CAMS Service Evolution

D5.1 Implementation of new primary aerosol sources: sea-salt aerosol schemes from machine learning and chamber experiments

Due date of deliverable	30/6/2025
Submission date	30/6/2025
File Name	CAMAERA-D5.1-v1.0
Work Package /Task	WP5 / T5.1
Organisation Responsible of Deliverable	HYGEOS/FMI
Author name(s)	Mikhaïl Sofiev / Samuel Remy / Nathan Capon / Rose-Cloé Meyer / Bruce Denby /Elin Cecilie Ristorp Aas
Revision number	1.0
Status	
Dissemination Level	PU

The CAMAERA project (grant agreement No 101134927) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

Primary aerosol emissions are key drivers of the skill of key regional and global CAMS aerosol products: PM2.5, PM10 and Aerosol Optical Depth (AOD). Some of these emissions are provided as inputs to the regional and global (IFS-COMPO) CAMS systems, such as for organic and black carbon, and some of them are computed online, such as for sea-salt aerosol and desert dust. This report summarizes efforts on improving the representation of sea-salt aerosols, road dust and desert dust emissions carried out in task 5.1 of Work Package 5. This task covered a wide range of activities, involving several modeling groups and models (IFS-COMPO, SILAM, EMEP, LOTOS-EUROS, EURAD-IM). The results achieved can be summarized:

- A statistical fit between whitecap fraction and windspeed, using a dataset of remote sensing derived whitecap fraction, has been derived. Whitecap fraction is a proxy for sea-salt aerosol emissions,
- Machine learning techniques has been used to estimate whitecap fraction using a range of meteorological and oceaning precursors and to export an inference model,
- This inference model has been used in IFS-COMPO, which represents the first use of ML techniques to represent an atmospheric composition process in the global CAMS system,
- A dataset of gridded road dust (non exhaust road) emissions over Europe, covering two years, has been created, using a gridded version of the NOn-exhaust Road Traffic Induced Particle emission model (NORTRIP),
- This dataset has been used in the EMEP model and shown to impact significantly and positively the simulation of PM10 in European countries most impacted by road dust emissions,
- A new dust emission scheme has been implemented in IFS-COMPO, adapted from the SILAM dust emission scheme. Several options has been tested, and a thorough evaluation showed an improvement of most of dust-related CAMS products,
- The representation of high latitude dust in IFS-COMPO has been improved, for both the current dust emission scheme and the proposed new dust emission scheme. The impact has been evaluated mostly over Iceland, where numerous studies focusing on local dust sources have been carried out.

Table of Contents

Table des matières

1. Introduction

1.1 Background

The European Union's flagship Space programme Copernicus provides a key service to the European society, turning investments in space-infrastructure into high-quality information products. The Copernicus Atmosphere Monitoring Service (CAMS, https://atmosphere.copernicus.eu) exploits the information content of Earth-Observation data to monitor the composition of the atmosphere. By combining satellite observations with numerical modelling by means of data assimilation and inversion techniques, CAMS provides in near-real time a wealth of information to answer questions related to air quality, climate change and air pollution and its mitigation, energy, agriculture, etc. CAMS provides both global atmospheric composition products, using the Integrated Forecasting System (IFS) of ECMWF - hereafter denoted the global production system -, and regional European products, provided by an ensemble of eleven regional models - the regional production system.

The CAMS AERosol Advancement (CAMAERA) project will provide strong improvements of the aerosol modelling capabilities of the regional and global systems, on the assimilation of new sources of data, and on a better representation of secondary aerosols and their precursor gases. In this way CAMAERA will enhance the quality of key products of the CAMS service and therefore help CAMS to better respond to user needs such as air pollutant monitoring, along with the fulfilment of sustainable development goals. To achieve this purpose CAMAERA will develop new prototype service elements of CAMS, beyond the current state-of-art. It will do so in very close collaboration with the CAMS service providers, as well as other tier-3 projects. In particular CAMAERA will complement research topics addressed in CAMEO, which focuses on the preparation for novel satellite data, improvements of the data assimilation and inversion capabilities of the CAMS production system, and the provision of uncertainty information of CAMS products.

1.2 Scope of this deliverable

1.2.1 Objectives of this deliverables

This deliverable reports on the following activities:

Machine learning approach for sea-salt aerosol emissions in the IFS with possible use in regional models

Sea-salt aerosol emission schemes based on a dataset of retrieved whitecap fraction

Implementation into the IFS and SILAM of a scheme that represents high latitude dust

Development of a gridded version of the NORTRIP road dust emissions scheme and implementation into EMEP, EURAD-IM and LOTOS-EUROS

2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP5 T5.1) was performed.

2.1 Deviations and counter measures

The gridded NORTRIP (GNORTRIP) dataset of non exhaust traffic emissions has been developed and released in mid May 2025. Four models (EMEP, EURAD-IM, LOTOS-EUROS and SILAM) are currently carrying out simulations using the GNORTRIP emissions, but results

from EURAD-IM, LOTOS-EUROS and SILAM are not available yet to be included in this report. The results with EMEP are presented in the relevant section.

Also, the bubble chamber approach for sea-salt aerosol emissions didn't yield the desired results, and FMI instead implemented a whitecap-based approach for sea-salt aerosol emissions.

2.2 CAMAERA Project Partners:

HYGEOS	HYGEOS SARL
ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
Met Norway	METEOROLOGISK INSTITUTT
RC.io	RESEARCHCONCEPTS IO
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMi
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT
FMI	ILMATIETEEN LAITOS
MF	METEO-FRANCE
TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
INERIS	INSTITUT NATIONAL DE L ENVIRONNEMENT INDUSTRIEL ET DES RISQUES - INERIS
IOS-PIB	INSTYTUT OCHRONY SRODOWISKA - PANSTWOWY INSTYTUT BADAWCZY
FZJ	FORSCHUNGSZENTRUM JULICH GMBH
AU	AARHUS UNIVERSITET
ENEA	AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

3 Sea-salt aerosol emission scheme from whitecap fraction and chamber experiments

The basis of the sea-salt aerosol emission module is the whitecap fraction dataset from the satellite measurements as provided by Anguelova et al. (2019) using meteorological precursors. This dataset immediately reveals that the retrieval frequency strongly affects the estimated whitecap fractions. From the four dataset frequencies two frequencies were selected for further analysis: horizontally polarized retrievals at 10GHz and 37GHz frequency. From these the H37HGz data indicates larger whitecap fractions than the data from H10GHz retrieval. This is illustrated in Fig. 1.1, where the annually averaged whitecap fraction is plotted, by averaging the different retrievals at each location.

