations

AIFS-Compo

Forecast → **Analysis** → **Forecast** → **Analysis** → • • •

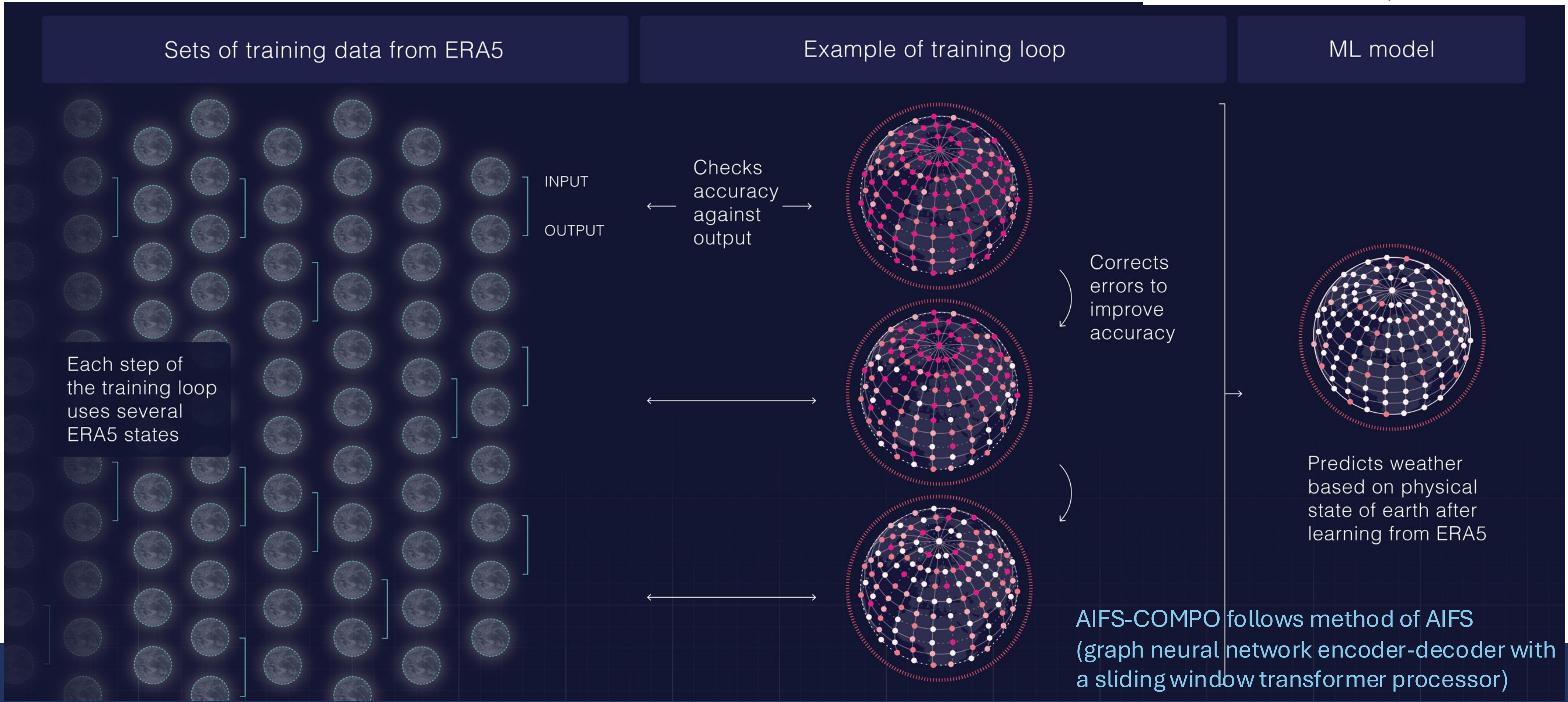
AIFS for NWP


AIFS - ECMWF's DATA-DRIVEN FORECASTING SYSTEM

A PREPRINT

Simon Lang* Mihai Alexe* Matthew Chantry Jesper Dramsch Florian Pinault Baudouin Raoult
Mariana C. A. Clare Christian Lessig Michael Maier-Gerber Linus Magnusson
Zied Ben Bouallègue Ana Prieto Nemesio Peter D. Dueben Andrew Brown Florian Pappenberger
Florence Rabier

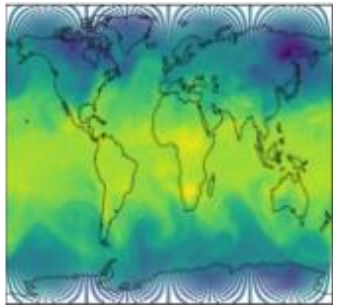
May 2024



AIFS-COMPO follows method of AIFS (graph neural network encoder-decoder with a sliding window transformer processor)

