CAMAERA

Harnessing Machine Learning and Deep Learning methods to forecast whitecap fraction and sea-salt aerosol emissions in the ECMWF Integrated Forecast System (IFS-COMPO)

Nathan Capon[1], Rose-Cloé Meyer[1], Samuel Remy[1], Magdalena Anguelova[2], Jean Bidlot[3], Josh Kousal[3], Thierry Elias[1] and Antonino Bonanni[3]

[1] HYGEOS, Lille, France

[2] NRL, Monterey, U.S.A

[3] ECMWF, Reading, U.K.

PROGRAMME OF THE EUROPEAN UNION

Copernicus Europe's eyes on Earth

IMPLEMENTED BY ECMWF

Coordinated by HYGEOS

CAMAERA : A HORIZON EUROPE PROJECT

The CAMS AERosol Advancement (**CAMAERA)** project is a Horizon Europe Project to support the development of the Copernicus Atmosphere Monitoring Service

CAMAERA focuses on aerosol modeling in CAMS:

- Improvement of the **regional** and **global** CAMS (IFS-COMPO) aerosol modeling capacities
- Assimilation of new sources of **data**
- Better representation of secondary aerosols and their precursors
- Foster exchanges between the regional and global components

Atmosphere Monitoring Service

CAMAERA

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

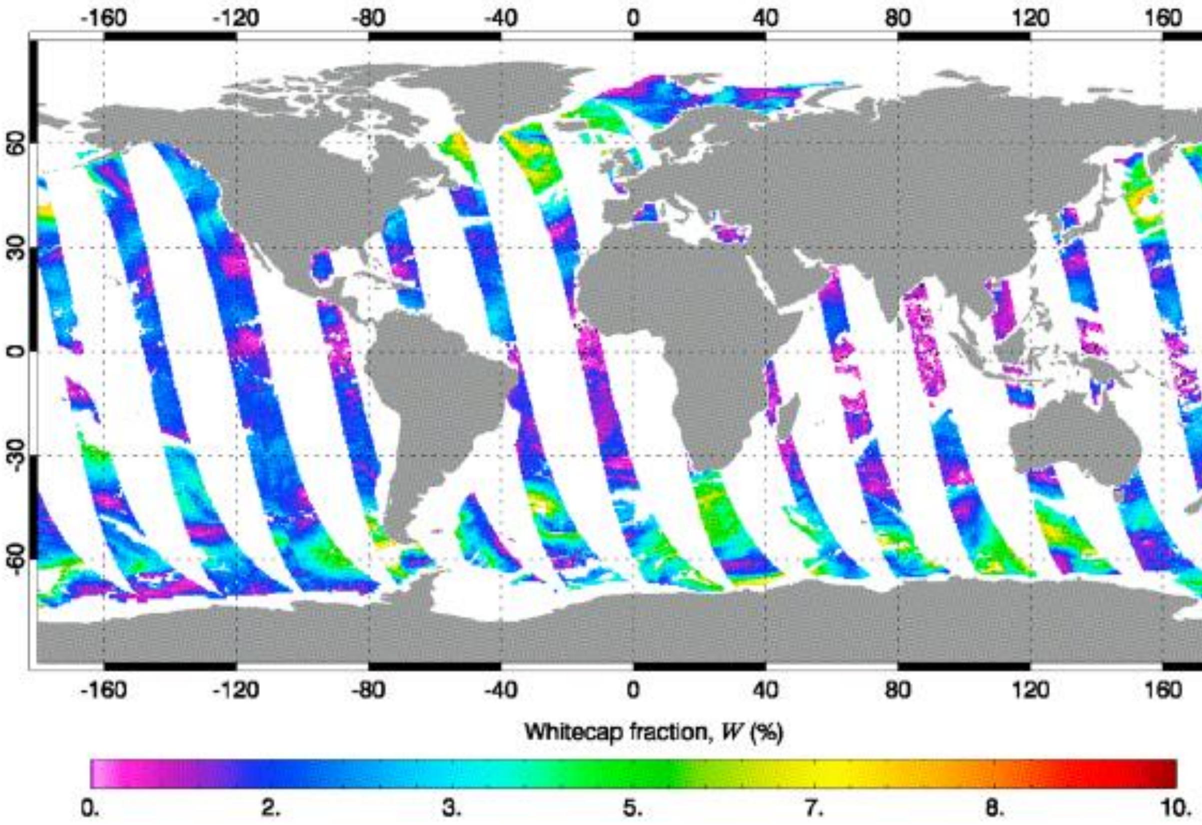
Dataset description

Ground truth : Whitecap fraction (WF) derived from remote sensing (Anguelova et al 2019)

