



PROGRAMME OF  
THE EUROPEAN UNION



IMPLEMENTED BY



#EUSpace

# A first attempt at predicting desert dust emissions with neural networks in IFS-COMPO

Samuel Remy, Nathan Capon, Rose-Cloé Meyer (HYGEOs)  
Jeronimo Escribano (BSC – now ECMWF)

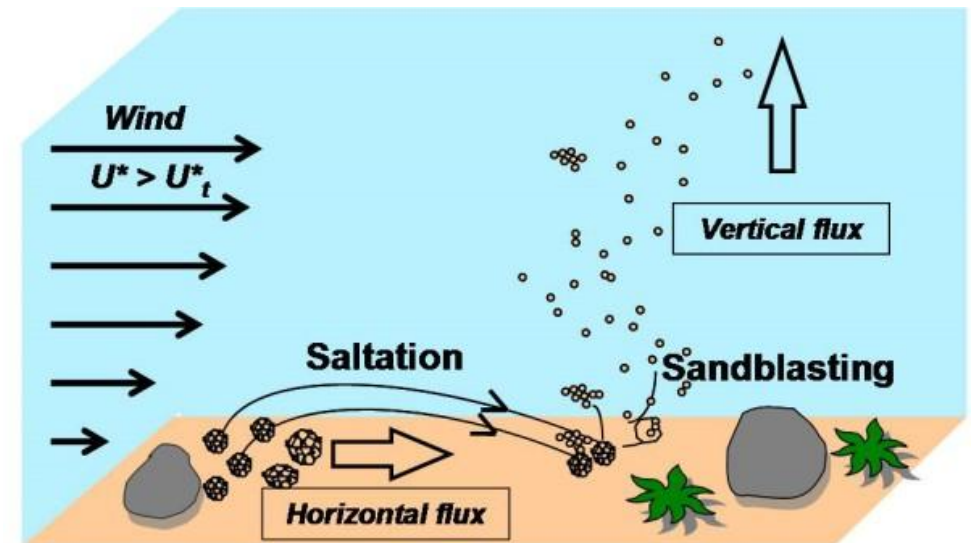


# Why use ML for desert dust emissions?

- Computational cost small compared to other processes (photolysis, wet deposition, etc.)
- But many sources of uncertainties impact desert dust emissions (inputs, resolution, representativeness, processes, etc.)
- In CAMAERA, we want to explore if there a possibility to train a statistical approach on a dataset of dust emissions that merges simulations and observations – a dust emissions analysis
- So the objective here is to improve **skill**, not **speed**.



CAMAERA



*Schematic from LISA representing the key processes for the production of desert dust aerosols.*



# Work plan

- Some desert dust emission datasets from flux inversion exist (Escribano et al 2017), but estimates depend a lot from model to model, so probably not applicable here.
- A first task is to create a reference dust emission analysis, computed with an offline ensemble Kalman smoothing applied to IFS-COMPO emissions.
- This offline dust inversion methodology can also have other uses:
  - Provides information on systematic/frequent biases of the dust emission scheme
  - Provides a comparison point to the future dust emissions from the IFS-COMPO emission inversion framework (work of Michael and Mel)
- Then, use this reference dust emission data to train a ML/NN algorithm offline.
- Finally implement in IFS-COMPO the capacity to call the inference model computed offline in order to estimate dust emissions



CAMERA



PROGRAMME OF  
THE EUROPEAN UNION



IMPLEMENTED BY



#EUSpace

# Offline inversion of desert dust emissions

Jeronimo Escribano, Samuel Remy



# Offline dust emission inversion how?

- **Data assimilation** of remote sensing dust AOD or AOD at 550nm to estimate dust emissions
- Modified workflow of the **Local Ensemble Transform Kalman Filter (LETKF)** used at BSC
- Provides scaling factors of the prior emissions (3 days temporal resolution, grid point spatial resolution)

**Assimilated DOD**  
VIIRS SNPP Deep Blue

|  |  |
|--|--|
| <b>IFS-COMPO Ensemble</b>                              | <b>LETKF (smoother)</b>                    |
| <b>20 members with perturbed emissions at the time</b> | <b>Control: dust emission</b>              |
| <b>100 active members</b>                              | - 3 days temporal resolution               |
|  | - Model horizontal resolution              |
|  | <b>Observations: Dust AOD or AOD 550nm</b> |

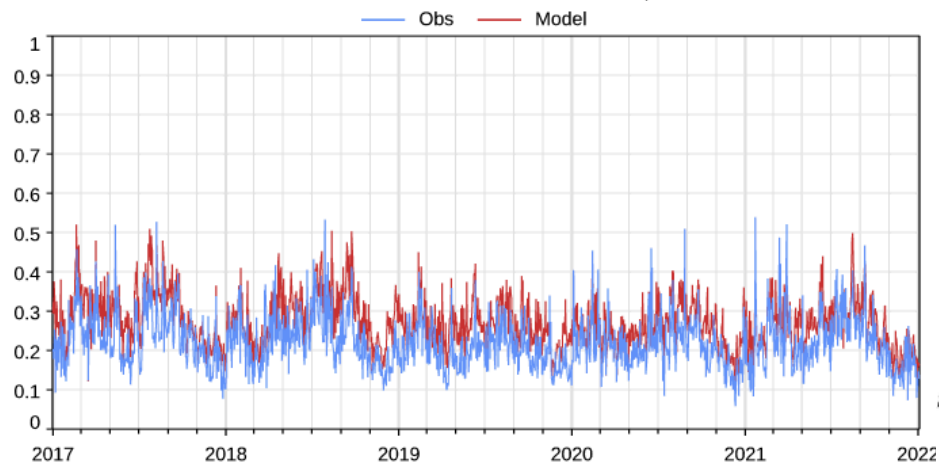
(BSC – J Escribano)



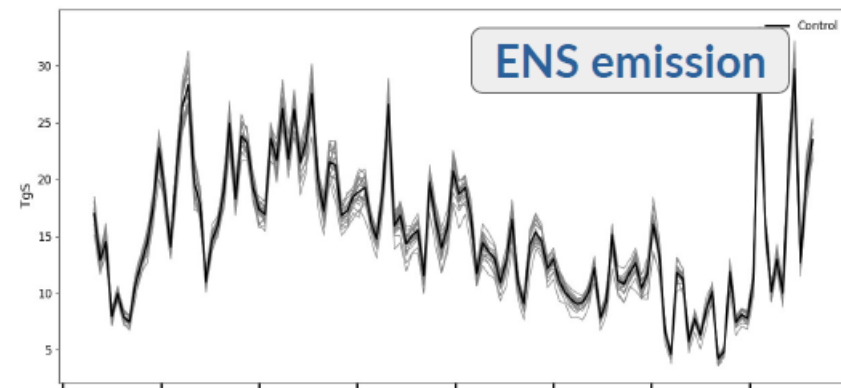
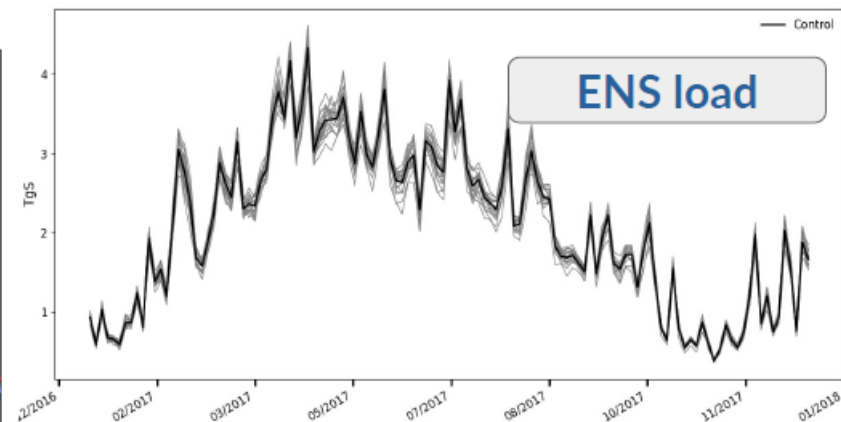
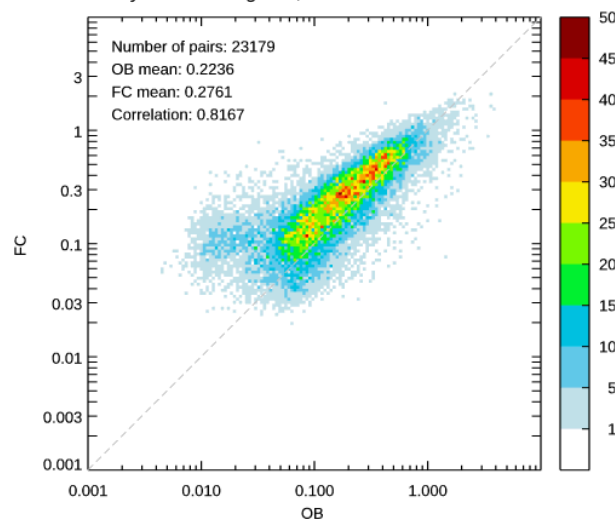
# Control run and ensemble

- 1 control and 20 perturbed members
- 15 days forecasts, cycling every 3 days, no data assimilation
- Meteorology nudged to ERA-5 during the forecast
- Perturbations of dust emissions applied in the first three days
- T511 (~50km grid cell) L137 resolution
- Aerosol only IFS-COMPO, pre cycle 50R1, no chemistry configuration
- Processed from 31/12/2016 to 1/1/2022

Mean. Model (i97q) against L2.0 Aeronet AOT at 500nm.  
23 sites in Desert AERONET. 31 Dec 2016 - 1 Jan 2022. 00Z, T+3 to 72. Ver0D 12.6.17.



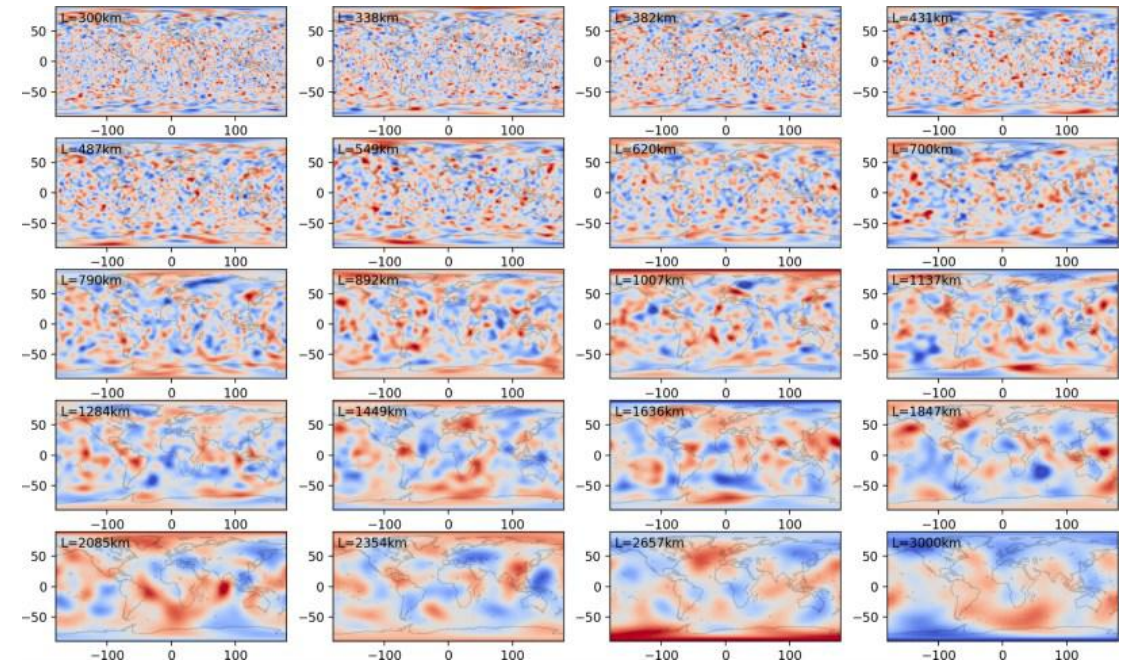
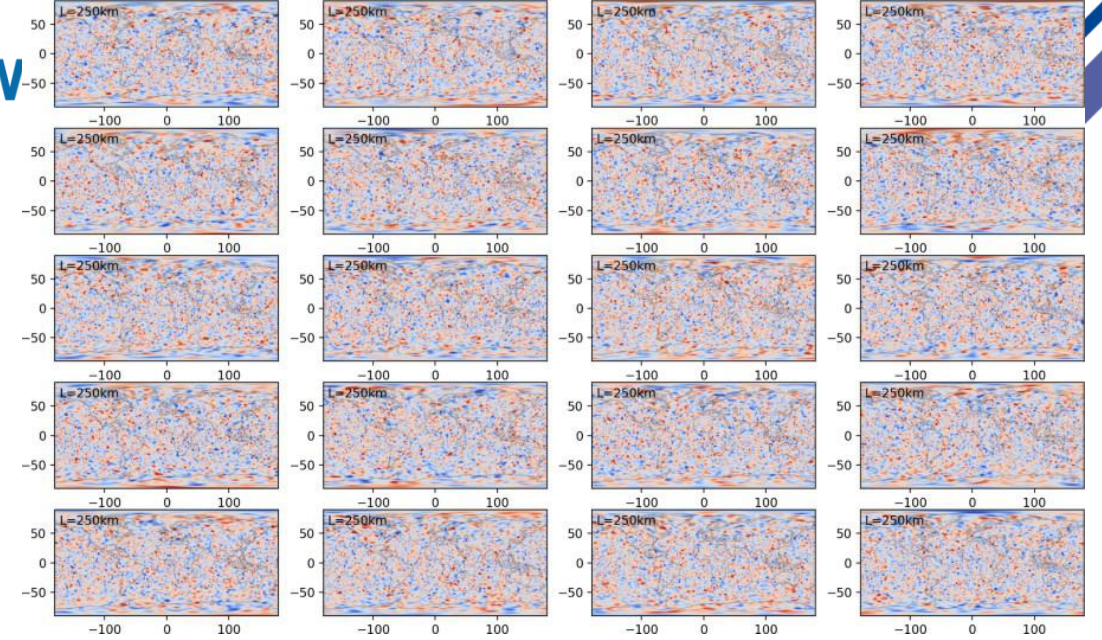
Model (i97q) vs L2.0 Aeronet normal @ 500nm  
31 Dec 2016 - 1 Jan 2022. 23 sites in Desert AERONET.  
Daily means using 00Z, T+3 to 72. Ver0D 12.6.17.