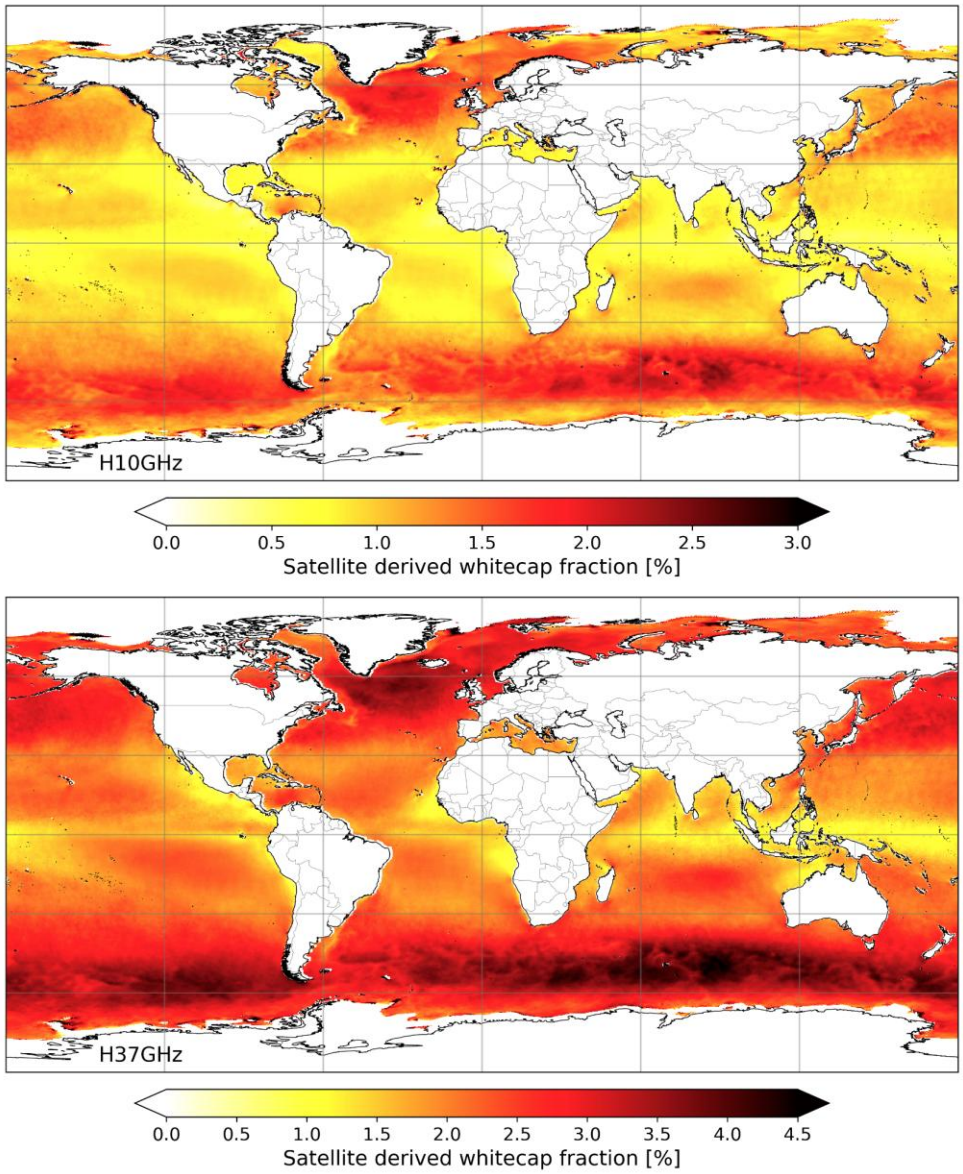


Figure 1.1: Satellite derived whitecap fraction [%], averaged over different overpasses and times for the full year 2014. The top panel is for horizontally polarized retrieval at 10GHz frequency, while the bottom panel is for horizontally polarized retrieval at 37GHz frequency. Note the different scales for different panels.

Since all the current models assume that the whitecap fraction depends most strongly on the wind speed, the ERA-Interim originated 10m wind speed (U_{10}) was selected as variable to use in the fit for velocity dependence. In Fig. 1.2 the 2D-histogram (contour plot) illustrates the number of cases that correspond to velocity - whitecap-fraction pairs when using 0.2m/s bin size for velocity, and 0.05% bin size for whitecap fraction. The blue circles additionally note the mean whitecap fraction inside each velocity bin, which is used to find the velocity dependence for the whitecap fraction. The traditional pure cubic dependence (red solid curve) clearly provides too low estimates at low velocities while this estimation is too large at highest velocities. Adding a constant off-set improves the situation (red-dashed), but a better fit is obtained by using squared dependence on velocity with additional constant term, as indicated by the black dashed line. Similar conclusions can be obtained if using median value for whitecap fraction (WCF) inside each velocity bin. However, when using the mean value the bias of the model becomes smaller.

When the whitecap fraction inside each bin is calculated using the mean, the optimum squared curve with constant term ($a*U_{10}^2+b$) becomes:

WCF[%] = 0.01261257*U_{10}^2+0.26117591 (H10GHz) Bias: 0.014005, r-value = 0.8867

WCF[%] = 0.01790888*U_{10}^2+1.20001489 (H37GHz) Bias:-0.007515, r-value = 0.8744

If median value is used to estimate the whitecap fraction inside each velocity bin, the optimum curves become:

WCF[%] = 0.01267127*U_{10}^2+0.19834430 (H10GHz) Bias:-0.044713, r-value = 0.8867

WCF[%] = 0.01801923*U_{10}^2+1.12888682 (H37GHz) Bias:-0.070904, r-value = 0.8744

Here velocity is in m/s, and the r-value is Pearson correllation coefficient of the model data to whitecap fraction data (all data points for 2024), while the bias is the mean difference of yearly averaged model-data and the yearly averaged satellite data for whitecap fraction. Figure 1.3 illustrates the yearly averaged model values together with the difference to satellite data for both frequencies.

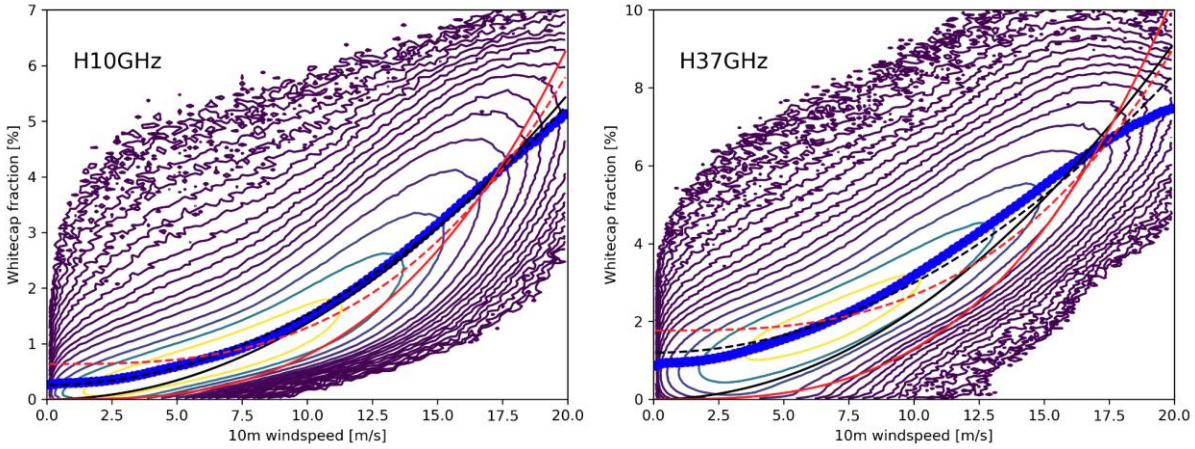


Figure 1.2: Contour plot for number of observed satellite retrieved whitecap fraction for the full year 2014 when using bin size of 0.2m/s for 10m wind-speed, and bin size of 0.05% for whitecap fraction. The used isosurfaces are: [10, 15, 25, 40, 65, 100, 150, 250, 400, 600, 1000, 1500, 2500, 5000, 10000, 20000, 50000]. Left panel is for H10GHz channel and right panel is for H37GHz channel. The blue circles denote the mean whitecap fraction inside each velocity bin, while the red solid line is the optimum traditional cubic fit for mean values. The red-dashed curve has additional constant term in addition to cubic term. The black solid curve is squared fit and the black dashed line has additional constant term.

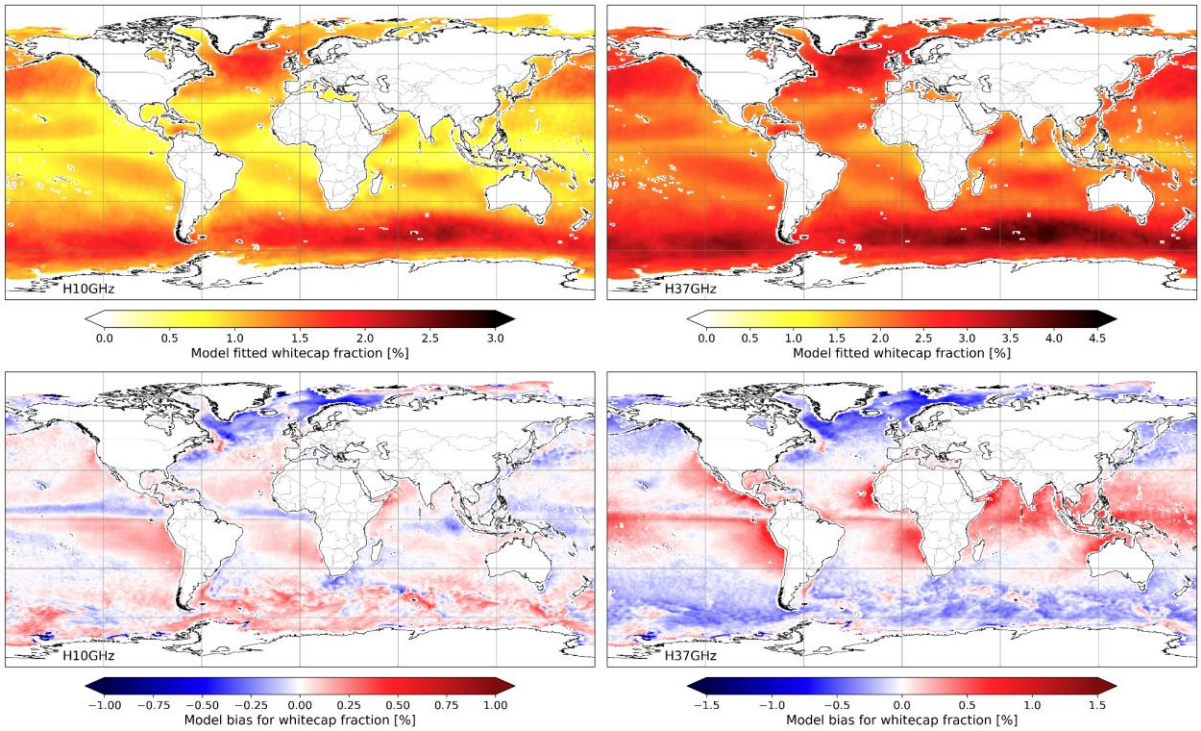


Figure 1.3: Time averaged (year 2014) modelled whitecap fraction (top panels) and the model bias (bottom panels) when the whitecap fraction is fitted with squared term plus a constant when using the mean to average the whitecap fraction in each velocity bin (black dashed line in Fig. F2). The left panels are for H10GHz channel, and the right panels for H37GHz channel.