Time range : 2 years of data with an hourly resolution

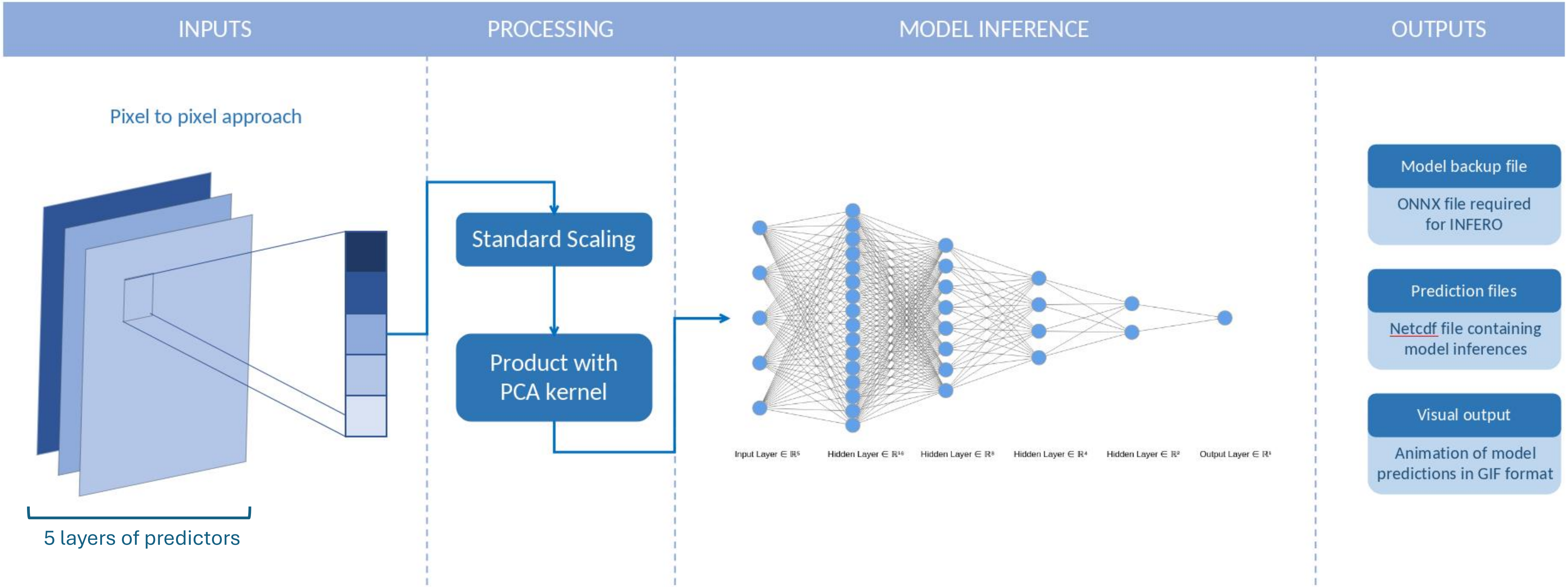
Predictors : 5 predictors collected from ERA5
- Wind Speed
- Wind Direction
- Sea Surface Temperature
- Mean Wave Period
- Significant Wave Height