(BSC – J Escribano)

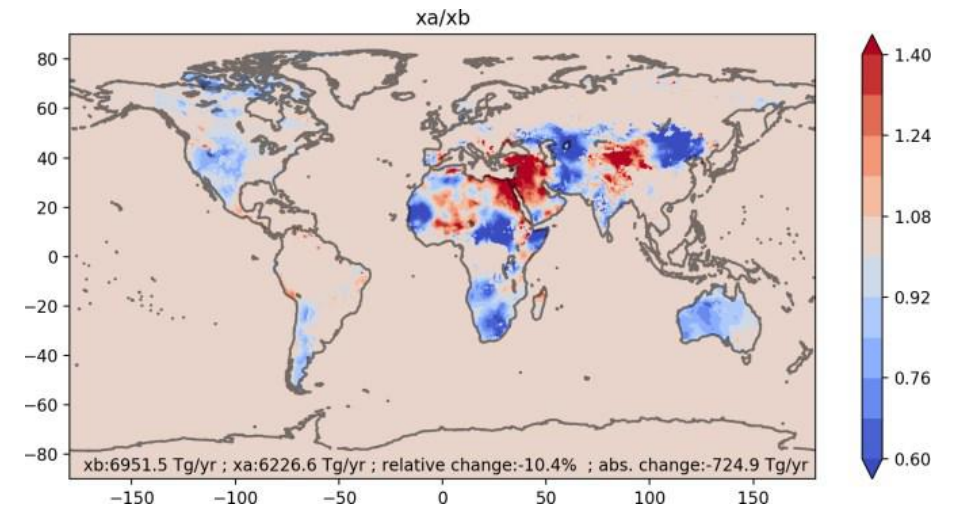
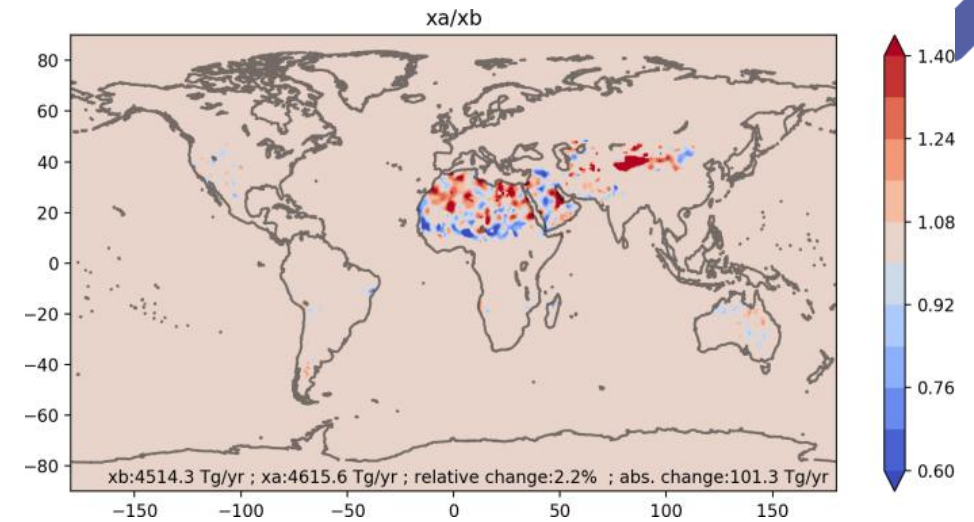
# Ensemble versions

- Version 1 with currently operational dust emission scheme and fixed correlation length of ensemble perturbations (250km)
- Version 2 with new dust emission scheme from CAMAERA WP5 (adapted from SILAM dust emission scheme), and with varying correlation length (300-3000km)
- Results of V2 shown



# Ensemble versions

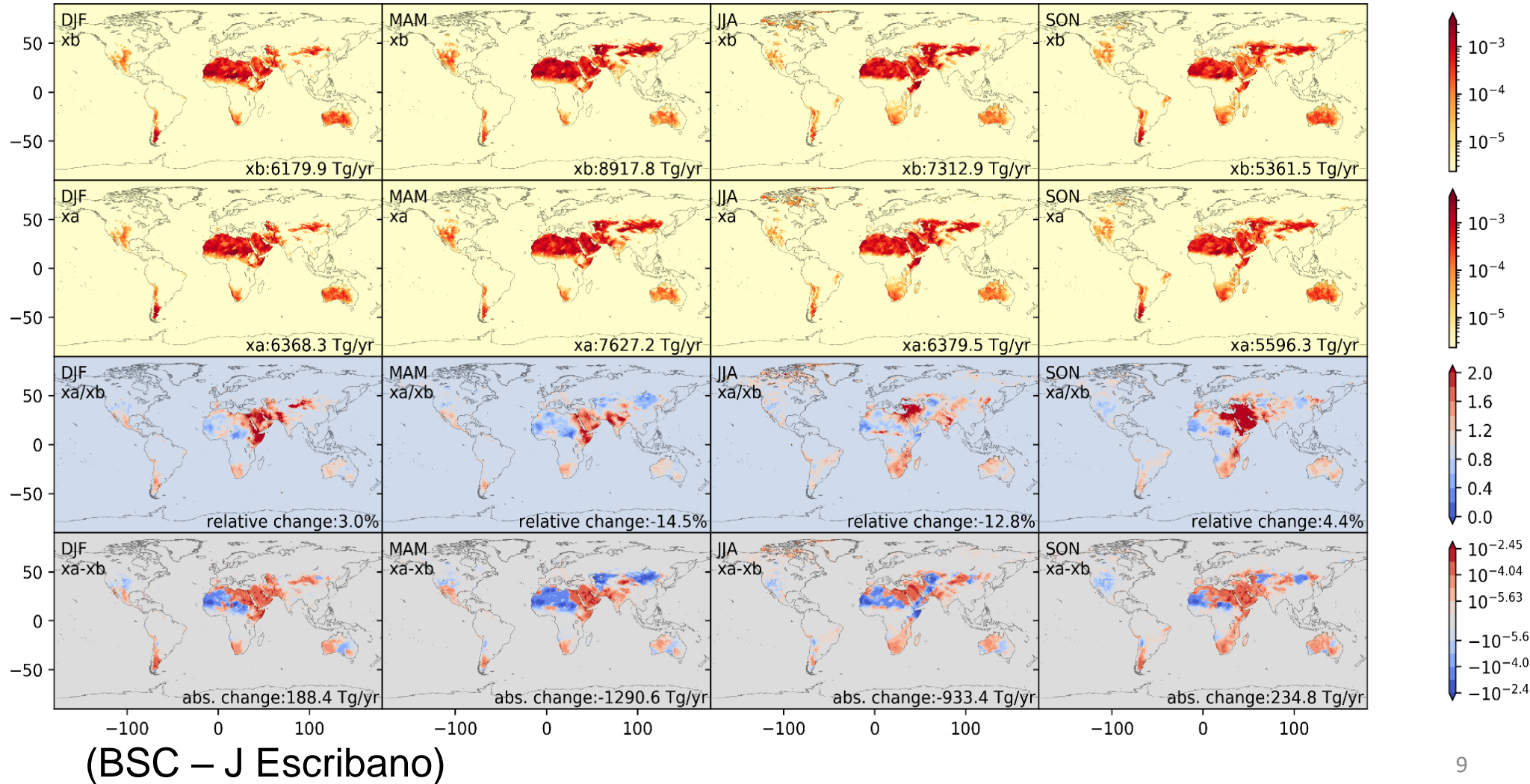
- Version 1 with currently operational dust emission scheme and fixed correlation length of ensemble perturbations (250km)
- Version 2 with new dust emission scheme from CAMAERA WP5 (adapted from SILAM dust emission scheme), and with varying correlation length (300-3000km)
- Results of V2 shown



(BSC – J Escribano)

# Results: emission changes

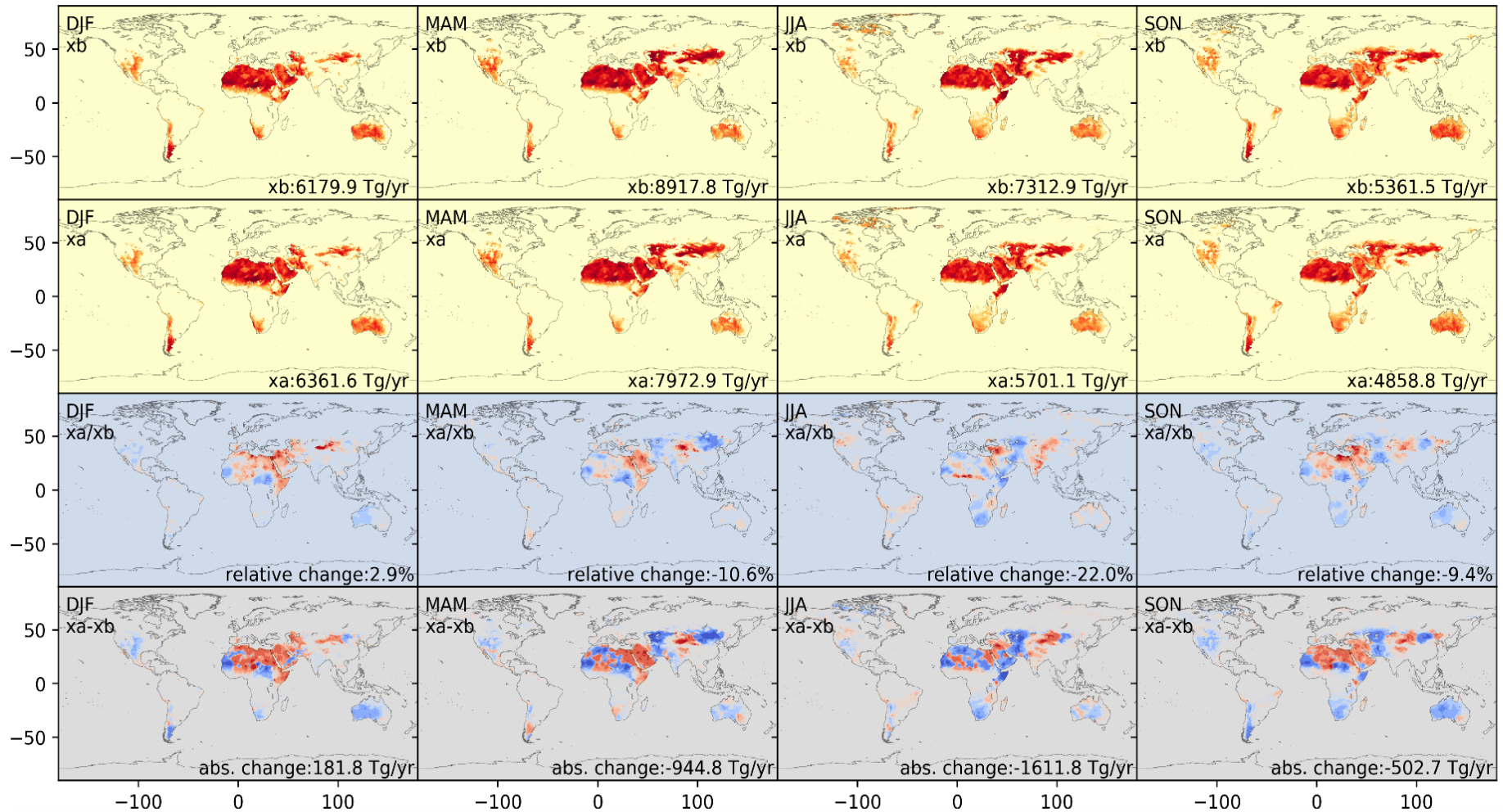
2017 only  
 Assimilation of DODx1.5





# Results: emission changes

2017 only  
Assimilation of AOD

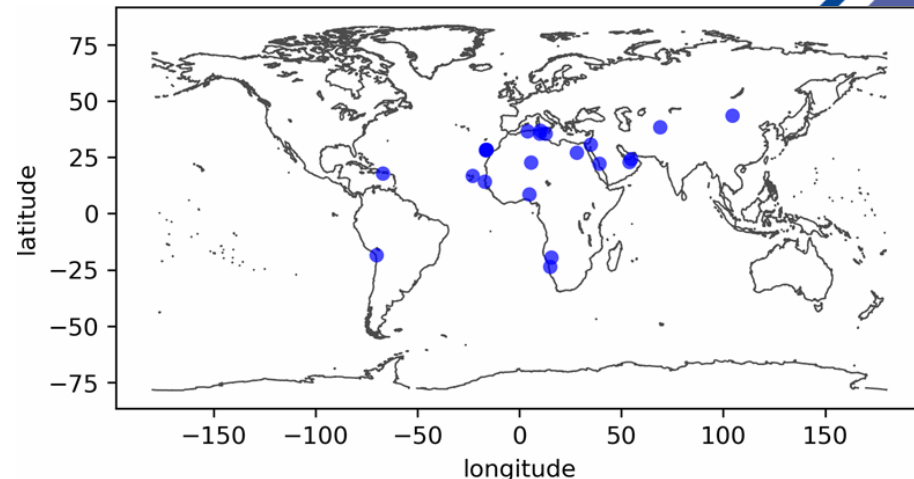


(BSC – J Escribano)