From AIFS to AIFS-Compo

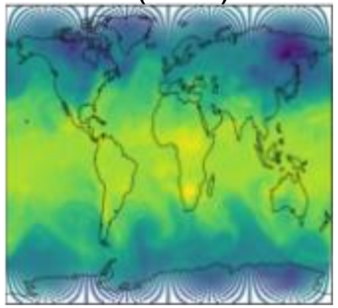
Input
Atmospheric state+AC variables:
X(t)



AIFS-Compo MODEL

encoder → *processor* → *decoder*

Output
Prediction:
X(t+3h)



Variables:
- Atmospheric composition variables
    - AOD, PMs, Reactive gases
    - Mixing ratio at pressure levels
- Upper-air variables at 13 pressure levels (t,v,u,w,z)
- Surface variables (temperature, winds, pressure, radiation)
- Static geographical features, location and time information as input forcing

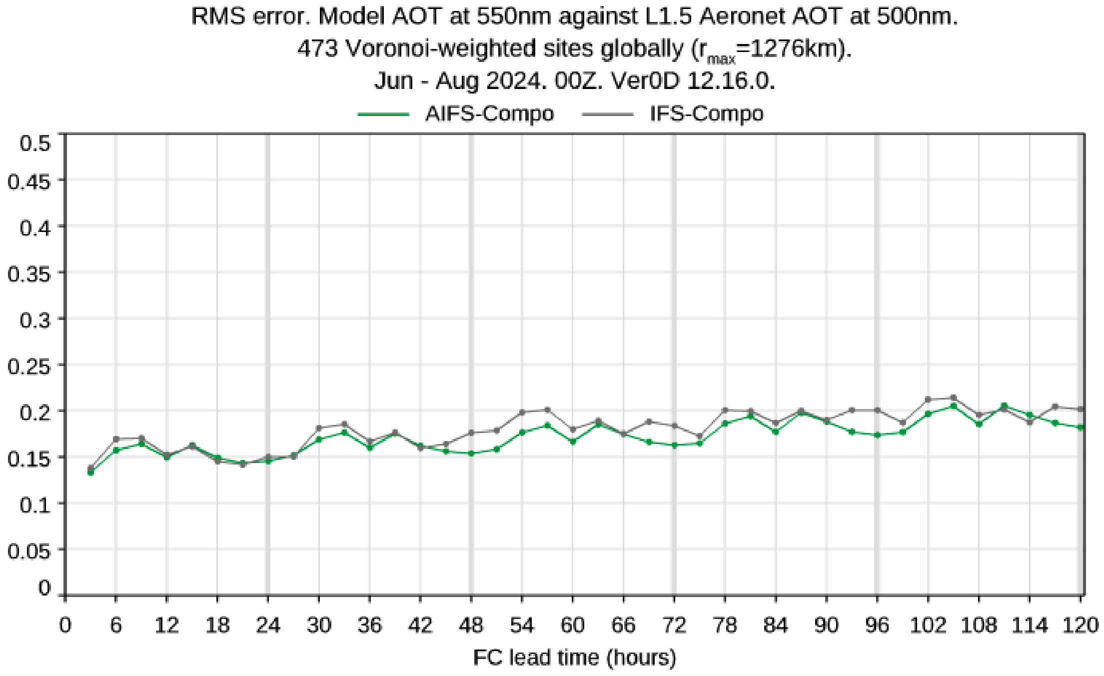
Training Scheme:
1. Train on EAC4
2. Fine-tune on operational analysis/forecast and lead times up to 36h