Dimension : around 200 millions pixels



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

OFFLINE DNN MODEL ARCHITECTURE



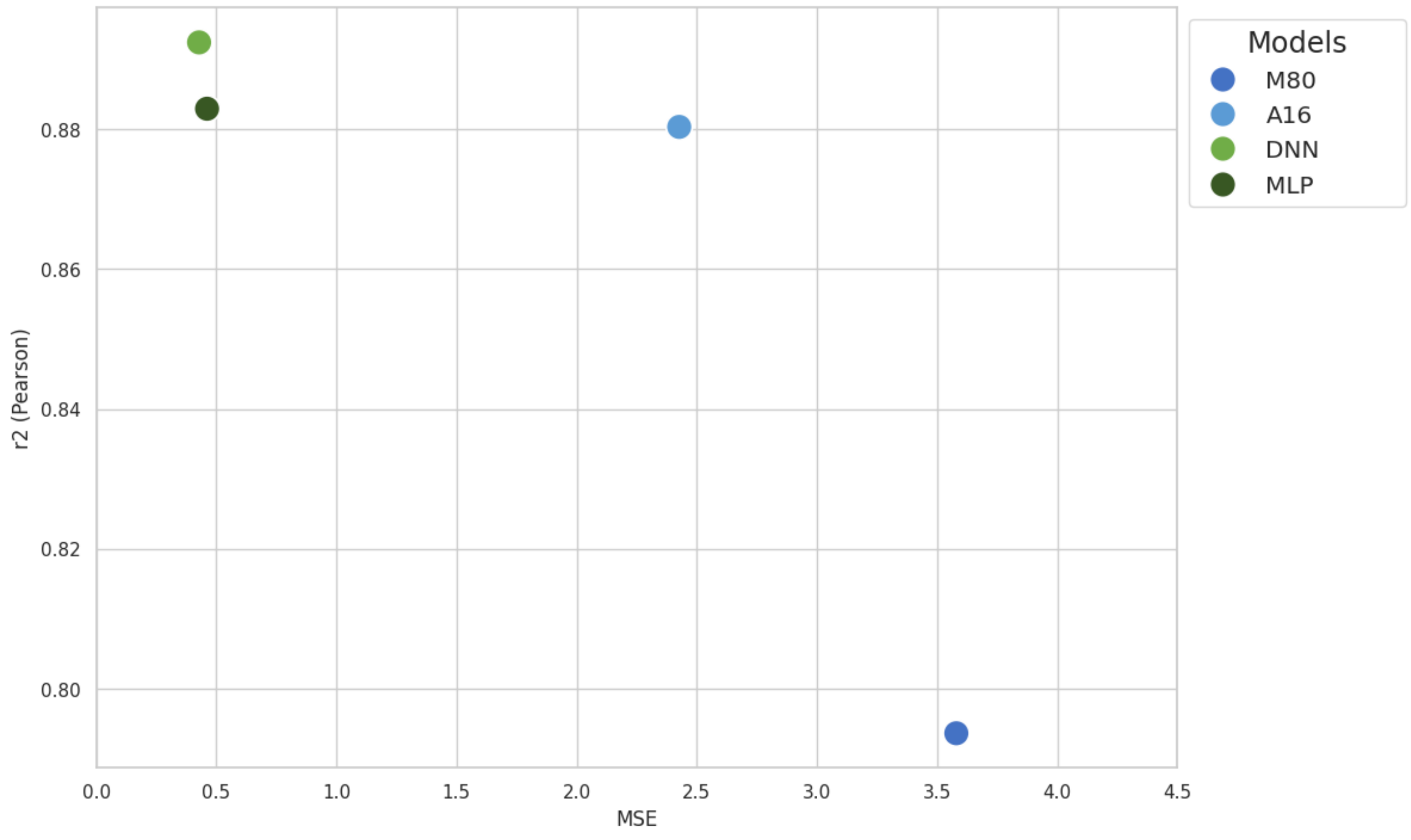
Description of the Deep Neural Network (DNN) model architecture and pre- and post-processing operations

RESULTS OF THE OFFLINE DNN MODEL AND COMPARISON WITH OTHER METHODS

Following results have been obtained by launching a **6 months** simulation

Main comments:

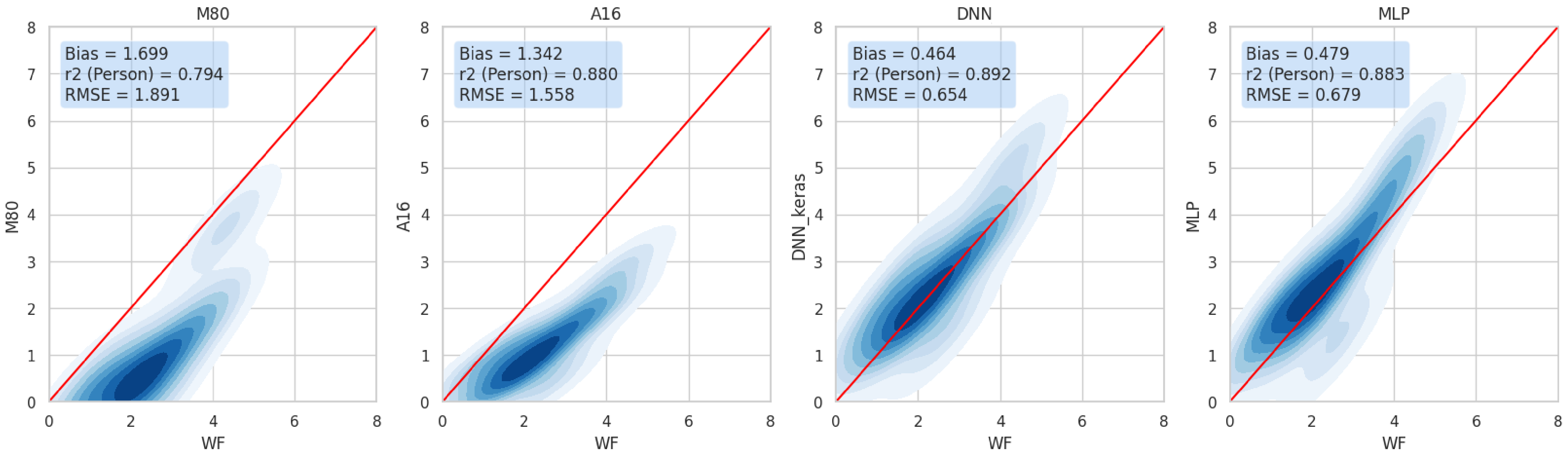
- Arithmetics models (Monahan 80, Albert 16) show a **low bias** as compared to our dataset
- DNN **outperforms** arithmetic models
- DNN manages to score better than a simple neural network architecture (MLP)
- No dependency to SST found



Pearson correlation score versus Mean Squared Error (MSE)

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

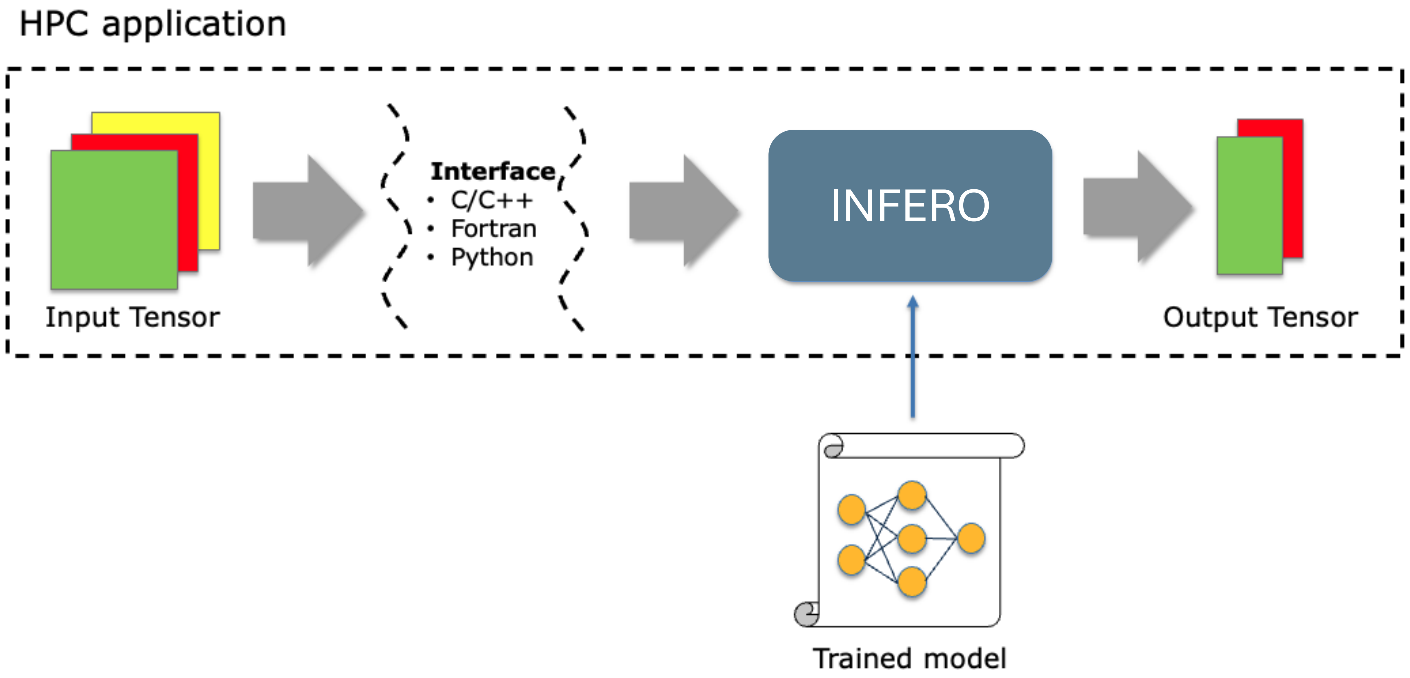
Incorporation of a **reduced version** of the DNN to make an initial estimate of its impact on the skill of simulated AOD over oceanic surfaces