Even if the number of observations for the year 2014 is large (over 100 million), one can also do a direct fit to all the data points. Since the previous estimation indicates that the most important term is likely the velocity squared term, we plotted the 2D-histogram (contour) in Fig. 1.4 such that the horizontal axis is the squared velocity with a bin size of 1.0 m^2/s^2. The fitted curves in this figure are the velocity squared fit with constant term (black dashed line), plus a similar fit (magenta) with a cubic term added ($a*U_{10}^3+b*U_{10}^2+c$). The optimum curves with cubic term are:

WCF[%] = $0.000122864540*U_{10}^3+0.0106628220*U_{10}^2+0.295313625$ (H10GHz) Bias: 0.002348, r-value = 0.8872

WCF[%] = $-0.000522215212*U_{10}^3+0.0281498948*U_{10}^2+0.898571366$ (H37GHz) Bias: 0.023044, r-value = 0.8776

As one can see the correlation of the model does not improve much when doing a direct fit to the data. Also, for 37GHz the bias is worse than when fitting with the data obtained by taking mean values inside each velocity bins. Interestingly, even if the direct fit should weight the small and intermediate velocities the most (since there are much less high velocity data points), visually the both the low and the high velocity data is followed better with these fits than with the fits where mean/median curve is used as a basis for fitting. For 10GHz data the cubic term tends to slightly increase the whitecap fraction at high velocities, while the 37GHz data indicates slower growth at high velocities and results in a small negative coefficient for the cubic term. However, all the fits have non-zero constant term, indicating that whitecaps might persist long-enough even in low wind conditions.

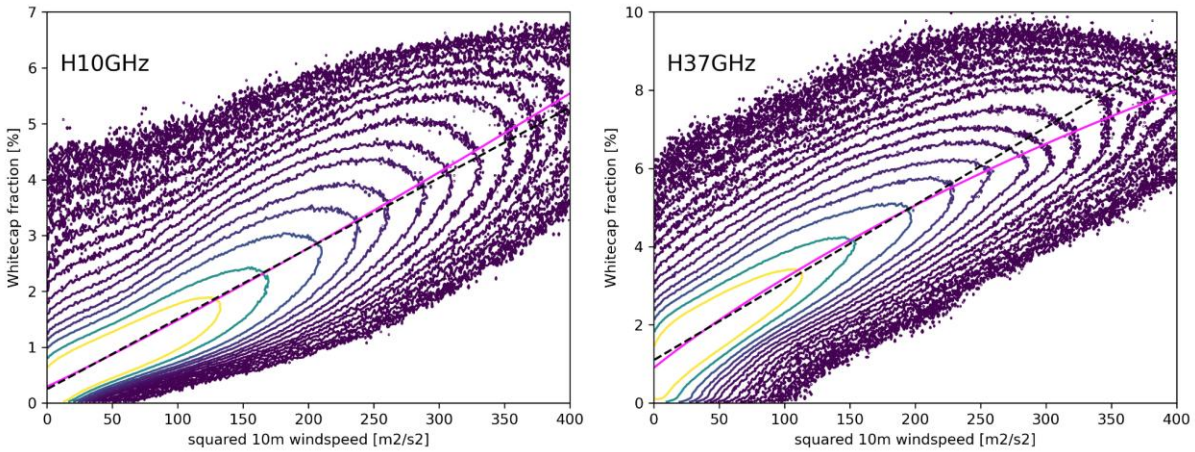


Figure 1.4: Contour plot for observed whitecap fraction when the horizontal axis is selected to be squared 10m wind-speed with bin size of 1.0 m^2/s^2, and vertical axis is whitecap fraction with bin size of 0.05%. Plotted isosurfaces are: [10, 15, 25, 40, 65, 100, 150, 250, 400, 600, 1000, 1500, 2500, 5000, 10000] . The black dashed line is velocity squared fit with constant off-set when using all the satellite observed points for year 2014. The magenta curve contains additionally a term proportional to velocity cubed in the fit. The left panel is for data in the H10GHz channel, while the right panel is for the H37GHz channel.

The modelled annual mean whitecap fraction and the bias from the direct fit is illustrated in Fig. 1.5. The performance is very similar to the case where the fitting was done using the curve generated by averaging the whitecap fraction inside each velocity bin. There are regions where the model either under-estimates (like Northern Atlantic) or overestimates the satellite derived whitecap fraction for both frequencies. However, there are also regions where the bias has different signs for different frequencies (Western Pacific close to Indonesia, and several sea regions closer to Antarctica). This illustrates that the satellite retrieval might catch different processes when the frequency is changed. Same thing is likely supported by the different dependence at high frequency.

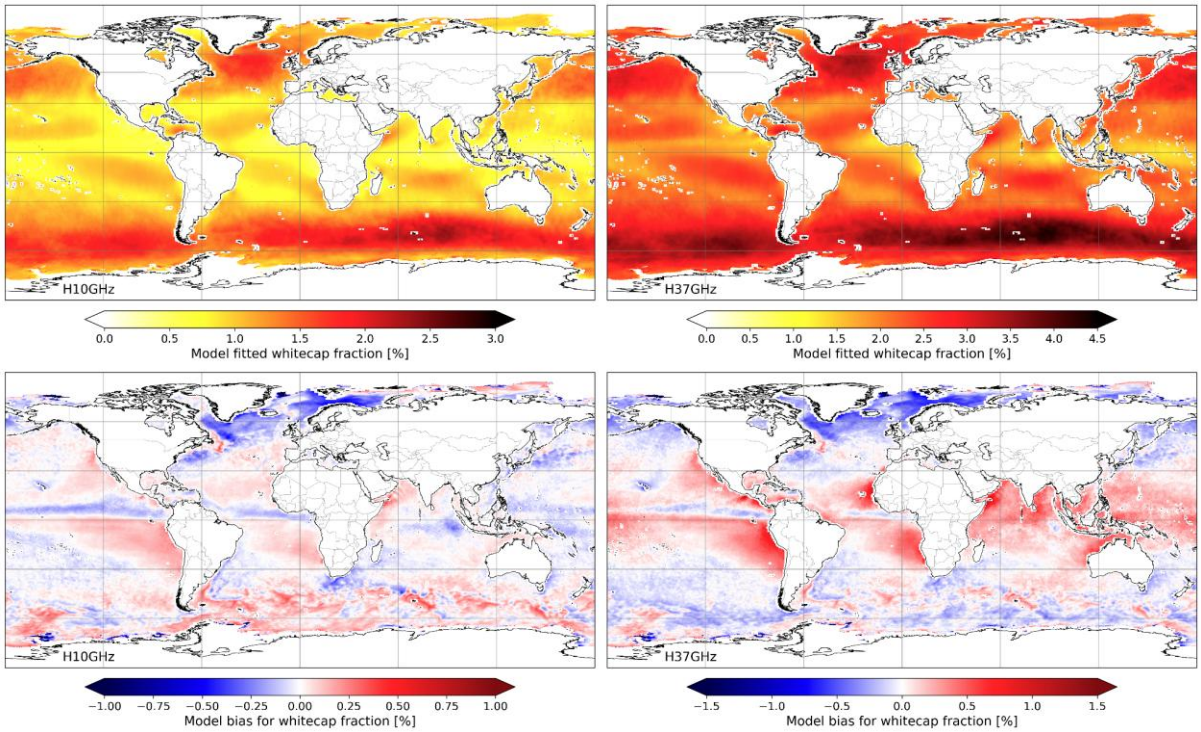


Figure 1.5: Time averaged whitecap fraction from the model where all the year 2014 data is fitted with formula: WCF = $a*U_{10}^3+b*U_{10}^2+c$ as given by the magenta curve in Fig. F4. The left panel is for data in the H10GHz channel, while the right panel is for the H37GHz channel.