Implemented in anemoi

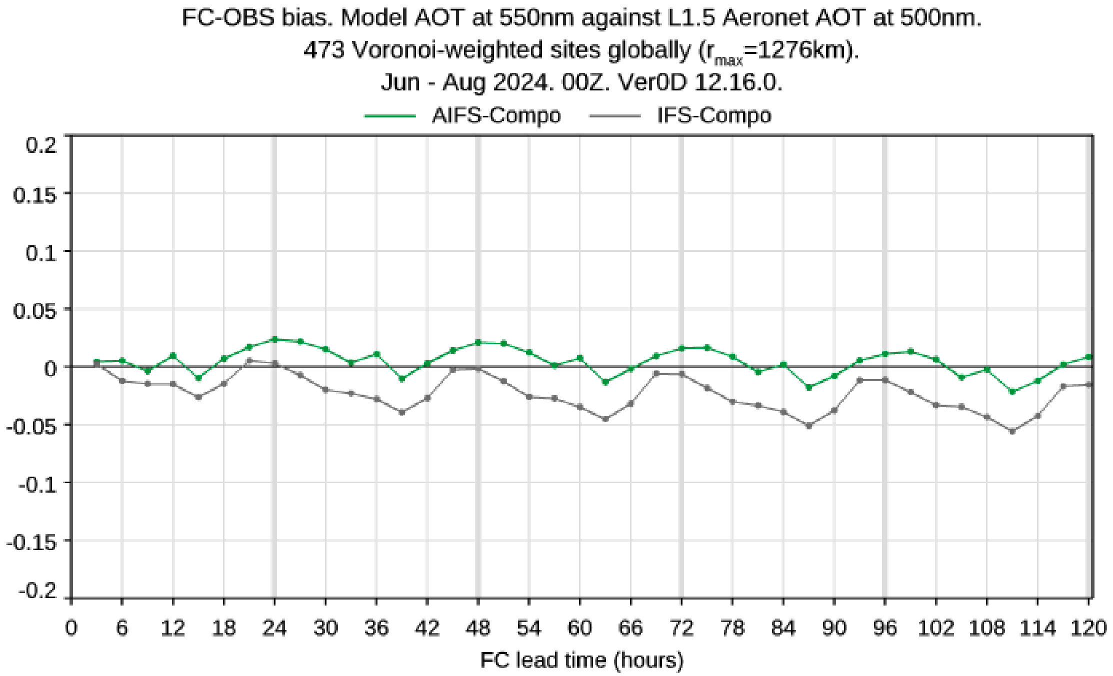
Results: AOD

AIFS-Compo 0.1

RMSE against observations



Bias against observations

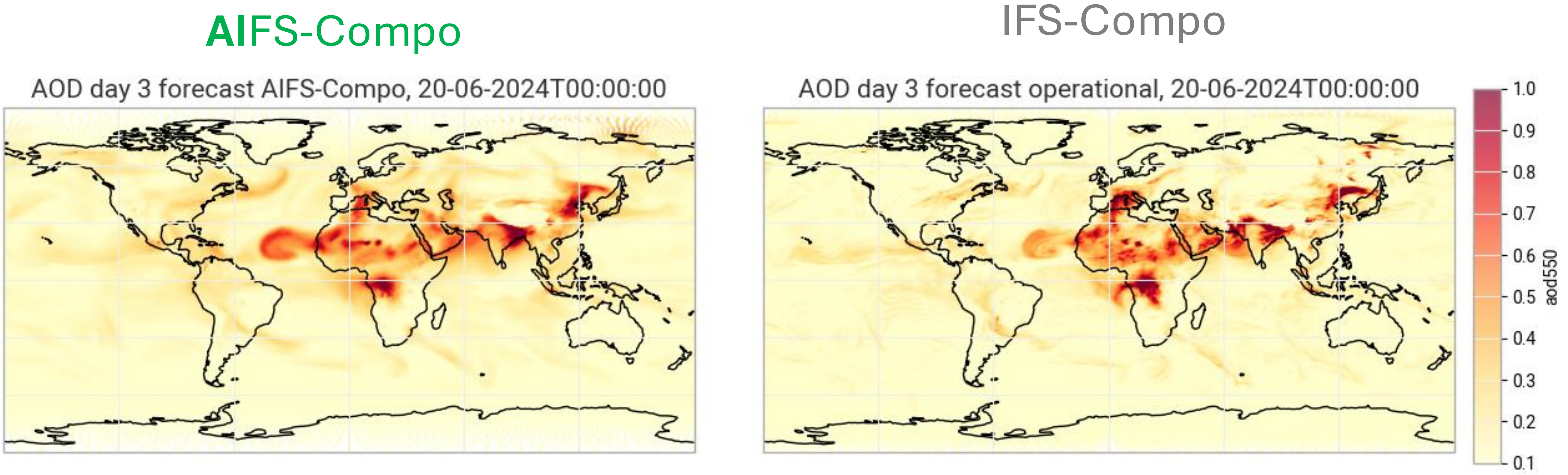


IFS-Compo

AIFS-Compo

Results: AOD

AIFS-Compo 0.1



Results summary

Good results for aerosols	Compared to observation AIFS-Compo beats IFS-Compo for AOD, PM2.5 and PM10 for almost all cases
Improvement needed for reactive gases & NWP	Reactive gases like no2, no, co, so2 and o3 are still not as good as operational forecast NWP variables are less well predicted
20 days stability	Verified forecast up to 10days Stability of AI model up to 20days using 3h timesteps
Large speedup	Time to create a 3 hourly 5-day forecast AIFS-Compo: 0.76s on 1GPU IFS-Compo: 1000s on 8000CPUs

Future work

AIFS-Compo 0.2 improvements for NWP vars + reactive gases

Emissions included as input

Additional variables: GHG, model level

Ensemble forecast

Higher resolution

Learning from observations

?

ありがとう
Arigatō