- Reduced model: only 2 predictors (Wind speed / Sea Surface Temperature)

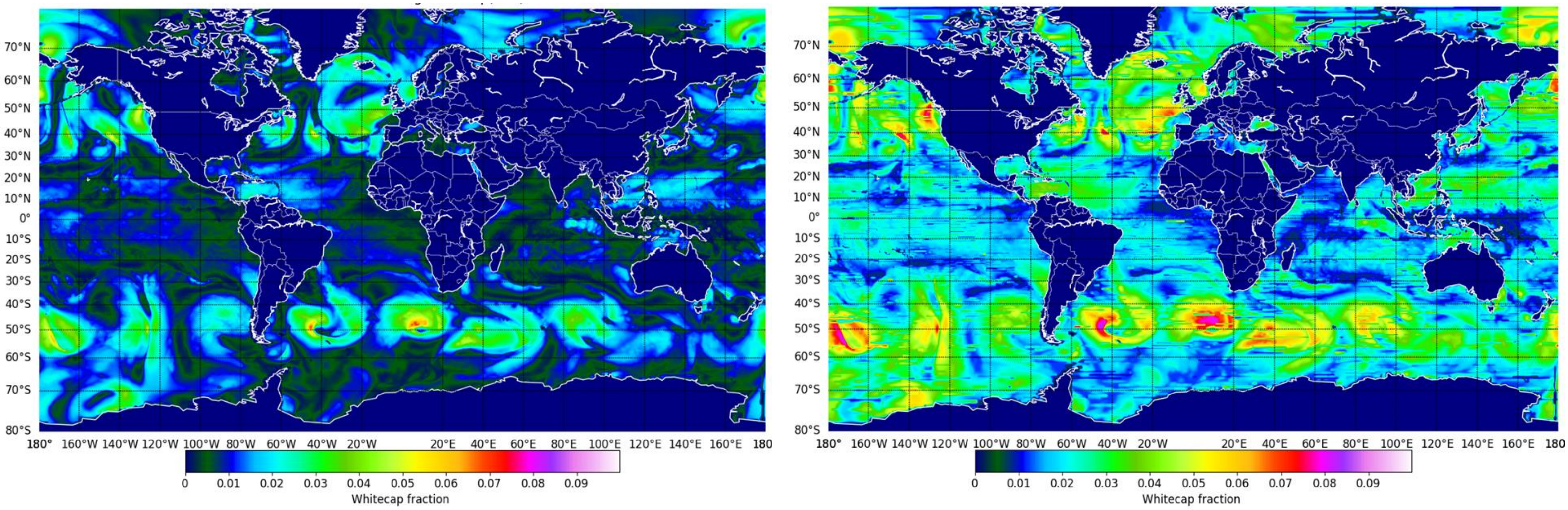
The **INFERO** library has been integrated into IFS-COMPO to interface with Deep Learning models