4 Sea-salt aerosol emission scheme using machine learning approach

This section aims to improve the Sea-salt aerosol emission scheme by focusing on Whitecap fraction estimation using Machine Learning approach. Whitecaps primarily occur when waves break or collide; they are a good proxy for sea-salt aerosol emissions. Estimating the amount of aerosol they produce is particularly important for forecasting cloud formation and correcting satellite observations. in IFS-COMPO, the whitecap fraction is estimated using polynomial fits from remote sensing satellite acquisitions. This study aims to compare such traditional models with machine learning approaches providing a benchmark, and as a second step, a machine learning approach is integrated into the global IFS-COMPO model.

4.1 Parameterizations and dataset

Currently, the whitecap fraction (WF) is estimated in the cycle 49R1 IFS-COMPO using the Albert 16 et al. (A16) parameterization (cf. Table 1), which uses 2 predictors: Wind Speed (WSP) and Sea surface temperature (SST). Based on that, the sea-salt aerosol emissions are derived using the Gong (2003) distribution. In this work, A16 has been compared with older fits such as Monahan et al. (M80), Callaghan et al. (CAL08), Salisbury et al. (S13) whose equations are given in Table 2.1, which also includes the new FMI parameterization presented in Section 3. The evaluation showed that A16 is by far the best estimate of whitecap fraction.

Table 2.1: Expressions of different arithmetic models

	Equation at 10.7 GHz	Equation at 37 GHz
M80	x	$3.84 \times 10^{-4} \times WSP^{3.41}$
CAL08	x	$4.82 \times 10^{-4} \times (WSP + 1.98)^3$
S13	$4.6 \times 10^{-3} \times WSP^{2.26}$	$3.97 \times 10^{-2} \times WSP^{1.59}$
A16	$a(SST) * (WSP - b(SST))^2$	$a(SST) * (WSP - b(SST))^2$
FMI	$1.27 \times 10^{-2} \times WSP^2 + 1.98 \times 10^{-1}$	$1.8 \times 10^{-2} \times WSP^2 + 1.29$

These whitecap fraction models will be compared to a new dataset from remote sensing designed by Anguelova et al. (2019). This dataset contains two years of whitecap fraction estimations with daily resolution and have been derived from WindSAT satellite acquisitions. This sensor is a multi-frequency polarimetric radiometer of the U.S. Navy, which has been launched for measuring ocean surface wind speed and direction. Given that this instrument operates primarily for frequencies in the microwave range, the frequencies are defined in GHz rather than in micrometers. Thus, knowing that emission and scattering properties of the whitecap vary according to wavelength, the dataset contains the whitecap fraction emissions derived at 10.7 GHz and at 37 GHz (cf. Figure 2.1) and will serve as a reference for the design of new models and the comparison of performance with arithmetic models.

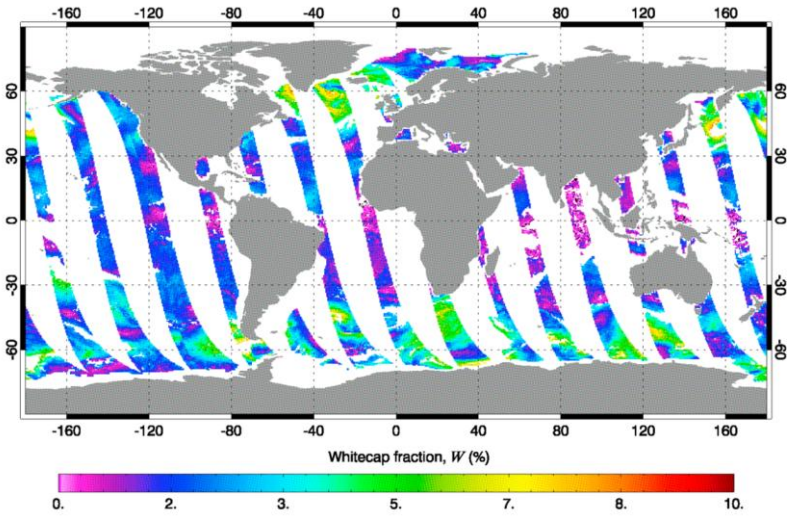


Figure 2.1 : Example of a daily map of whitecap fraction from several Windsat acquisitions.

To begin with, it is important to focus on the inputs of the models. To evaluate the performances of the arithmetic models, WSP and SST must be part of the inputs. Moreover, knowing that whitecap emissions mainly occur from water projection, it is relevant to add marine predictors, such as wave period and height. Finally, after several studies, it appears worthwhile to also consider turbulent kinetic energy from waves flux as input, especially for major events. Table 2.2 contains all the inputs taken into consideration for this study. Most of them are provided by the ERA5 reanalysis; the wave height total and the turbulent kinetic energy from ocean waves have been provided by an ECWAM reanalysis kindly made available by J. Bidlot and J. Kousal from ECMWF.

Table 2.2: Description of inputs for model training

Acronym	Fullname	Resolution	Unit	Provider
SST	Sea Surface Temperature	hourly	Kelvin	ERA5
WSP	Wind Speed	hourly	m/s	ERA5
WD	Wind Direction	hourly	Degree	ERA5
MWP	Mean Wave Period	hourly	s	ERA5
SWH	Significant Wave Height	hourly	m	ERA5
WHT	Wave Height Total	hourly	m	HINDCAST
PhiOC	Turbulent kinetic energy from ocean waves into the ocean	hourly	W/m²	HINDCAST

Then, the objective is to use the dataset of whitecap fraction from remote sensing and this set of predictors to estimate the whitecap fraction with deep neural networks and, as a second step, to integrate a selection of inference models into IFS-COMPO.

4.2 Machine Learning models

The first part of our work is to develop machine learning models using different methods and assess their performance.

4.2.1 Methodology for offline training

It is possible to train different models on the constructed dataset. As a first step, we carry out a benchmark of well-known Machine Learning algorithms, such as Random Forests, Gradient Boosting, Ridge regression, etc on our dataset. The objective is to better apprehend the data and identify where classical models can struggle, as well as to select the best performing machine learning model to simulate whitecap fraction.

The second and main source of interest is to implement the model designed in a publication of Zhou et al. from 2023, which is a traditional Deep Neural Network (DNN) of 3-layer deep. This model shows promising results but authors have used the same set for validation and testing, which seems to be hazardous to assess performance. Furthermore, after training this architecture on our data, we found that we can improve the performance by adding a pre-processing part, which corresponds to a scaling and a PCA transformation. We also try to tune the model by adapting the loss function and hyper-parameters used during the training to adapt the architecture to our dataset. The resulting architecture is described in Figure 2.2.

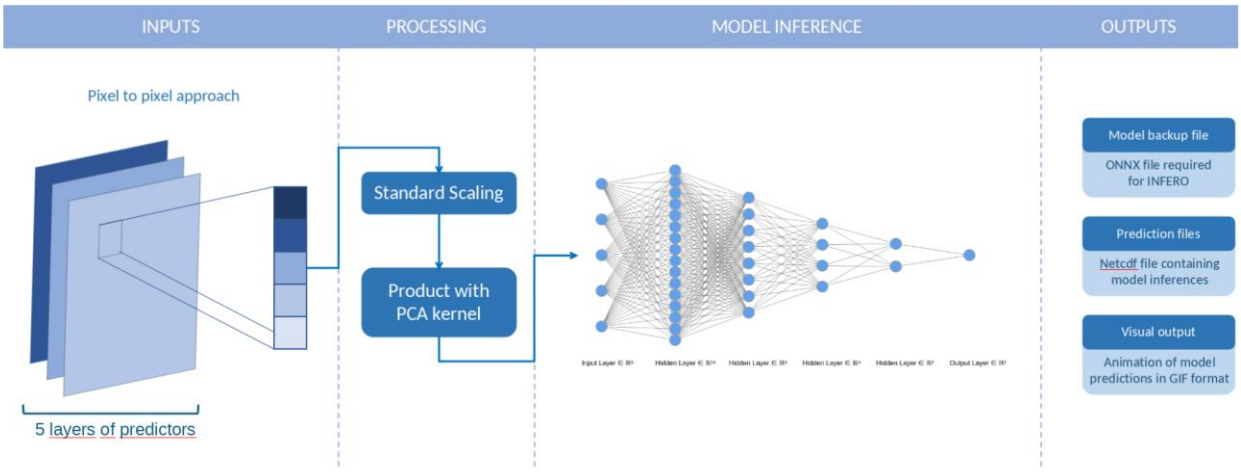


Figure 2.2 : Description of the Deep Neural Network (DNN) architecture and pre- and post-processing operations

All previous models use a pixel-to-pixel approach, which means that they use only the physical information of a single pixel to estimate whitecap fraction emission for this one. Despite the low resolution of our reference (0.5 degree by pixel), it seems interesting to use information of the neighbouring pixels as it could reduce the impact of outlier in our predictors. Thus, we also implemented spatial Deep Learning architectures, such as Convolutional Neural Network and U-Net.