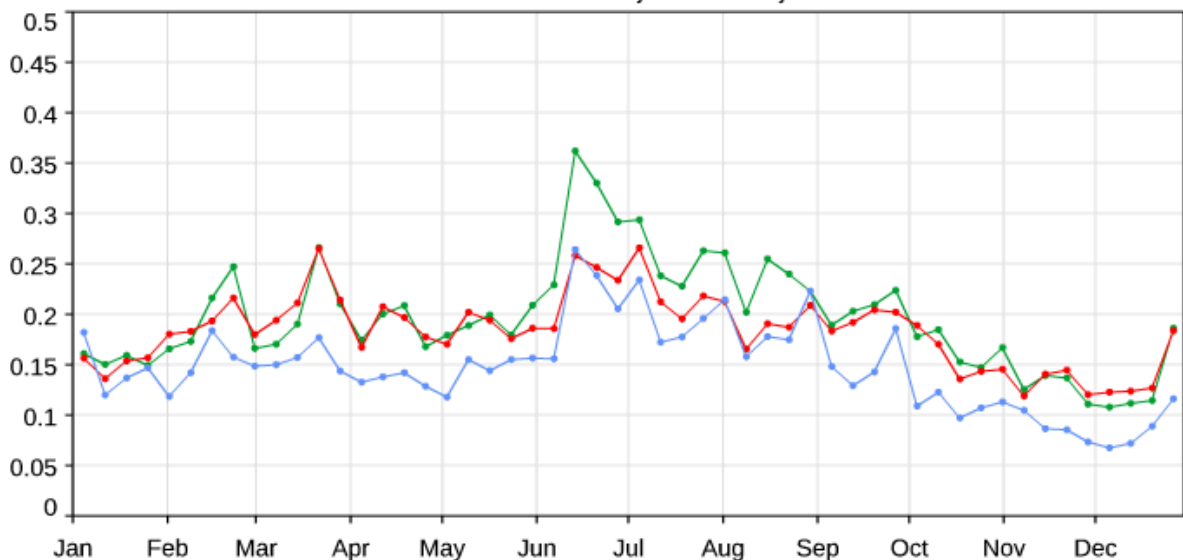
# Use of scaling factors in IFS-COMPO

- Use of daily desert dust emissions scaling factors in IFS-COMPO dust emissions
- Test in forecast only (no DA) simulations over several years
- Often lower emissions in summertime
- RMSE vs AERONET improved in summertime
- No improvement in correlation



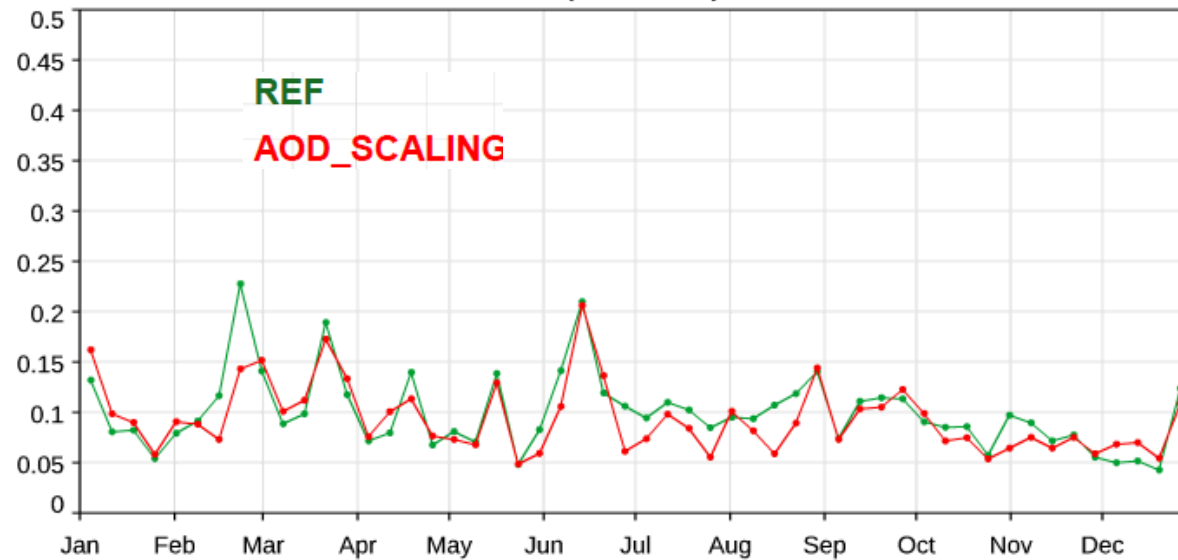
Mean. Model against L2.0 Aeronet AOT at 1020nm.  
21 sites in desert AERONET. 1 Jan - 29 Dec 2020. 00Z, T+3 to 24. Ver0D 12.19.9.

— Obs — j3fk — j3fl



RMS error. Model against L2.0 Aeronet AOT at 1020nm.  
21 sites in desert AERONET. 1 Jan - 29 Dec 2020. 00Z, T+3 to 24. Ver0D 12.19.9.

— j3fk — j3fl





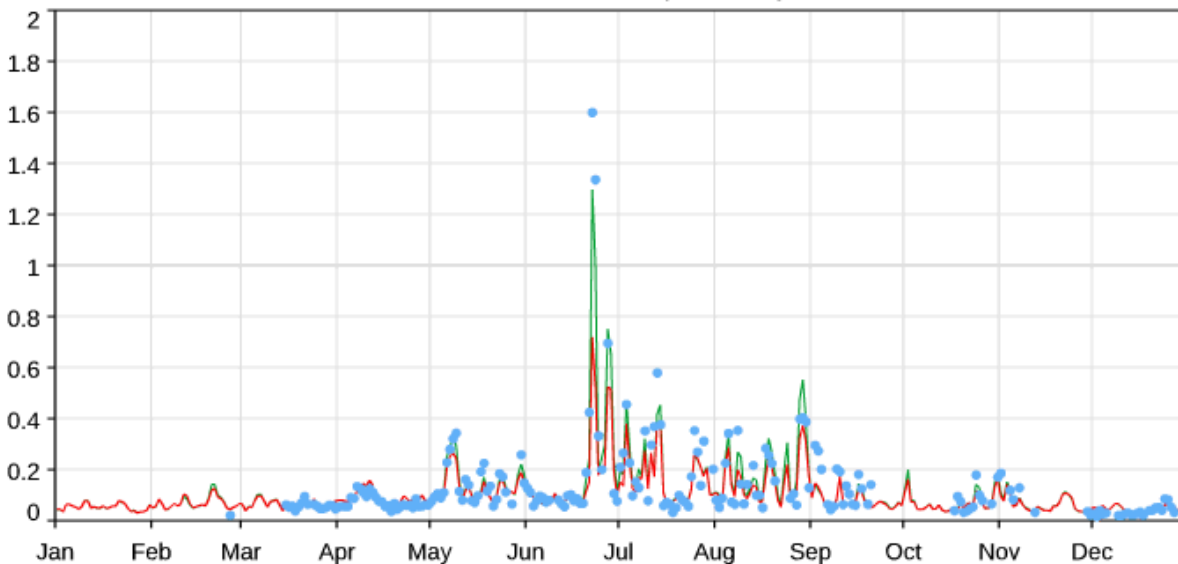
# Use of scaling factors in IFS-CO

- Use of daily desert dust emissions scaling factors COMPO dust emissions
- Often lower emissions in summertime
- RMSE vs AERONET improved in summertime
- No improvement in correlation

Comparison of j3fk & j3fl and L2.0 Aeronet AOT at 1020nm over La\_Parguera (17.97°N, 67.05°W).

1 Jan - 29 Dec 2020. Daily means using 00Z, T+3 to 24. Ver0D 12.19.9.

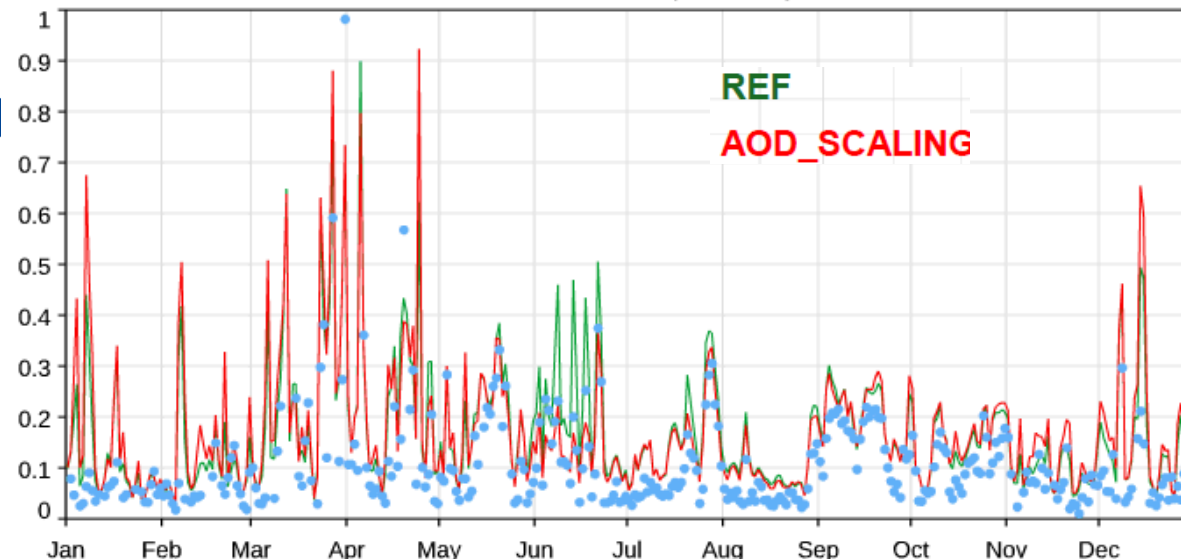
● L2.0 Aeronet — j3fk — j3fl



Comparison of j3fk & j3fl and L2.0 Aeronet AOT at 1020nm over SEDE\_BOKER (30.86°N, 34.78°E).

1 Jan - 29 Dec 2020. Daily means using 00Z, T+3 to 24. Ver0D 12.19.9.

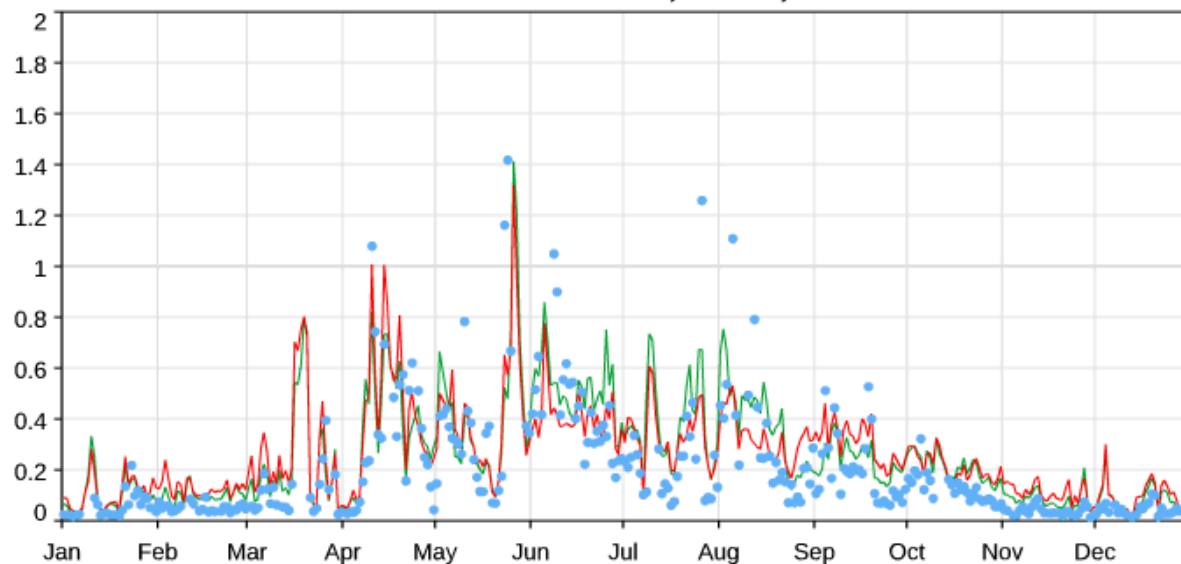
● L2.0 Aeronet — j3fk — j3fl



Comparison of j3fk & j3fl and L2.0 Aeronet AOT at 1020nm over Tamanrasset\_INM (22.79°N, 5.53°E).