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

Representation of the incorporation of our model into IFS



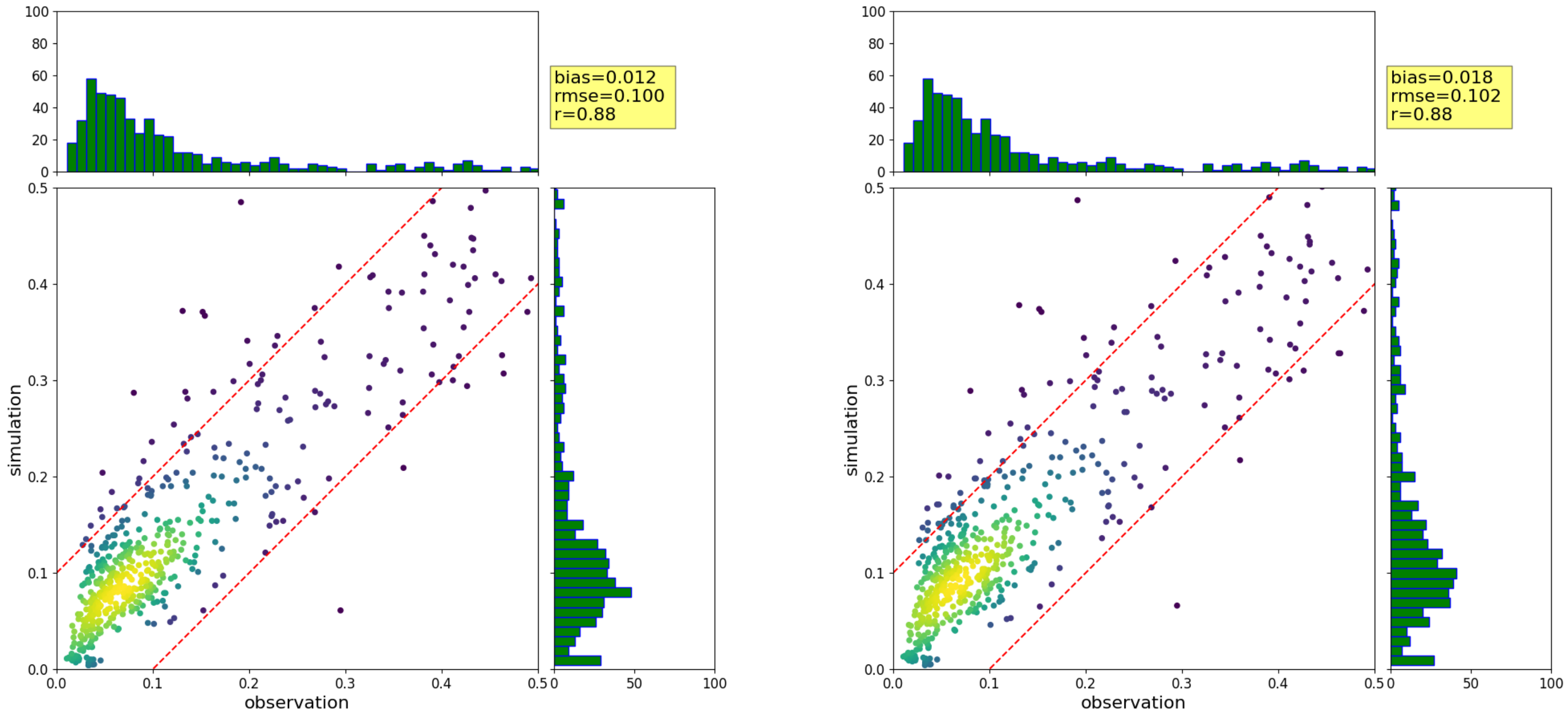
EXAMPLE OF IFS-COMPO SIMULATED WHITECAP FRACTION



Simulated whitecap fraction by IFS-COMPO on 1/1/2017 0UTC, using the operational A16 scheme (left), and with deep learning model enabled through the INFERO library (right).