4.2.2 Results and performances of the offline models

To train our model, the dataset has been split into three sets, named respectively *train*, *validation* and *test*. The *train* set is used to evaluate model performances and to adjust model parameters to improve those metrics.The *validation* set is a control set that prevents model to “overfit”, which means that the model adapts specifically to the *train* set and stops doing generic learning. Finally, the *test* set is a set that has never been encountered during training and allows us to predict the model's performance on a new data set that is different from the one used during training. In our case, the *test* set is every acquisition collected between January and June 2017 and *train* and *validation* sets are obtained with a random split of acquisitions collected over the year 2014.

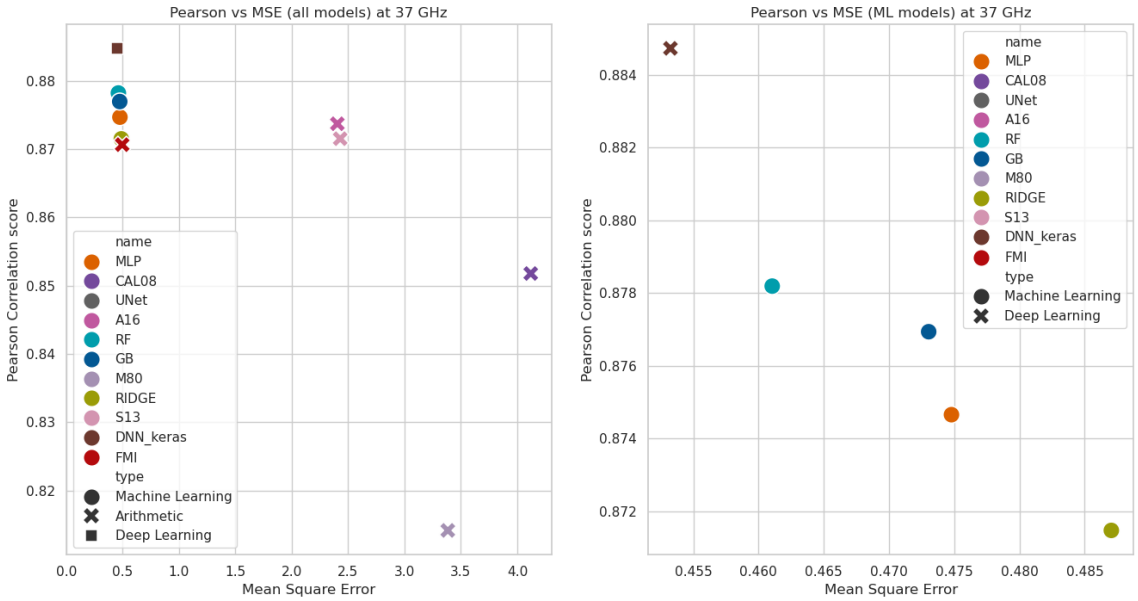


Figure 2.3 : Comparison of all models according to Pearson correlation score and Mean square Error (MSE) for whitecap fraction at 37 GHz

Every model has been trained on those sets and the following results are those obtained on the *test* set. FIgure 2.3 illustrates performances of every model that has been tested. The different markers distinguish the type of approach that has been used. Among Machine Learning models, there are ensemble models (Random Forests [RF] / Gradient Boosting [GB]) and regression models that include a regularization term (Ridge) as well as small neural network (MLP). It is clear from this figure that machine learning models outperform arithmetic models, although they seem to plateau. This stagnation may be due to our input data, given that none of the different architectures seem to gain any other advantages compared to the others.

To better visualize the performance of certain models, Figure 2.4 shows an evaluation against the *test* set of January-June 2017 for the M80, A16, DNN and FMI whitecap fraction models. It points out that the correlation of all the models are high, ranging from 0.81 (M80) to 0.87 (FMI, A16) and 0.88 (DNN) and that the DNN and the FMI models are the best in terms of RMSE. But this figure also highlights the biases of arithmetic models from the literature, which is close to two and explains the poor performance observed on Figure 2.3. Knowing that they have been trained on our dataset, the other models are properly unbiased. However, the FMI model shows a slight curvature, which allows us to assert that it tends to overestimate low values of WF and underestimate high values, whereas the DNN model seems to be well balanced with little dispersion around the identity line in red.

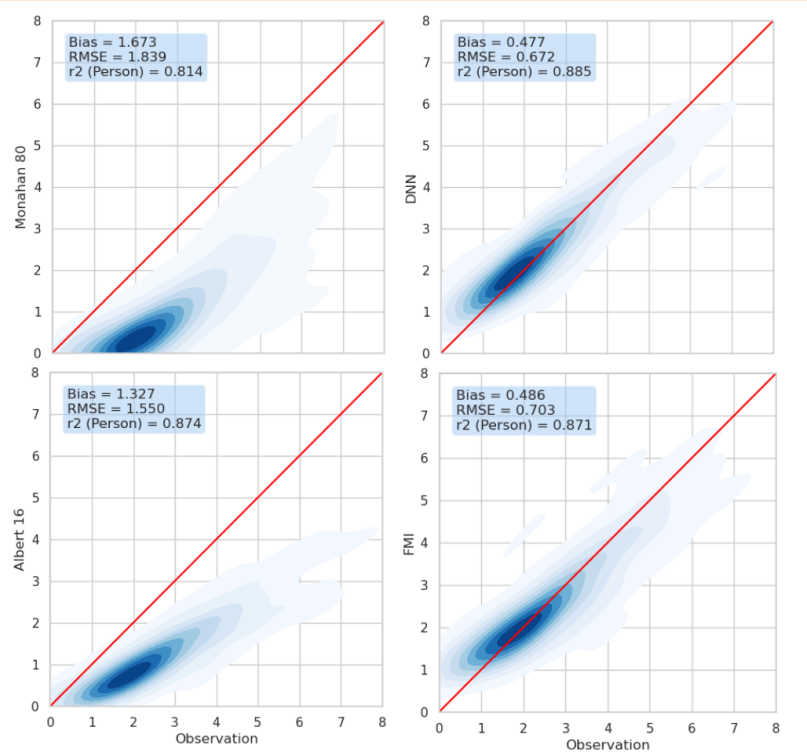


Figure 2.4 : Density scatterplots comparing observed to simulated whitecap fraction for four different models. The bias, RMSE and Pearson correlation coefficients are also shown.

Based on these evaluations, we can conclude that DNN is the most promising model. Also, comparing the A16 against the other classical models such as M80, CAL08 and S13 (the latter two are not shown) show that the A16 outperforms the others for the skill of the prediction of whitecap fraction. The newly implemented whitecap fraction model presented in Section 3 (FMI entry) shows very good skill scores, much better than A16.

Previous figures only highlight global performances of the models. However, we are also very interested in models correlation over time with observations. Figure 2.5 shows predicted whitecap fraction at 37 GHz for various models for January 2017 at a specific location as well as the values from our reference dataset. Firstly, all models struggle to retrieve major events with anti-correlation at certain times (cf. points between 17 and 21 January). Secondly, it is clear that all models exhibit very similar variations, which implies that they all depend strongly on wind speed.

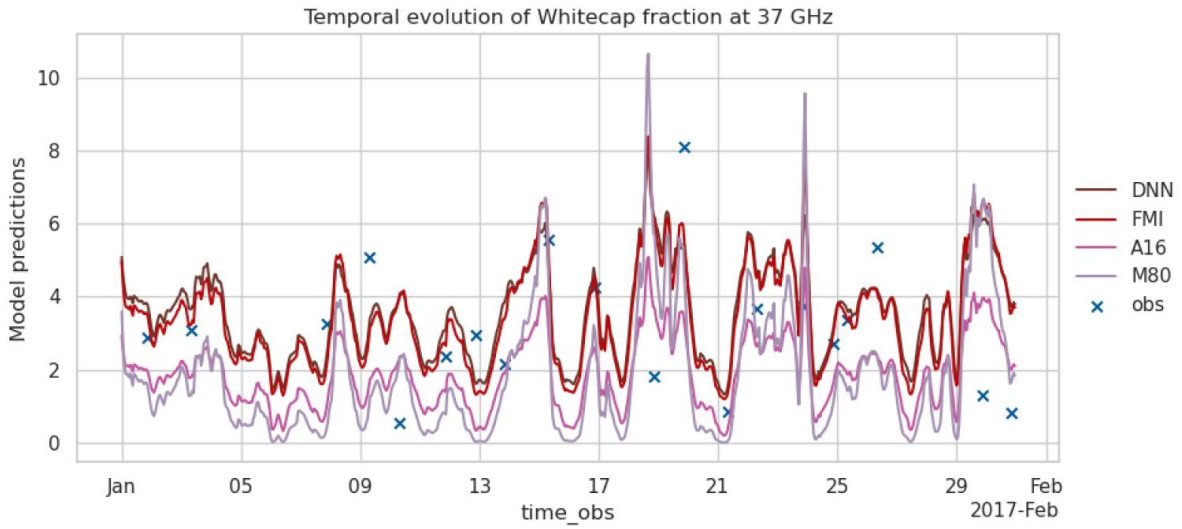


Figure 2.5 : Time series of simulated and observed Whitecap Fraction of several models over 1 month (January 2017) on a single pixel at 30°N, 50°W

Thus, our DNN model does not seem to take advantage of the other provided variables that are provided as input, except for minor variations, which explains the slight improvement in global performance. It is obviously difficult to draw real teaching from this figure, given that it was generated over only one month and for a single point but,, by looking at the variation deviations between our reference and wind speed, we can wonder if a piece of data is missing to predict whitecap fraction.