1 Jan - 29 Dec 2020. Daily means using 00Z, T+3 to 24. Ver0D 12.19.9.

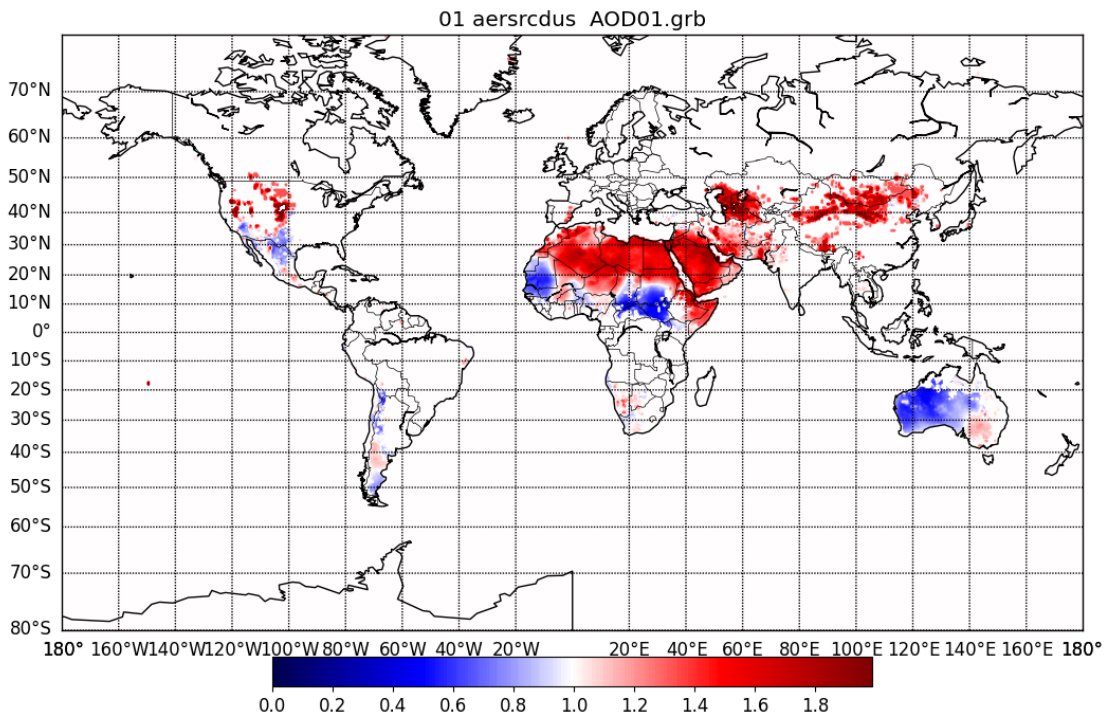
● L2.0 Aeronet — j3fk — j3fl



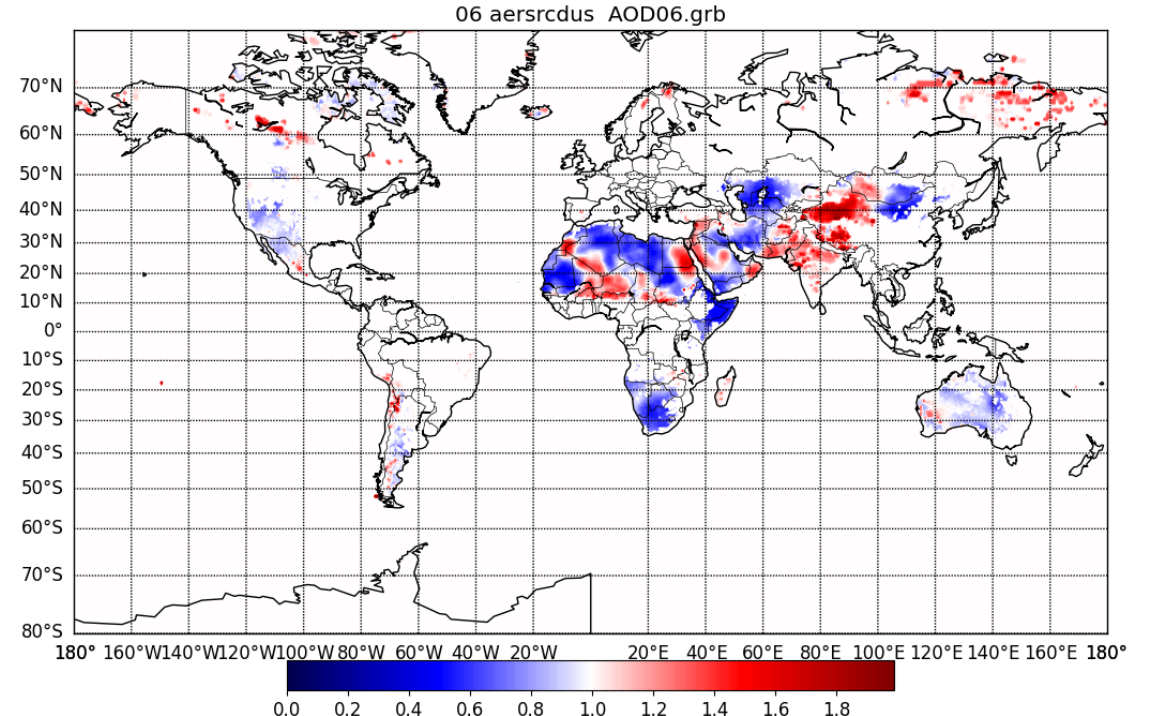


# Computation of mean monthly emissions weighted scaling factors

- Could be used to debias IFS-COMPO dust emissions
- Currently being tested



January



June



PROGRAMME OF  
THE EUROPEAN UNION



IMPLEMENTED BY



#EUSpace

# Training of a neural network to estimate desert dust emissions

Nathan Capon



# Inference model versions

- V0 (technical test) with 5 predictors – March 2026
- V1 with 15 predictors – End of April 2026
- V2 with various updates and use of log (dust emissions) as reference – June 2026



## Input data - predictors

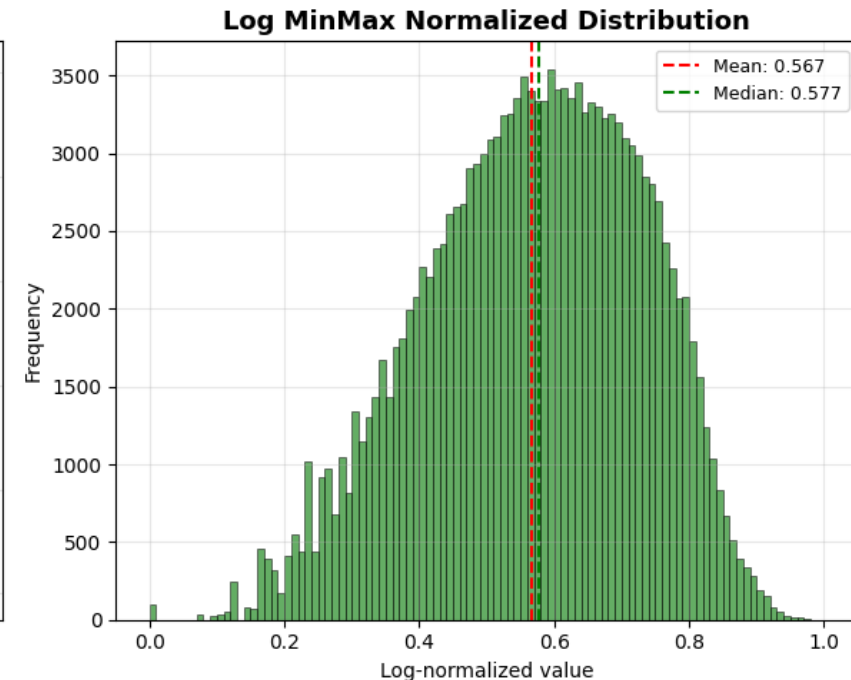
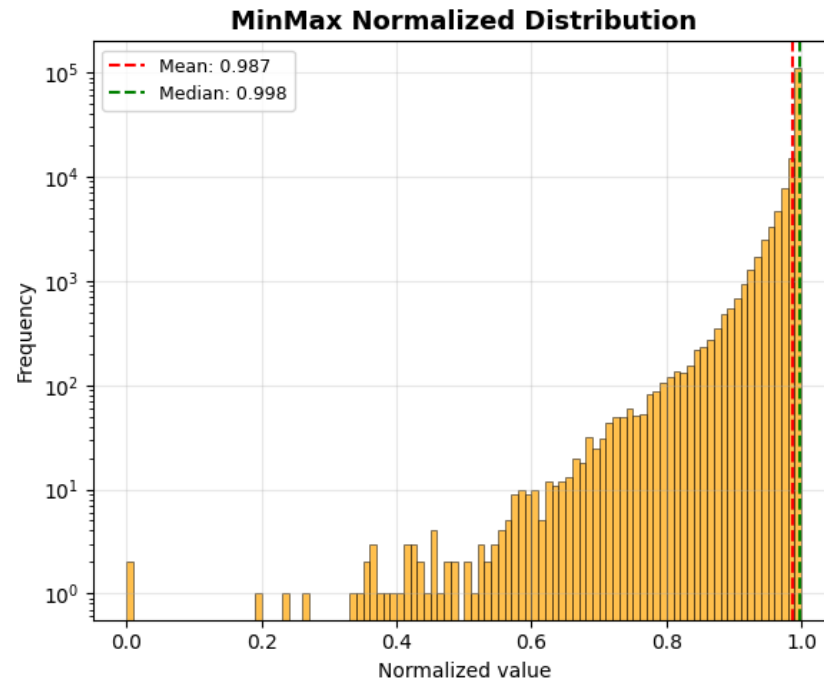
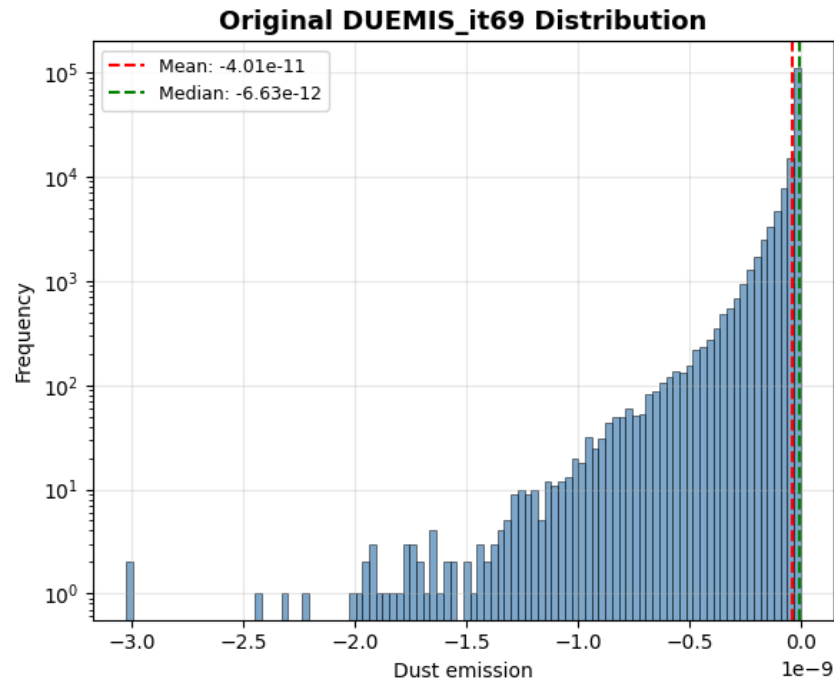
- 15 predictors from cycle 50R1 IFS-COMPO simulations
- The inputs include meteorological variables, surface properties (fixed), climatological input

| Wind               | Static Soil         | Seasonal Soil         | State Variables |
|--------------------|---------------------|-----------------------|-----------------|
| Wind speed         | Clay fraction       | Roughness length      | Soil moisture   |
| Wind gusts         | Sand fraction       | Surface albedo        | Snow depth      |
| Friction velocity  | Bare soil fraction  | LAI (Leaf Area Index) |                 |
| Threshold velocity | Orography std. dev. |                       |                 |



# Input data - reference

- Data from the dust emissions analysis – AOD control variable
  - Training: 2018–2020 (36 monthly files) → 90% train / 10% validation
  - Test: 2017 (12 monthly files)
  - Grid:  $361 \times 720 = 260\text{K}$  spatial locations per timestep
- Dataset heavily tilted towards small values
- In a v2, in order to reduce the relative weight of small values, use of  $\log(\text{normalized emissions})$  as a reference





# Model selection

- Three constraints guided the choice of a model

| Constraint                | Implication                        | Excluded Models                      |
|---------------------------|------------------------------------|--------------------------------------|
| IFS-COMPO parallelization | Only local pixel values accessible | CNNs, UNets, Transformers            |
| Continuous output         | Smooth predictions required        | Decision trees, Random Forests, GBDT |
| Non-linear relationships  | Complex feature interactions       | Ridge, Lasso, Linear Regression      |

- The most suitable model is a feed-forward Neural Network (DNN)



# Hyperparameter optimisation

Finding the best hyperparameters is like searching for the highest point in a mountain range without a map (in foggy conditions)