IMPACT ON SKILL SCORES OF SIMULATED AOD

Skill scores of simulated AOD are very close. Larger impact noted in simulated Na/Cl at surface versus EBAS observations.

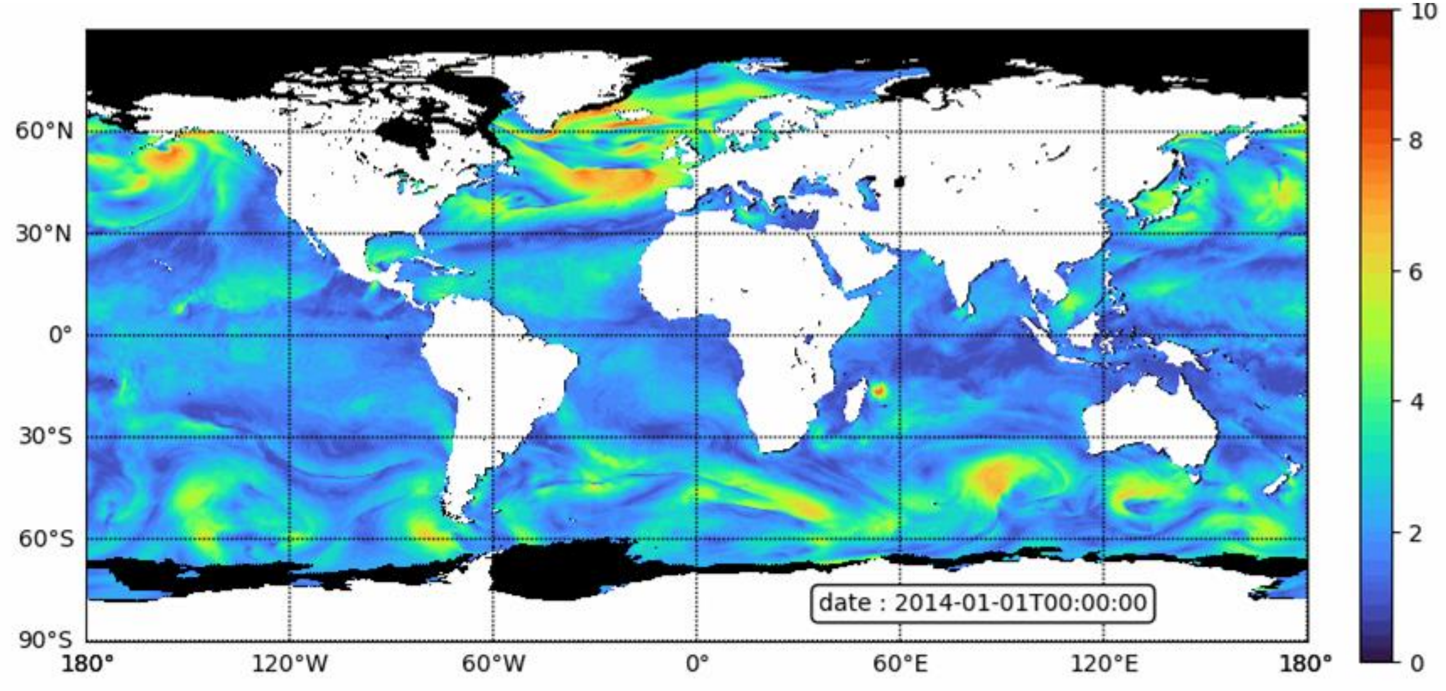


Simulated and observed AOD at 500nm from the MAN network, daily values in 2017

INCOMING DEVELOPMENTS

Coming next

- Improve offline model calibration and test new hyperparameters (loss function, input data representation, etc.)
- Implementation of models using spatial information (Unet, Transformers, etc.)
- Extend IFS-COMPO implementation to use precursors from the wave model (dissipation of turbulent energy from breaking waves, etc.)
- Use of DNN in IFS-COMPO for other purposes : desert-dust emissions