To try to find an answer to this problem, we have set up convolutional neural networks, which are based on abstracting patterns and spatial variations. Despite the advantage that this could have provided in terms of signal continuity and information extrapolation, the results for these models did not surpass those of DNN. We therefore have no real teachings on the interest of other inputs for predicting the output, except that the Significant Wave Height is highly correlated with wind speed and can replace wind speed without degrading performance.

4.3 Implementation in IFS-COMPO

The second part of our work is to integrate some of our models into IFS-COMPO. In the following, we present the method developed to use the models trained offline in the IFS-COMPO implementation, and show some results and evaluations of different methods implemented.

4.3.1 Integration of the machine learning model in IFS-COMPO

For models that use a pixel-to-pixel approach, we start with the DNN as it brings the best results. To incorporate it, we used the INFERO library which has been integrated into IFS-COMPO. To describe the operation, a backup model written as a ONNX file is provided to IFS-COMPO. The model data is first loaded and initialized with a Fortran interface. Then, in the routine computing the whitecap fraction, the arithmetic parametrization is replaced by a call to a function preprocessing the data and calling INFERO. The workflow is illustrated in Figure 2.6.

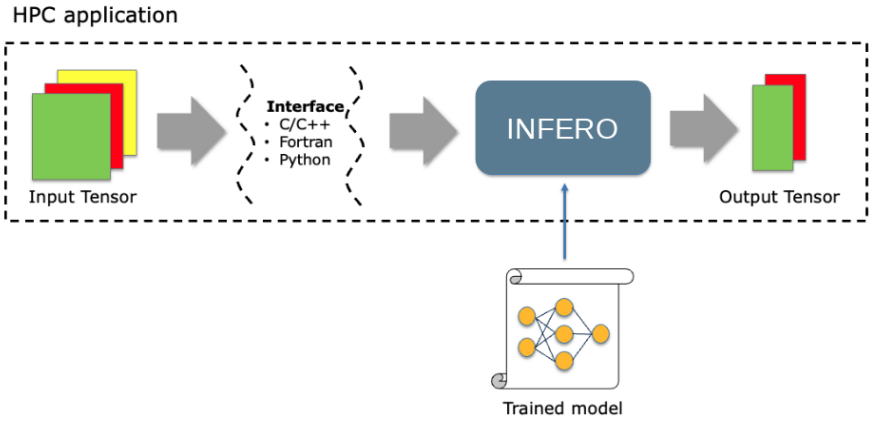


Figure 2.6 : Representation of the incorporation of our model into IFS-COMPO

By doing so, it is possible to run the DNN model designed offline directly inside IFS-COMPO. Moreover, knowing that every Machine Learning model evoked before could be saved into ONNX format, it is possible to run all pixel-to-pixel models with the strategy defined in Figure 2.6. In addition, the input tensor used by INFERO can be composed of data from several pixels, which means that the calls to the library can be adapted to the structure of the data distributed on each thread in IFS-COMPO. The input and output data from each pixel are independant.

In our case, the input tensor is composed of the different predictors chosen in the training of the machine learning model given in Table 2. Some of the parameters are available in the routine calculating the sea-salt aerosols, but the wave parameters are not known by this part of the code. It has thus been necessary to communicate the wave parameters to the aerosol model. However, the aerosol calculations are done before the wave model, which means that at the first time step, we do not have access to the information on the wave parameters. This means that we have to use one parameterization at the first step, for example A16. For the other time step, we can use the machine learning model.

The integration of machine learning in this environment using INFERO seems to be limited for spatial models. These models indeed need to use interactions and communications between neighbouring pixels, which is challenging in IFS-COMPO. The first issue is that grid points are not aligned or regularly located, as it is the case usually in the spatial machine learning model. In addition, the code structure is massively parallel, as the domain is first divided into several regions and allocated on the processors using distributed memory, then the pixels in each region are organized in batch, and several threads are used with sharing memory. It is difficult to retrieve the information of neighbouring pixels. Thus, the integration of this type of models seems compromised and was a key factor in our declining interest in these methods.

4.3.2 Results and evaluation

Different schemes for the whitecap fraction estimation have been implemented in IFS-COMPO using the INFERO library. In the following, we present the results for five cases : the neural networks estimation at 10 GHz and 37 GHz, the FMI scheme at 10 GHz and 37 GHz, and Albert 16 scheme. Five forecast only (ie without data assimilation) experiments have been launched for the full year 2017, at the usual TL511L137 (40km grid cell) resolution. The experiments are called A16 : this is the operational configuration, which serves as a reference, NN-10 and 37: neural network at 10 and 37 GHz, and FMI-10 and -37, the FMI scheme at 10 and 37 GHz.

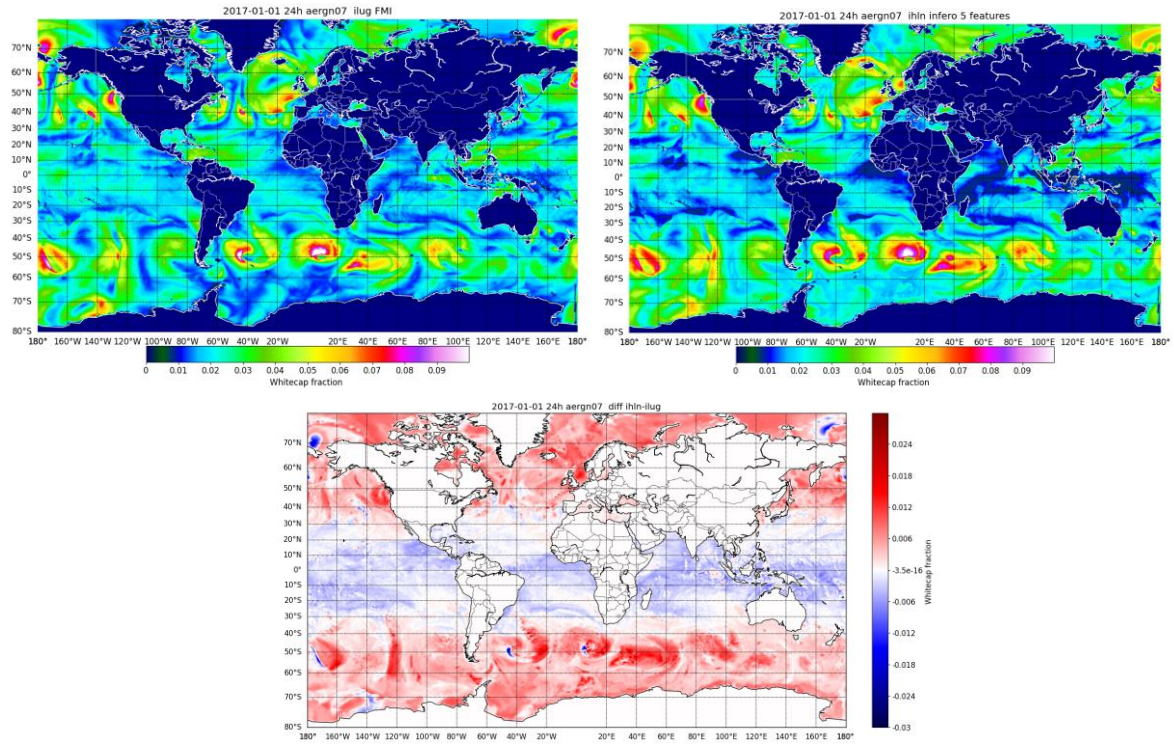


Figure 2.7: Simulated whitecap fraction by the FMI-37 (left) and DNN-37 (right) experiments and the absolute difference (bottom) for January 1st 2017

In Figures 2.7 and 2.8, we show snapshots of the simulated whitecap fraction for January 1st 2017 using 37 GHz (Figure 2.7) and 10 GHz (Figure 2.8) based algorithms. The whitecap simulated by the 37 GHz algorithms is much higher than that simulated with the 10 GHz algorithms, which reflects the differences between the remote sensing whitecap fraction at these two wavelengths: the whitecap fraction at 37 GHz is more or less twice as high as for the two schemes at 10 GHz and the A16 scheme. At 37 GHz, the FMI-37 and DNN-37 show quite similar results; DNN-37 simulates slightly higher whitecap fractions at high latitudes, and slightly lower values over tropical and equatorial areas. DNN-10, FMI-10 and the A16 scheme show very similar results. The very difference whitecap fraction values between the 10 and 37 GHz algorithms that it is necessary to adapt the computation of the sea-salt emissions from the whitecap fraction depending on the wavelength.. The scaling factor between the whitecap fraction and the sea-salt emissions has been divided by 2.25 for the 37 GHz algorithms in order to match as much as possible the evaluations for the sea-salt emissions.