# Hyperparameter optimisation

Hyperparameters:

| Category      | Parameters                              |
|---------------|---|
| Architecture  | Layers, Hidden dim, Activation, Dropout |
| Training      | Batch size, LR, Weight decay, Grad clip |
| Optimizer     | Type, Scheduler                         |
| Normalization | MinMax, Standard                        |
| Loss          | MSE, MAE, R <sup>2</sup>                |

Possible approaches:

| Approach                     | How it works           | Problem                                   |
|------------------------------|------------------------|---|
| <b>Random Search</b>         | Sample randomly        | Wastes trials on already-explored regions |
| <b>Grid Search</b>           | Test every combination | Too slow (combinatorial explosion)        |
| <b>Bayesian Optimisation</b> | Learn from past trials | Smart & efficient exploration             |

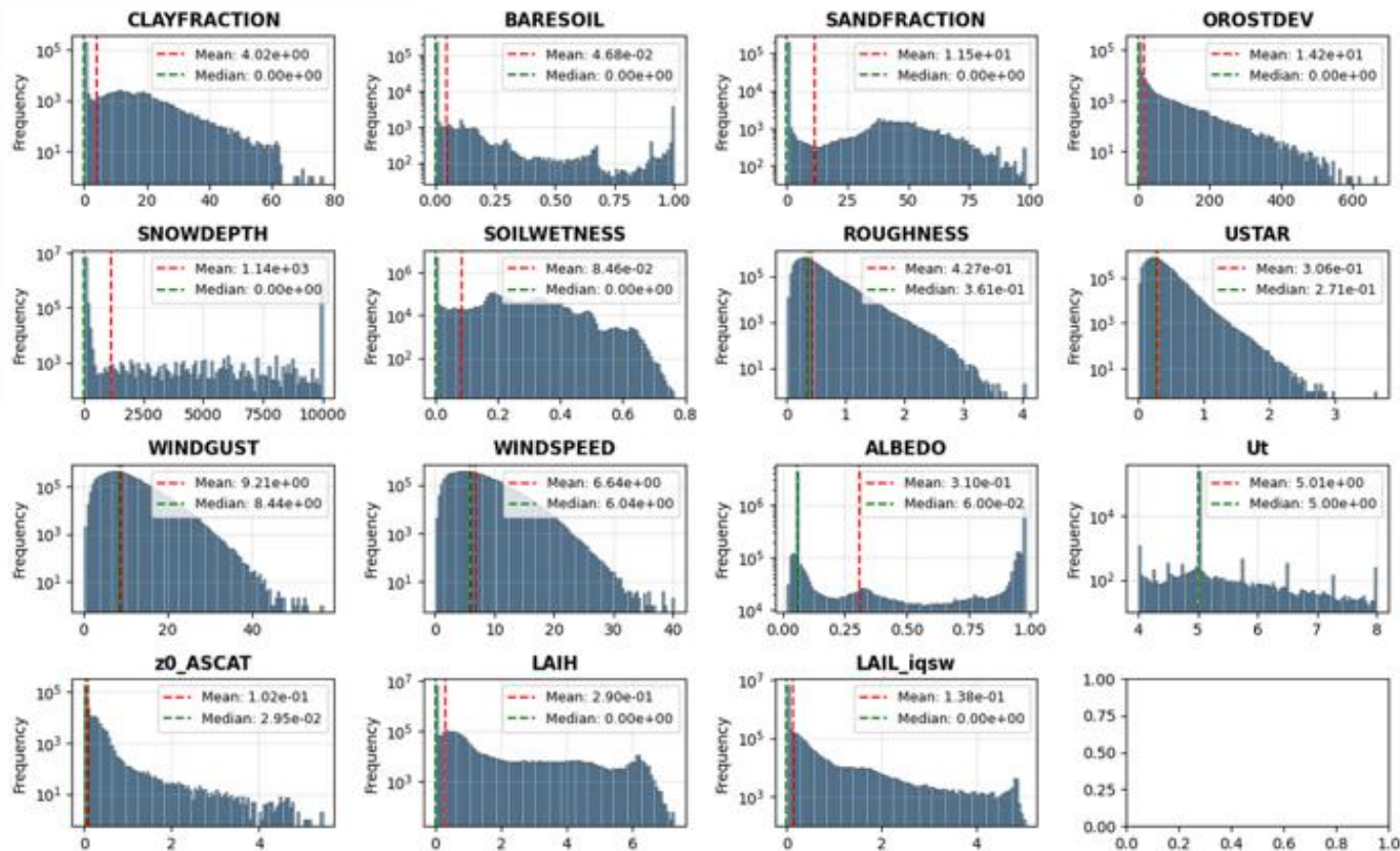
Key idea: Build a probabilistic model (surrogate) of the objective function Use it to decide where to sample next. 100 trials executed.



# Hyperparameter optimization – input norm

**Min-Max normalization** preferred over Standard (z-score) normalization

**Why?** Our predictors do not follow a Gaussian distribution. Applying z-score normalization would not be appropriate, making Min-Max scaling the better choice for input features.

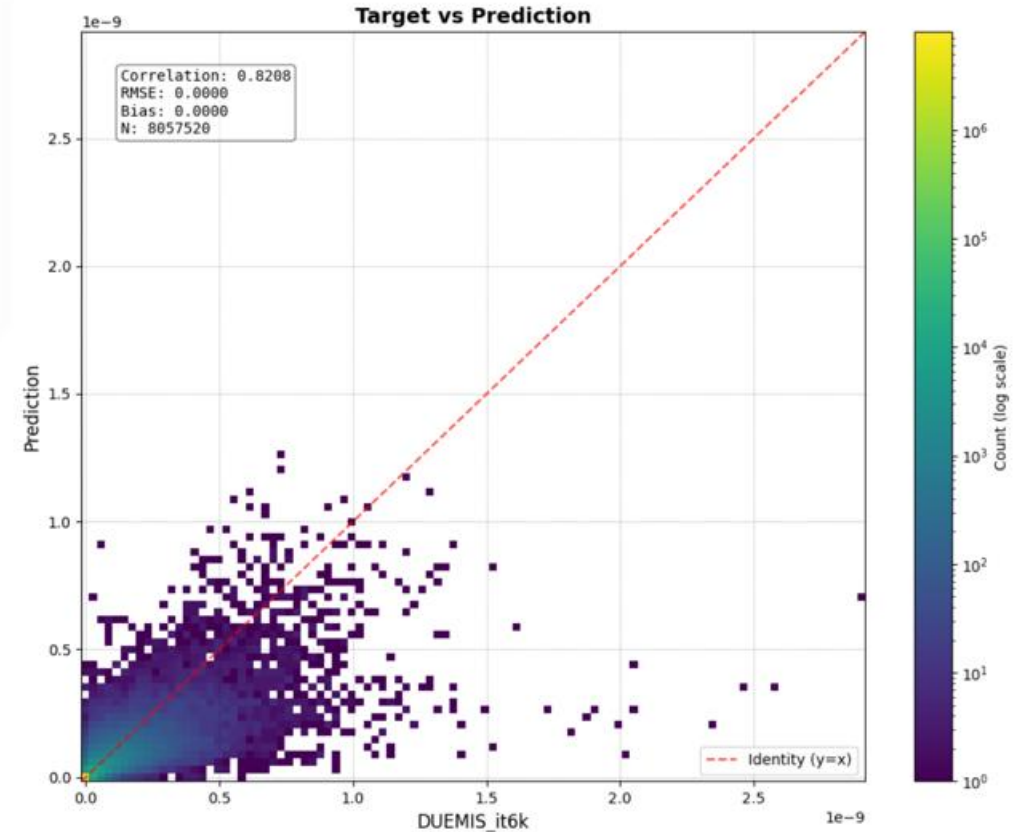




# Final pre-trained inference model architecture

Final configuration for our DNN model:

| Hyperparameter | Value           |
|----------------|-----------------|
| Layers         | [256,256,256]   |
| Activation     | GeLU            |
| Normalisation  | Inputs & Output |
| Dropout        | 0.2             |
| Loss           | r2              |
| Optimizer      | AdamW           |
| learning rate  | 1e-4            |
| batch size     | 256             |

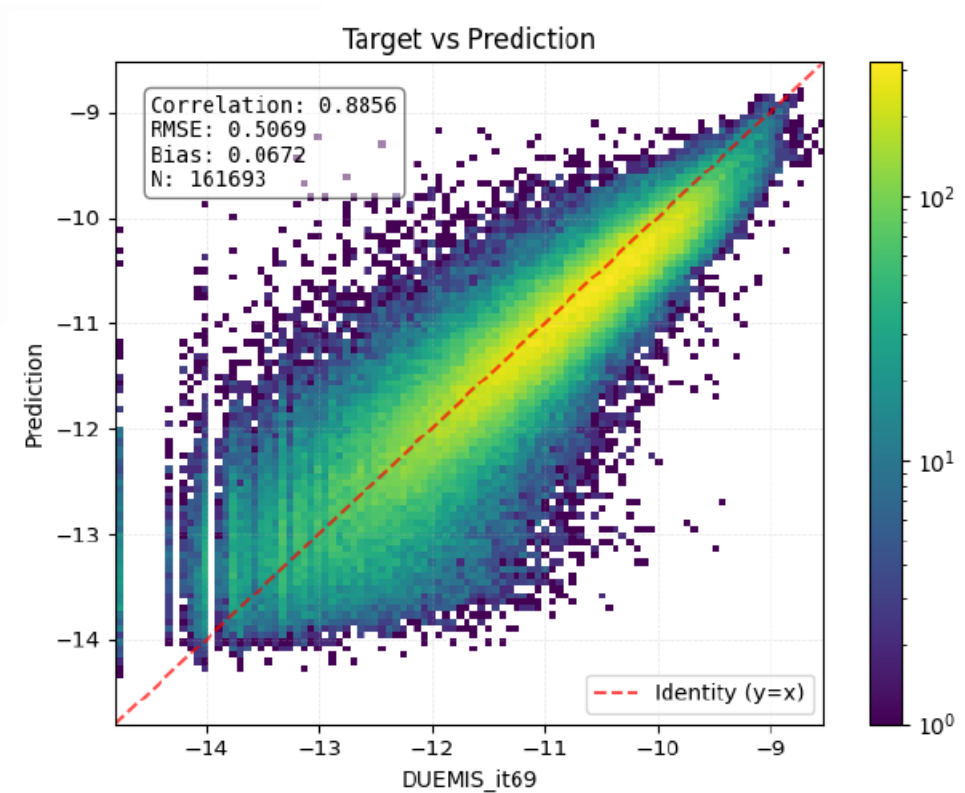




# Final pre-trained inference model architecture -v2

Final configuration for our DNN model:

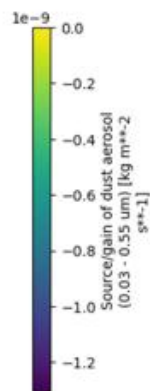
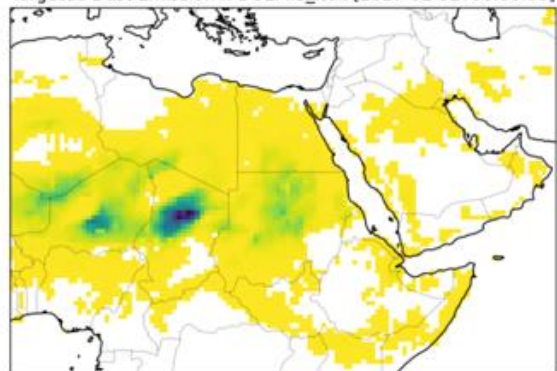
| Hyperparameter | Value           |
|----------------|-----------------|
| Layers         | [256,256,256]   |
| Activation     | GeLU            |
| Normalisation  | Inputs & Output |
| Dropout        | 0.2             |
| Loss           | r2              |
| Optimizer      | AdamW           |
| learning rate  | 1e-4            |
| batch size     | 256             |



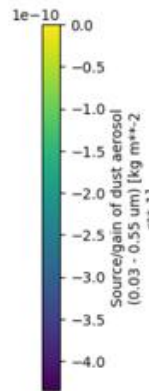
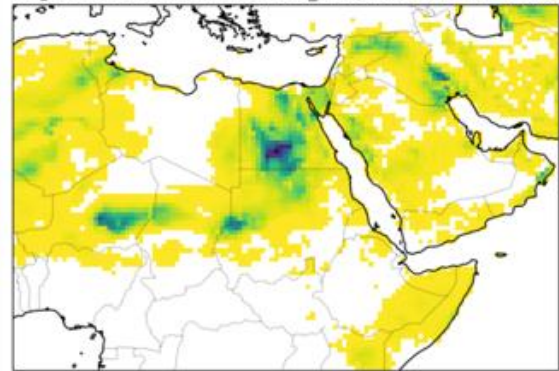


# Final pre-trained inference model architecture

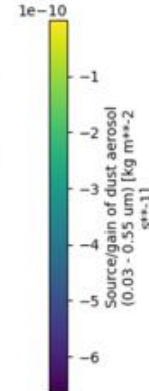
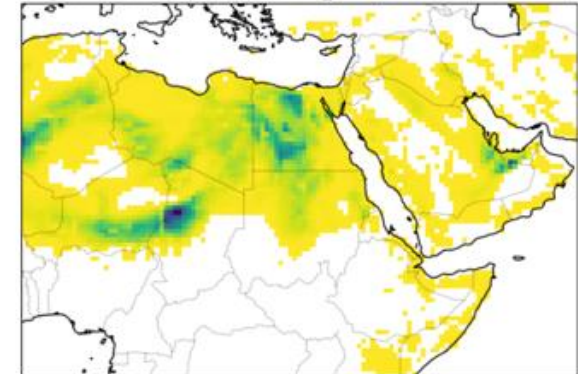
Targeted Dust Emission :: DUEMIS\_it6k [2017-01-01T00:00:00]