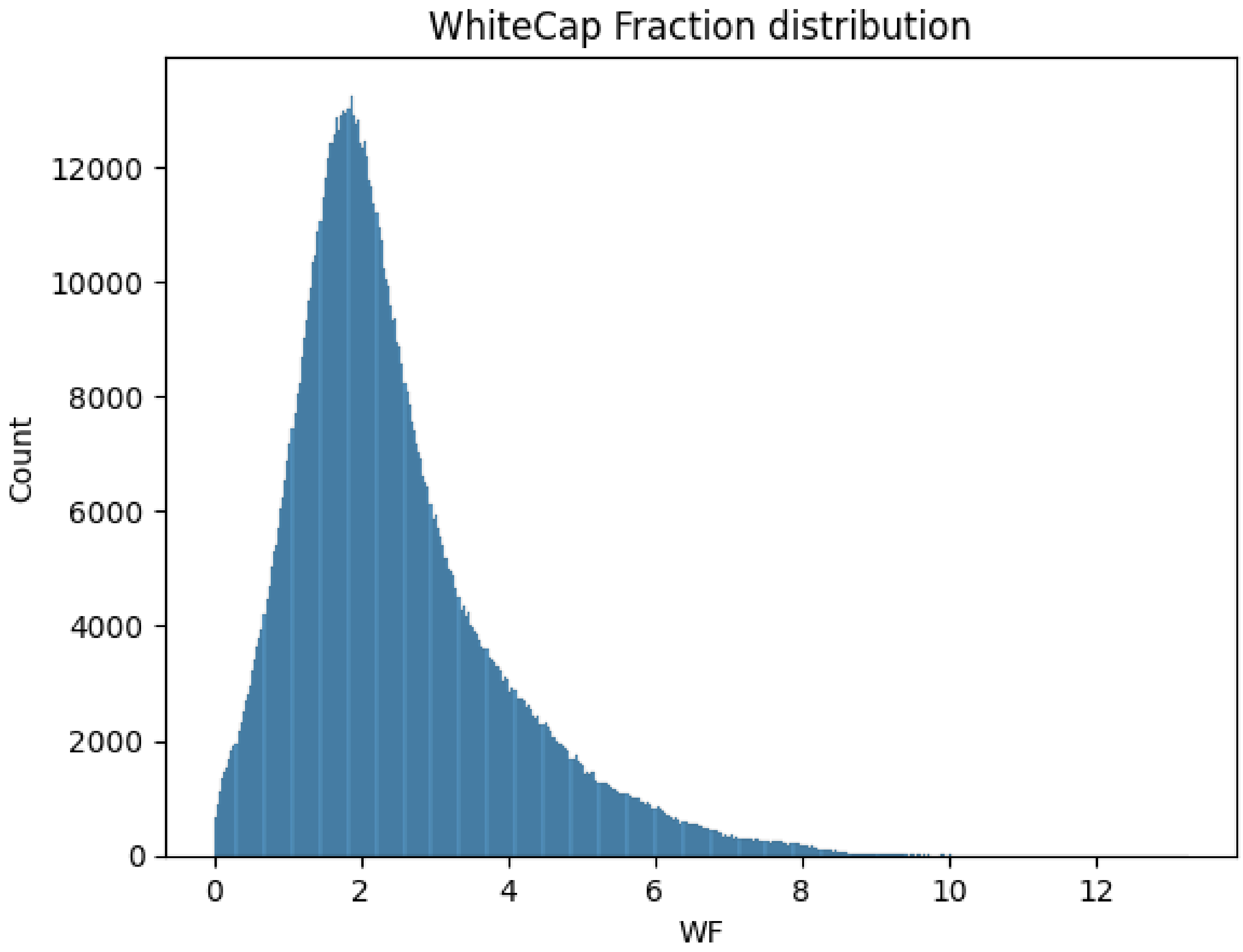
Animation of Whitecap fraction estimated from our DNN model

THANKS FOR LISTENING

CAMAERA

Coordinated by HYGEOS

WHY LOSS FUNCTION IS IMPORTANT IN OUR CASE



The **loss function** measures how well or poorly a model's predictions **match** the actual outcomes.

By adjusting the model's parameters to reduce the loss, the model learns to make **more accurate predictions**. It's crucial because it **guides** the model's **learning process**.

In our case, an adapted loss function enables us to prevent the model from ignoring minorities.