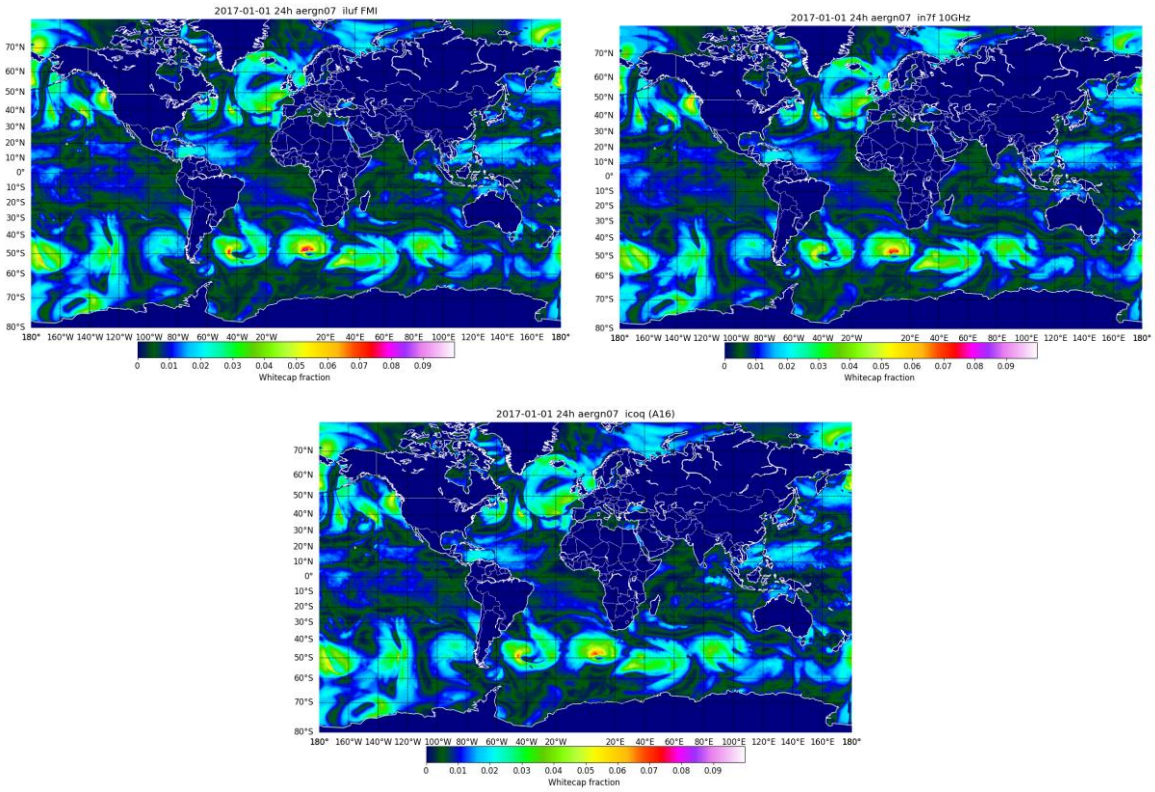


Figure 2.8: Simulated whitecap fraction by IFS-COMPO at 10 GHZ using the FMI scheme (left), our deep learning model (right) and A16 (bottom) for January 1st 2017

The evaluation of simulated AOD over a selection of AERONET sites impacted by sea-salt aerosols (AERONET stations Ragged Point,Reunion St Denis, Noumea, Midway Island, Key Biscayne,,Cape San Juan, Edinburgh, Cabo da Roca, ARM Graciosa, American Samoa, Amsterdam Island, Andenes, Birkenes) shown values that are exactly similar for all five experiments most of the year. At 500 nm, the five schemes are higher than the observation, except for a peak at the end of September. At 1020 nm, the simulations fit the curve of the observations. A way to discriminate the performances of the five schemes is to consider the correlations, given in Table 2.3. The schemes at 10 GHz and the A16 schemes have very close correlation scores, while the two schemes at 37 GHz have slightly smaller correlations. However, this can also be due to the computation from whitecap fraction to sea-salt emissions that could perhaps be improved.

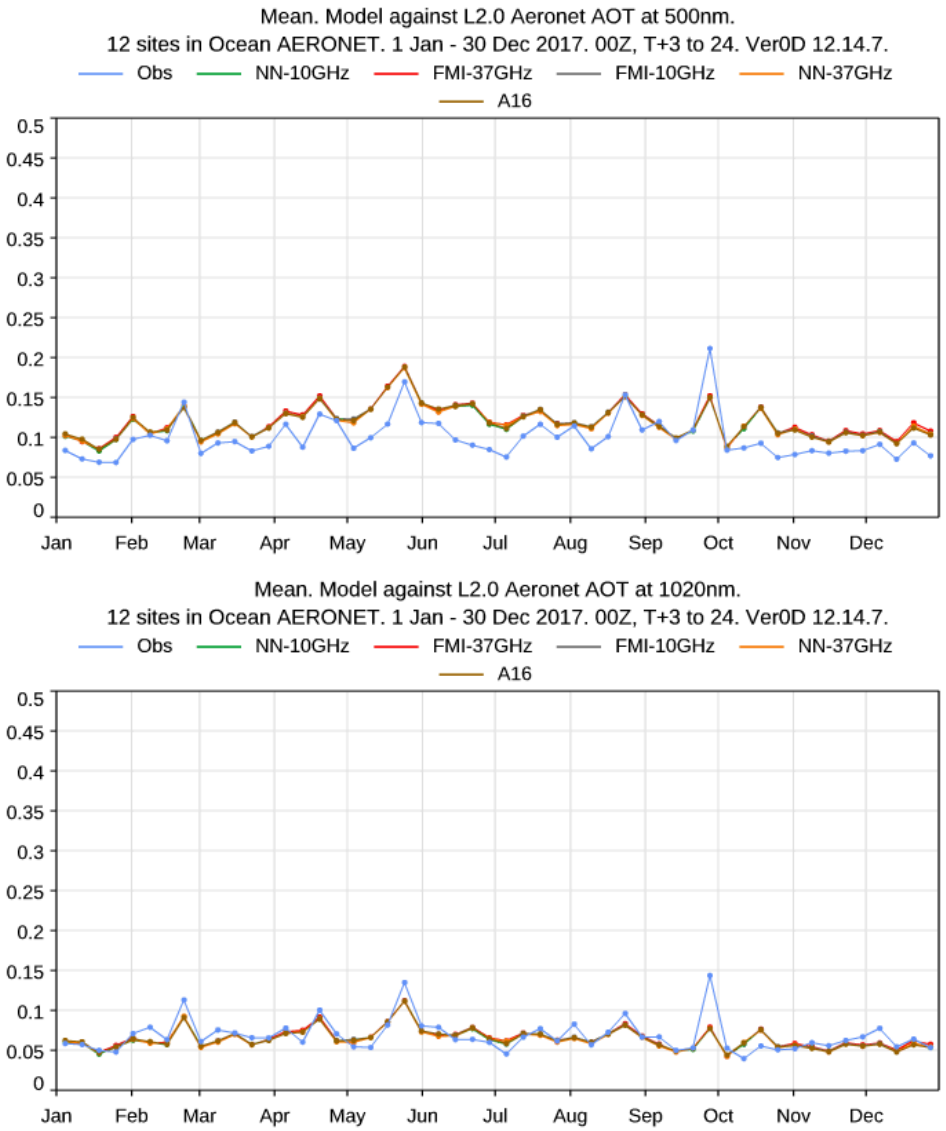


Figure 2.9: Evaluation of AOD at 500 and 1020nm for the full year 2017 compared to Aeronet stations selected to be representative of sea-salt aerosol impacted sites.

Table 2.3 : 2017: correlation between observed and simulated daily AOD at 500nm and 1020nm, European PM2.5 and PM10 in 2017 over all stations (background rural stations for PM10).