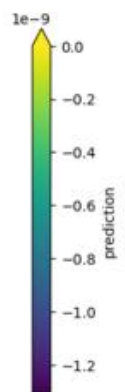
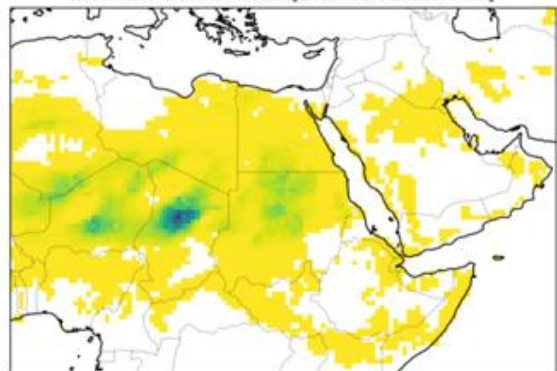
Targeted Dust Emission :: DUEMIS\_it6k [2017-06-11T00:00:00]



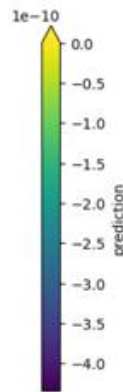
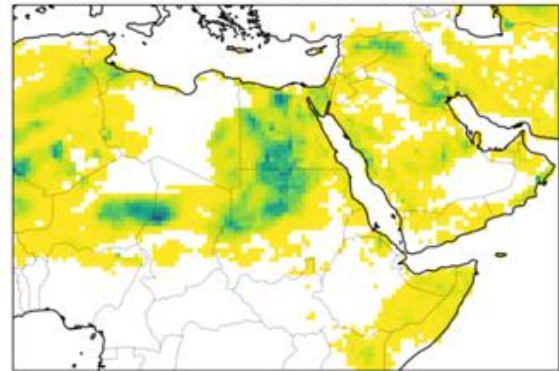
Targeted Dust Emission :: DUEMIS\_it6k [2017-10-12T00:00:00]



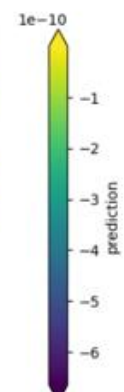
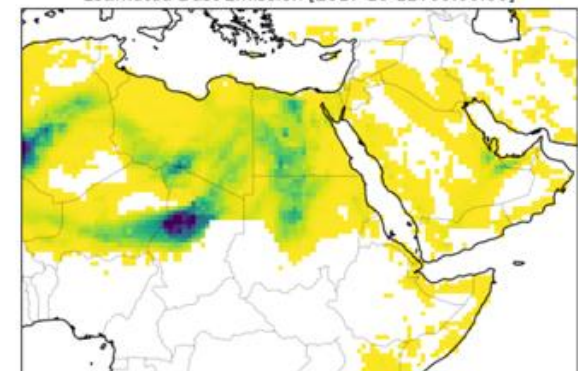
Estimated Dust Emission [2017-01-01T00:00:00]



Estimated Dust Emission [2017-06-11T00:00:00]



Estimated Dust Emission [2017-10-12T00:00:00]





PROGRAMME OF  
THE EUROPEAN UNION



IMPLEMENTED BY



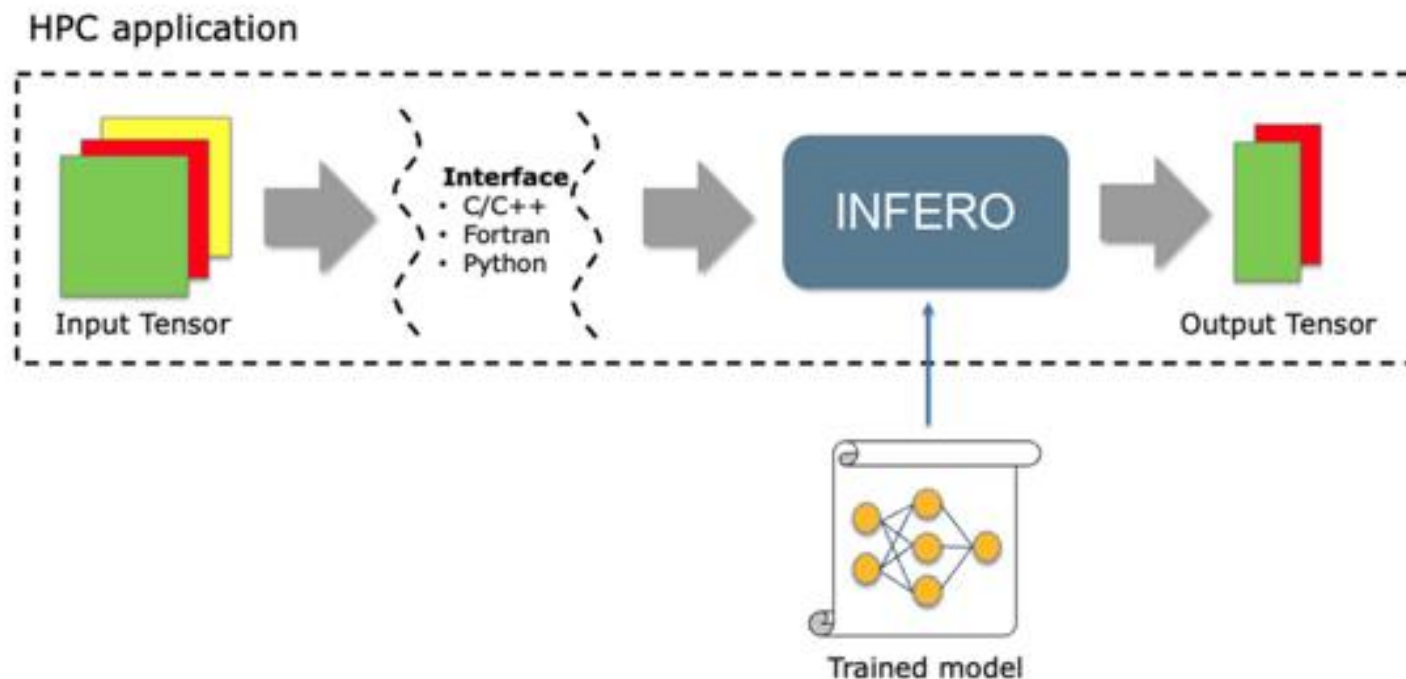
#EUSpace

# Implementation of desert dust emissions from neural network in IFS-COMPO

Rose-Cloé Meyer, Samuel Remy

# Use of the inference model in IFS-COMPO

- Use of the ECMWF INFERO library in cycle 50R1 IFS-COMPO (<https://infero.readthedocs.io/en/latest/>)
- INFERO allows to run a pre trained inference model (exported in ONNX format) in IFS-COMPO





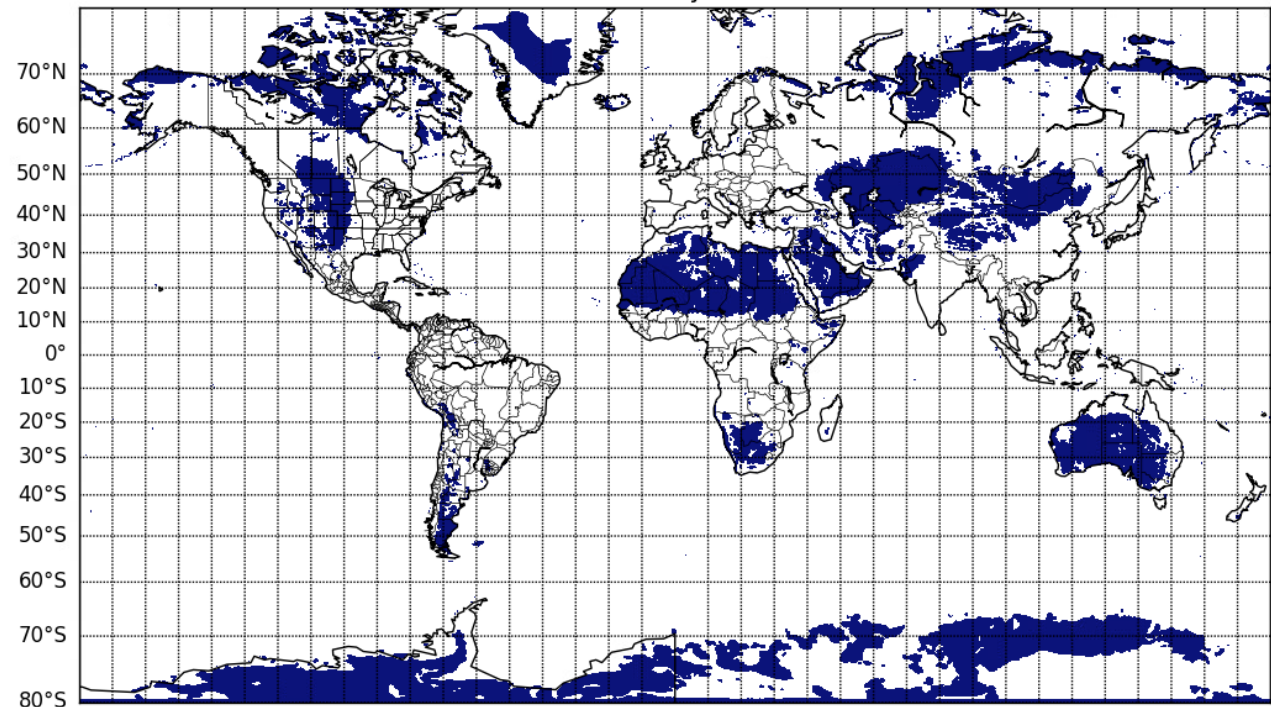
# Use of the inference model in IFS-COMPO

The use of desert dust emissions from NN has been tested into IFS-COMPO forecast only experiments with the following specifics :

- Cycle 50R1 + new dust emission scheme from CAMAERA WP5, adapted from SILAM + new dry deposition (proposed for 51R1)
- 2017 and second half of 2025 tested

Three experiments are shown:

- reference experiment using the physical scheme
- experiment using the full 15 predictors neural network for dust emissions
- experiment using a similar configuration but only in the blue area – outside, desert dust emissions are always null.



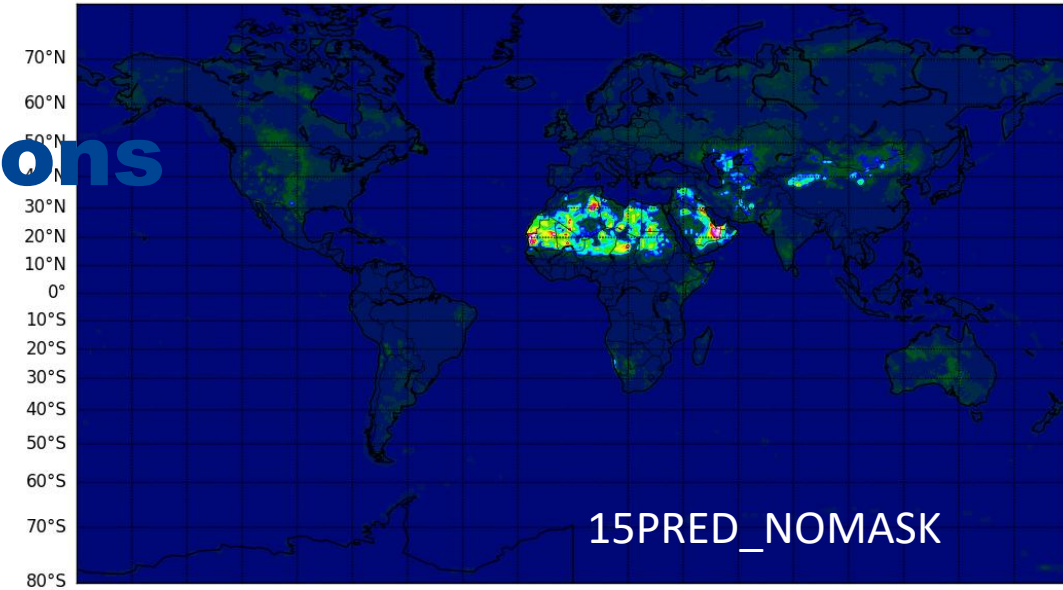


# Simulated desert dust emissions

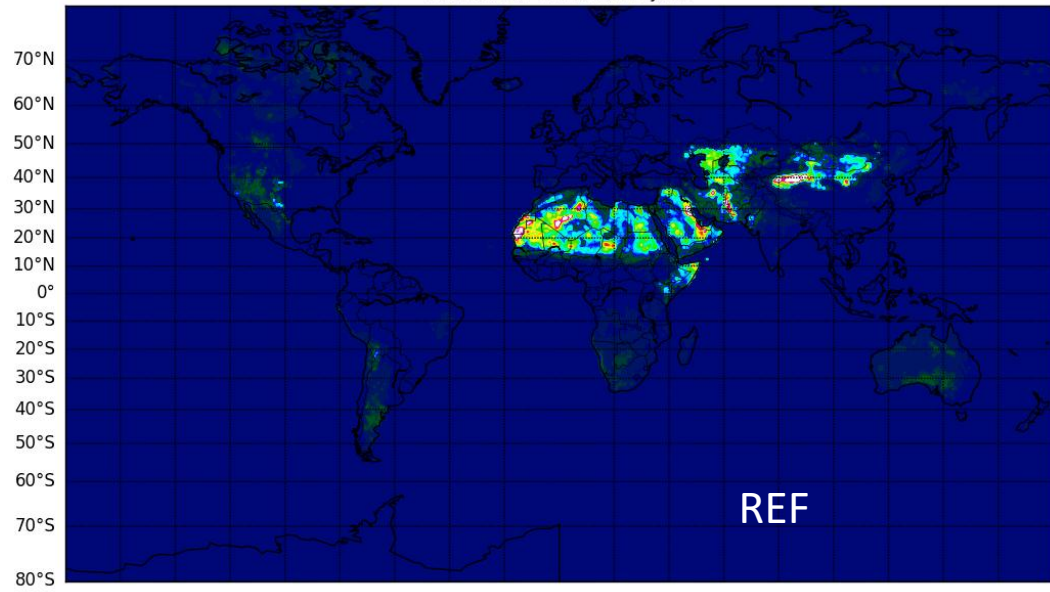
Simulated dust emissions averaged in June 2025:

- General features are quite similar
- High latitude dust sources visible over Canada
- NN predicts lower emissions over Taklimakan
- For experiment with (15 predictors, no mask), large areas with low but non null emissions

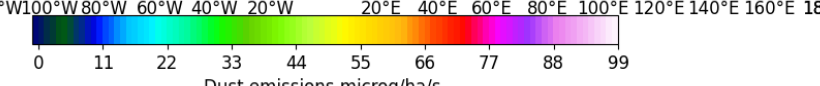
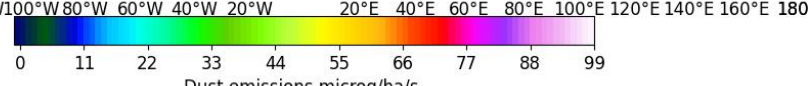
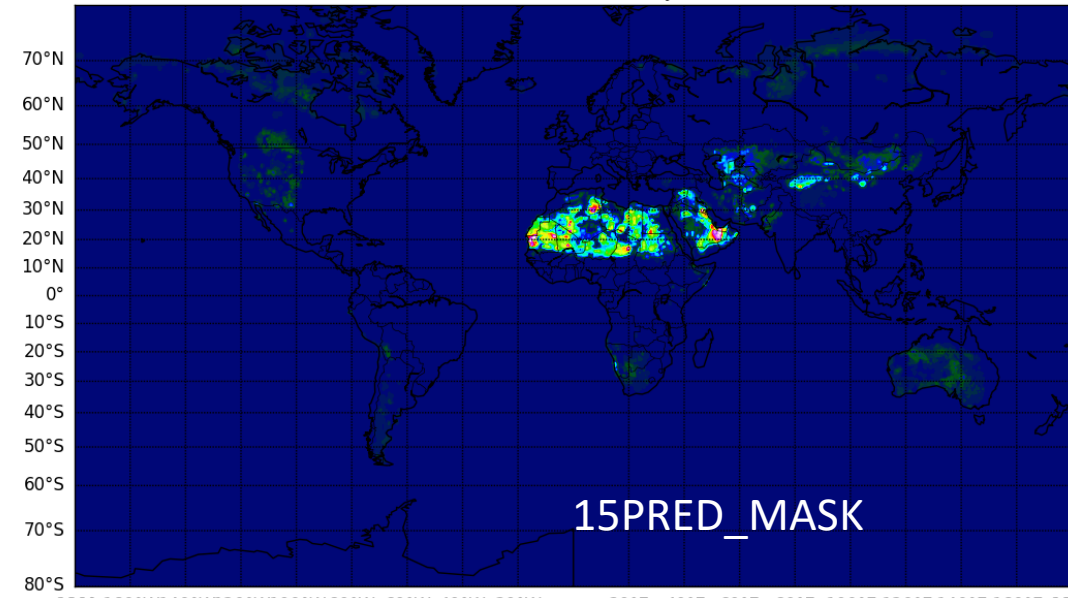
20250601 aersrcdus j2vq



20250601 aersrcdus j1ks



20250601 aersrcdus j53r

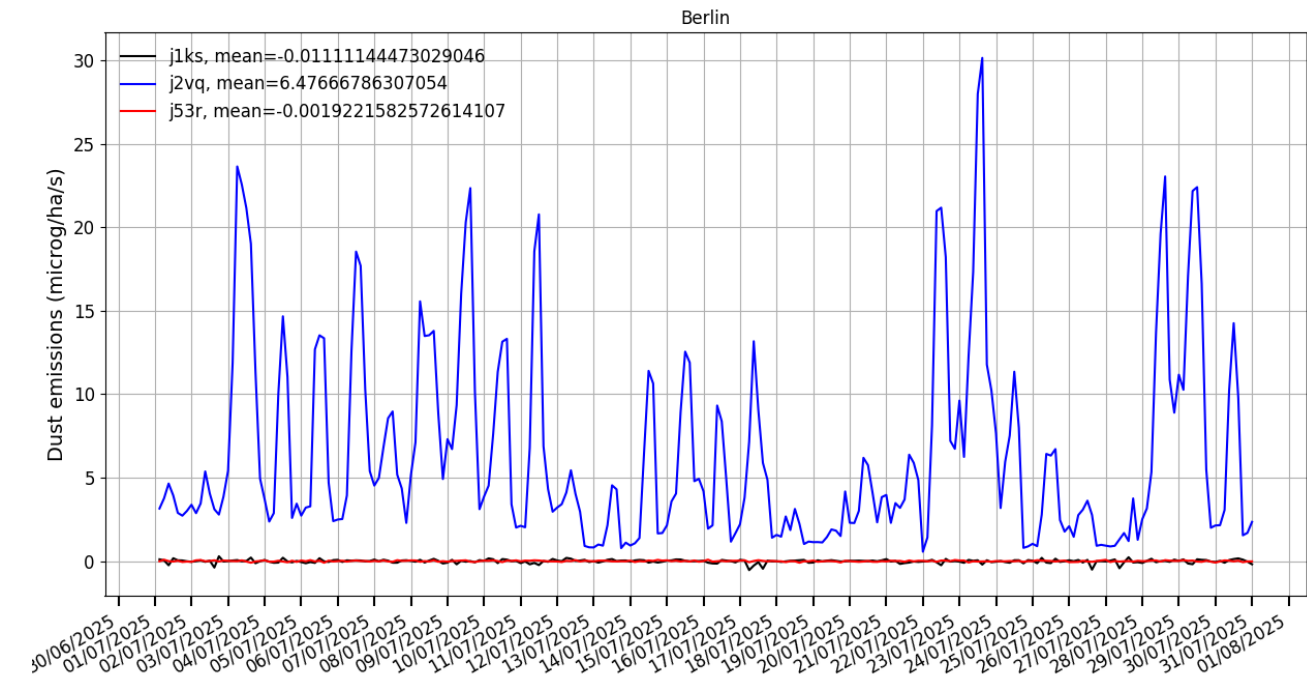
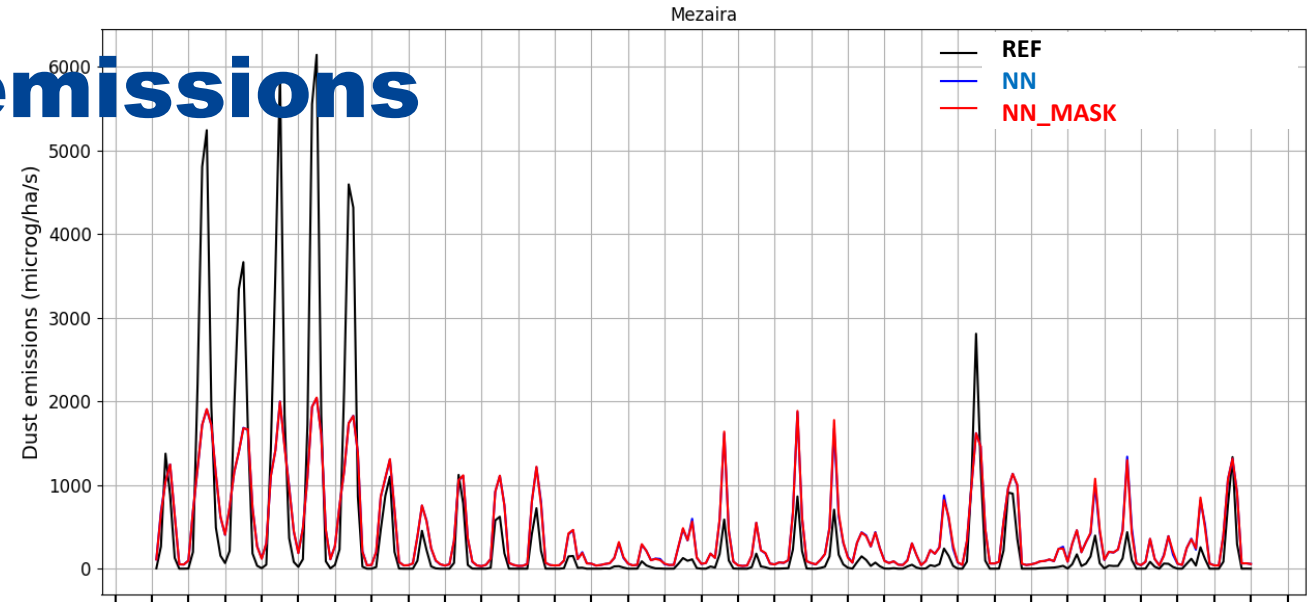




# Simulated desert dust emissions

Timeseries of simulated dust emissions over Mezaira (UAE – dusty region) and Berlin (Germany, not dusty!) show:

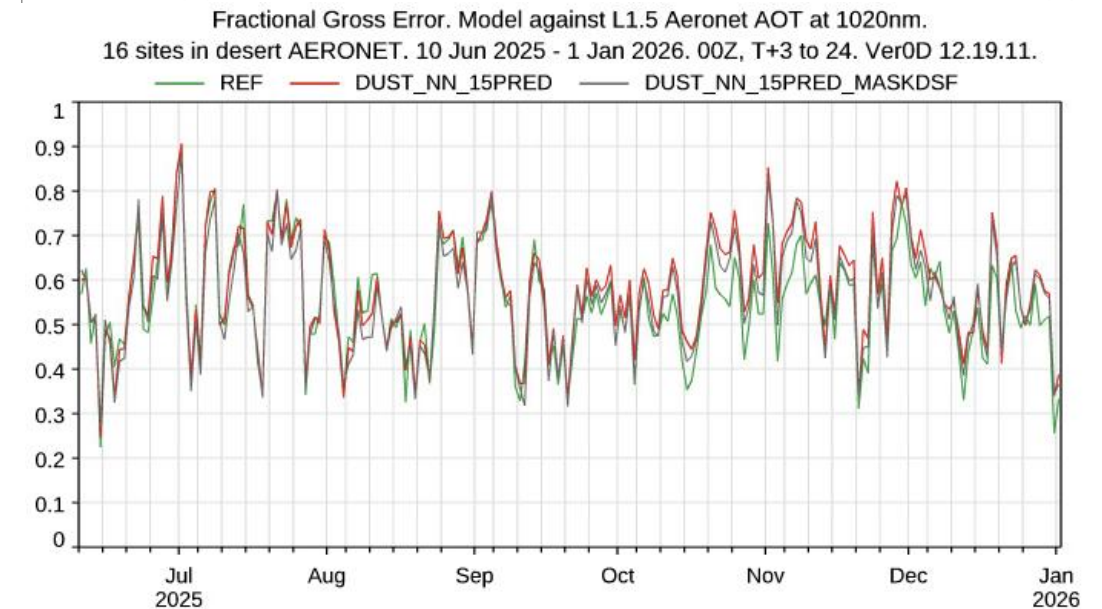
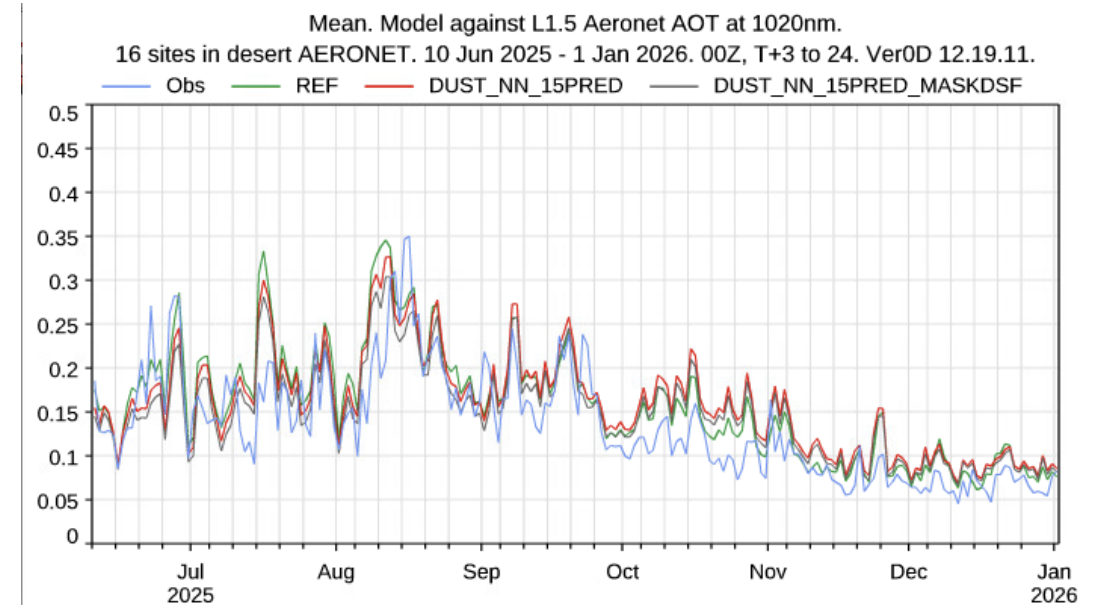
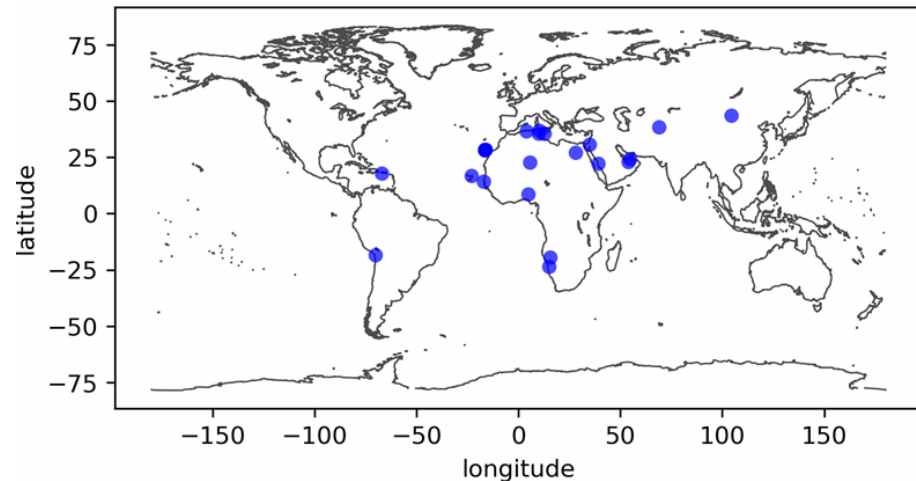
- NN without mask predicts low but still significant dust emissions over Berlin, effectively masked out with the use of a DSF based mask
- The occurrence of dust emissions in Mezaira is forecasted in quite a similar way
- The diurnal cycle shows a lower amplitude with NN : no null values at night, lower maxima in daytime





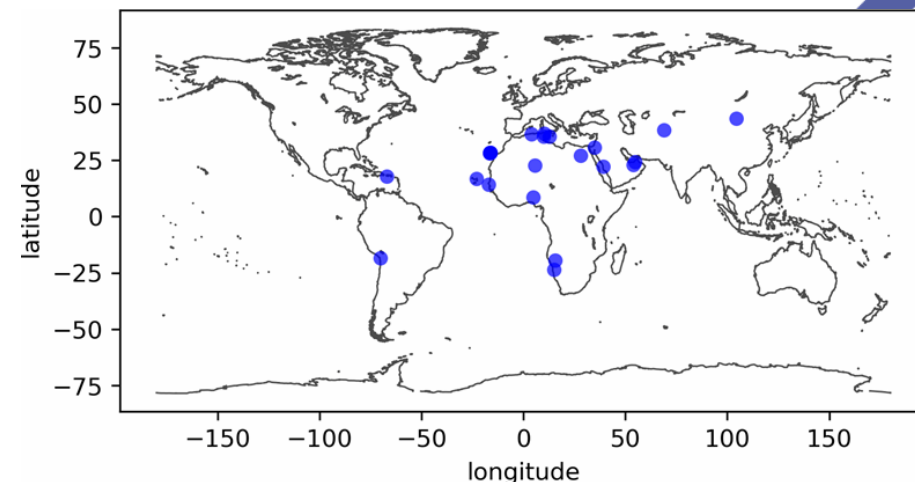
# Impact on simulated AOD

- A comparison of simulated AOD at 1020nm over a selection of « dusty » AERONET stations show comparable results
- No strong signal on error – sometimes better, sometimes worse

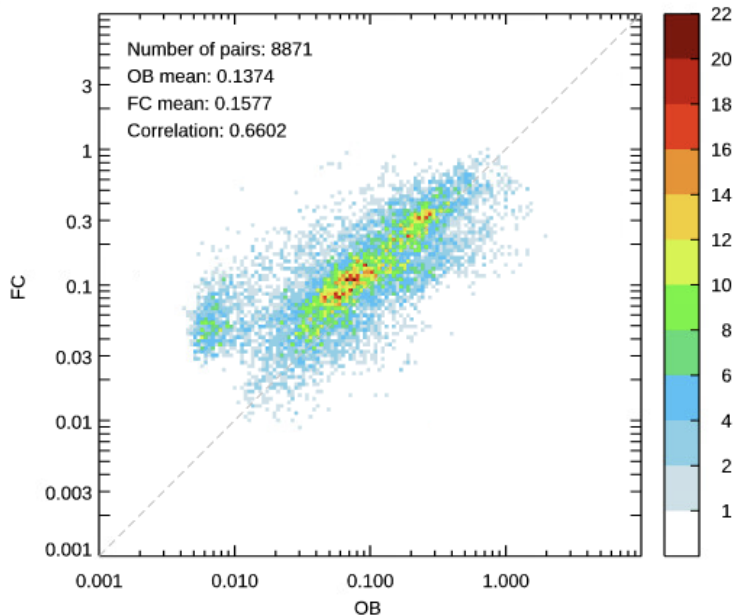


# Impact on simulated AOD

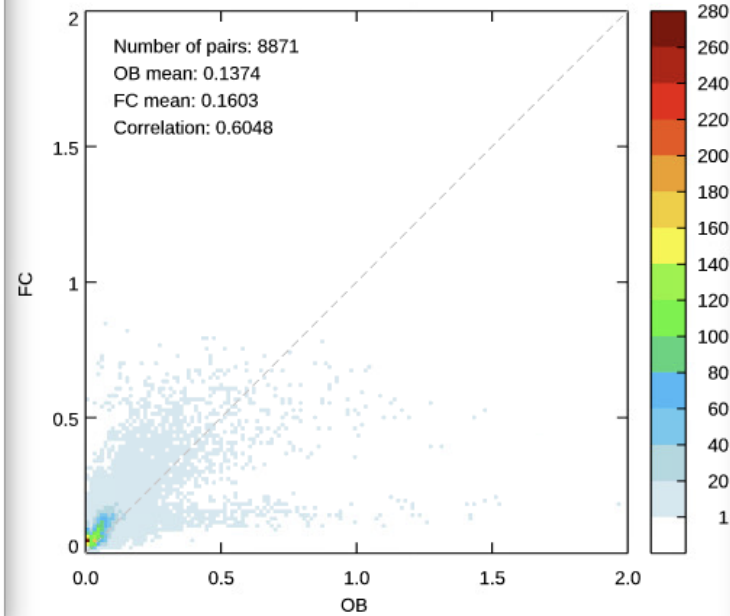
- A comparison of simulated AOD at 1020nm over a selection of « dusty » AERONET stations show comparable results
- No strong signal on error – sometimes better, sometimes worse
- Correlation is clearly degraded



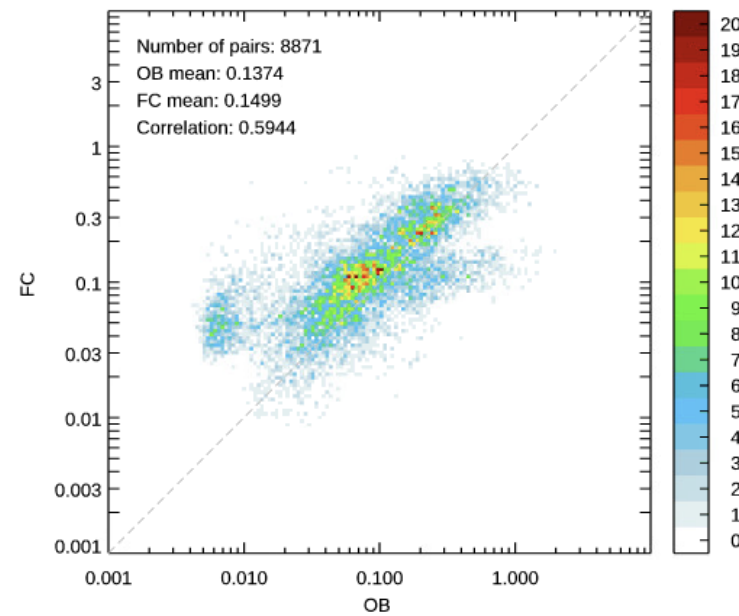
Model (REF) vs L1.5 Aeronet Any AE @ 1020nm  
10 Jun 2025 - 1 Jan 2026. 16 sites in desert AERONET.  
00Z, T+3 to 24. Ver0D 12.19.11.



Model (DUST\_NN\_15PRED) vs L1.5 Aeronet Any AE @ 1020nm  
10 Jun 2025 - 1 Jan 2026. 16 sites in desert AERONET.  
00Z, T+3 to 24. Ver0D 12.19.11.



Model (DUST\_NN\_15PRED\_MASKDSF) vs L1.5 Aeronet Any AE @ 1020nm  
10 Jun 2025 - 1 Jan 2026. 16 sites in desert AERONET.  
00Z, T+3 to 24. Ver0D 12.19.11.

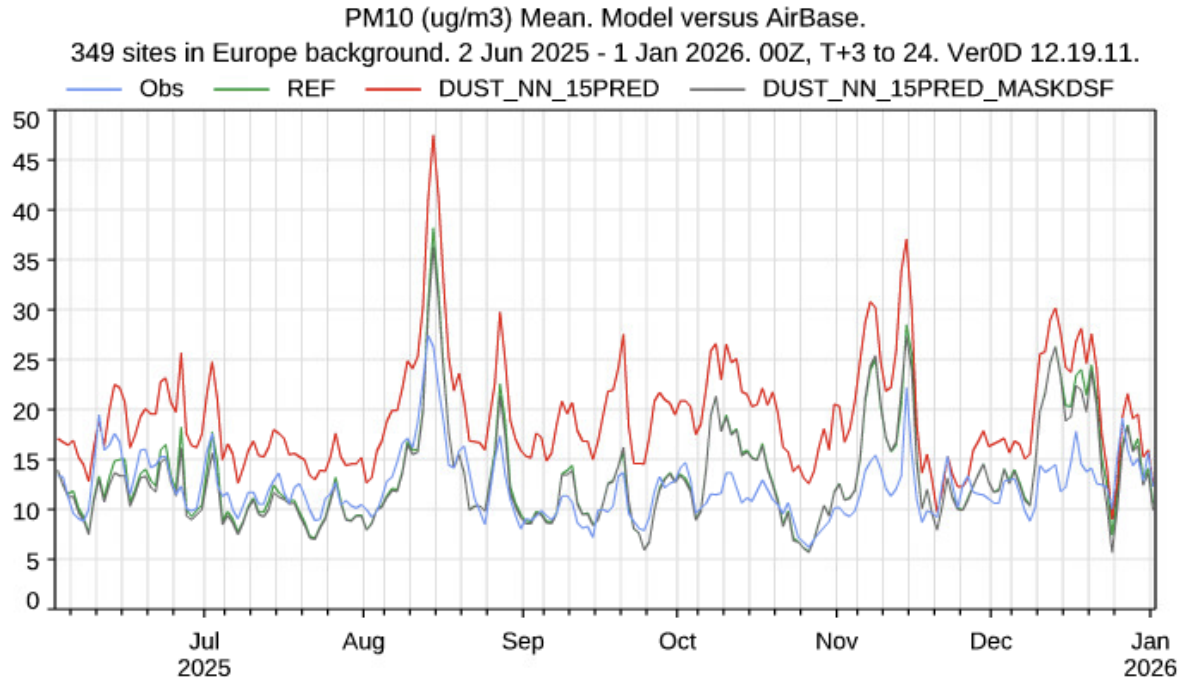




# Impact on simulated PM10 in Europe

A comparison of simulated vs observed PM10 over background rural stations in Europe show:

- A clear degradation from the small and persistent emissions with the NN
- Problem easily removed by the masking
- Skill scores very close





## Conclusion – what we have learnt and ongoing work

- NN dust emissions is possible but hard – much harder than whitecap fraction (for SS emissions) : skill scores on the test dataset are much lower than the training/validation
- The training is done on an analysis of (non null) desert dust emissions – which means that the NN doesn't know about null emissions => this needs to be taken into account elsewhere
- Theoretically, it could be possible to train a model including null emissions, but not sure it is worth it – probably better/easier to use NN only in areas where desert dust emissions are supposed to happen
- So some information from IFS-COMPO is useful/needed to use NN desert dust emissions => towards a “mixed” approach, NN with some physical constraints
- Ongoing work and next steps:
  - Retrain offline inference model with updated hyperparameters – error metric
  - Work on feature analysis – which predictors have more weight and “why”?
  - Work on predictors : some should be removed, and maybe some should be added, such as  $(u - u_t)^3$  for example?
  - Test convolution approach (U-Net) – not implementable in IFS-COMPO but could be informative