	A16	NN-10	FMI-10	NN-37	FMI-37
AOD 500	0.71	0.72	0.71	0.69	0.70
AOD 1020	0.76	0.77	0.77	0.74	0.76
PM 2.5	0.63	0.63	0.63	0.63	0.62
PM10	0.55	0.55	0.55	0.54	0.54

The PM evaluations also show very similar results for the five experiments. For PM2.5, the simulations are lower than the observation in wintertime, and closer to the observations for summertime. For PM10, the simulations are lower than the observations for the whole year. The correlations are very similar for the five schemes (see Table 3).

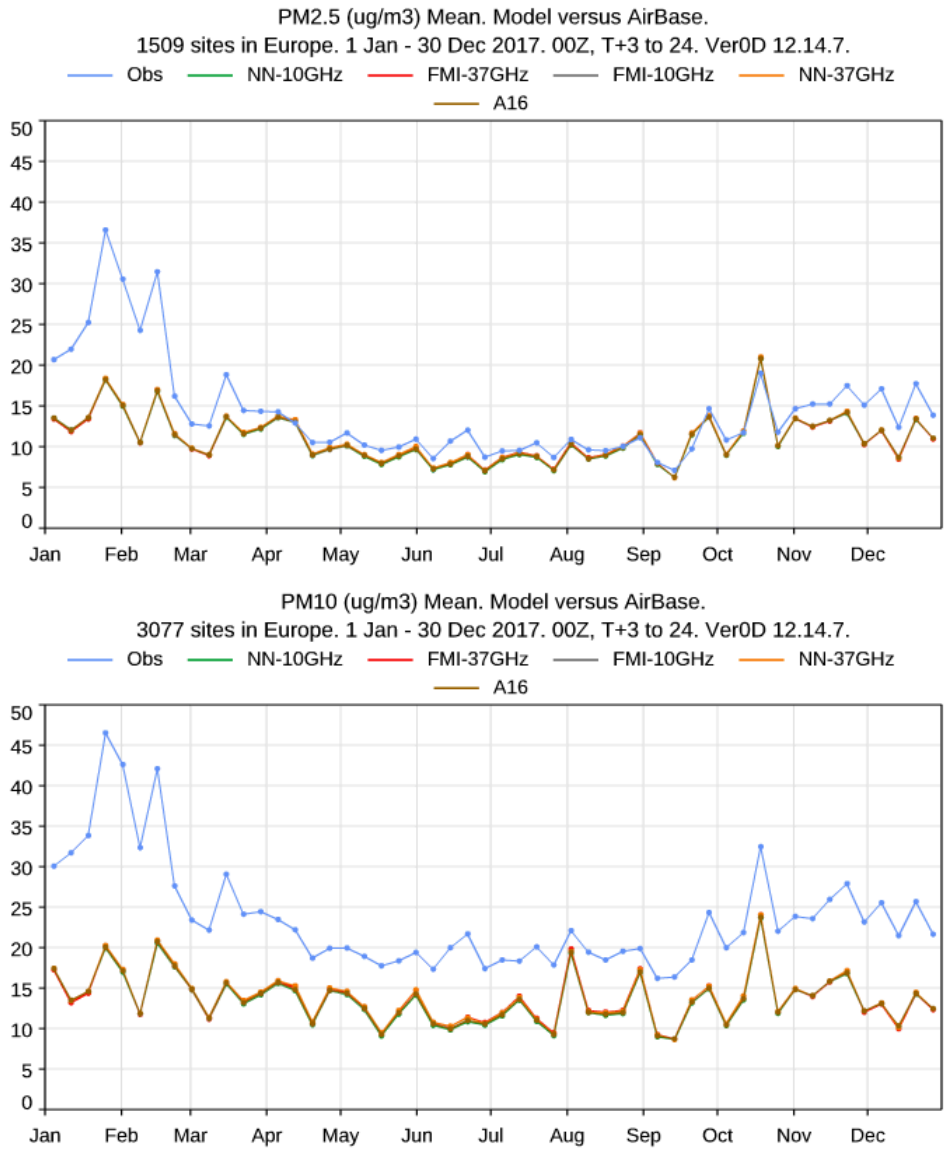


Figure 2.10 : Evaluation of PM2.5 and PM10 for the full year 2017 compared to Airbase

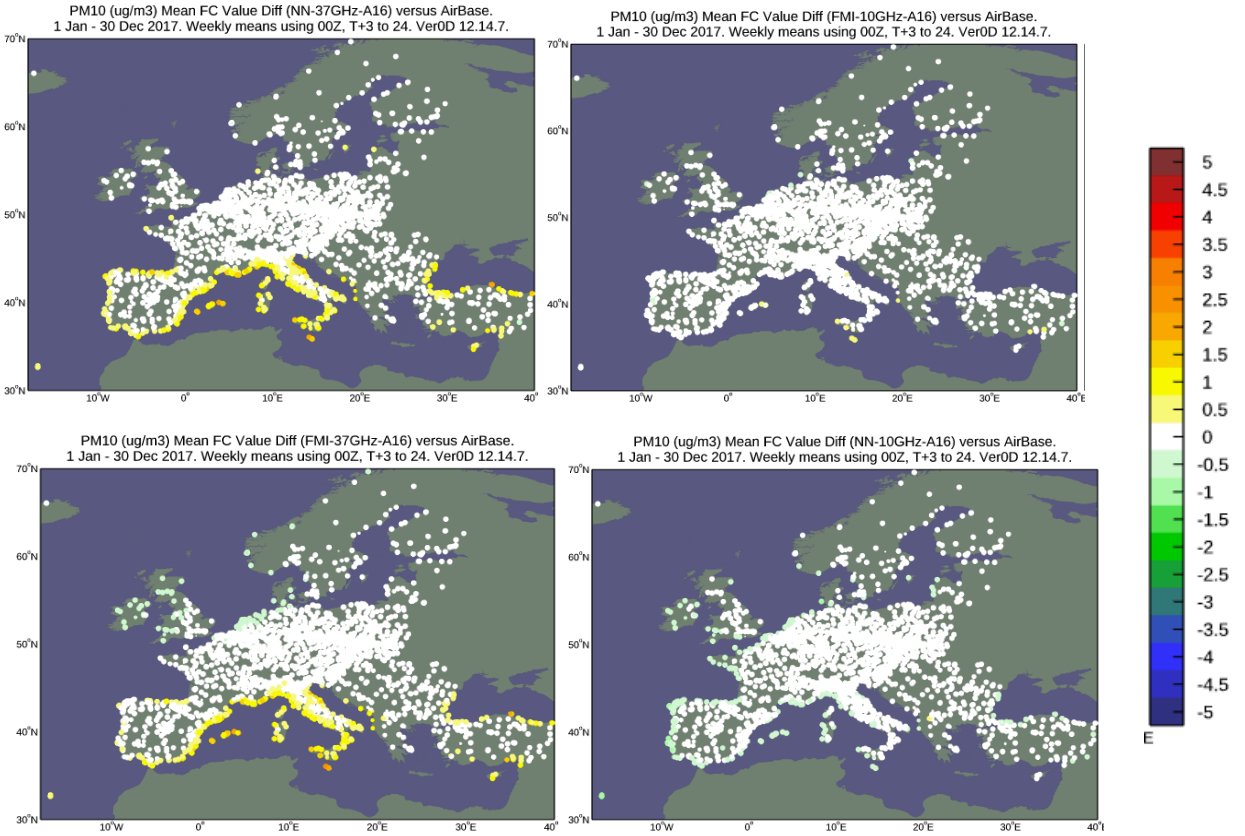


Figure 2.11 :2017 average; comparison of the absolute difference of the bias in simulated PM10 as compared to Airbase/EEA observations over Europe. Difference NN-36 minus A16 (top left), FMI-37 minus A16 (top right); FMI-37 minus A16 (bottom left) and NN-10 minus A16 (bottom right).Yellow - orange values indicate that the experiment is more biased than A16; green indicate the reverse.

In Figure 2.11, we compare A16 with the four other schemes for the simulation of PM10 in 2017 over all European stations to get a better understanding of the differences. For many stations that are far away from the coastline, the impact of the changes in sea-salt aerosol emission schemes are expectedly very small. The two 37 GHZ experiments (FMI-37 and NN-37) bring a small degradation of the bias along the coasts of Spain (along the Mediterranean only) and Italy, while NN-10 brings a small improvement along the Atlantic coast, from Portugal to France, Ireland and UK.

Next, we focus on the evaluations for Na and Cl at surface and their contribution to PM10 as provided by the EMEP/EBAS database over European stations. For the surface evaluations, the simulations are slightly higher than the observation for most of the year 2017, except for some peaks in March, October and December. In the first part of the year, the neural network at 37 GHz is higher than the other schemes, in wintertime, the FMI scheme at 37GHz is lower, and in summertime, the neural network scheme at 10 GHz is lower. However, the five simulations remain close to each other. For the PM10 evaluations, the simulations are significantly higher than the observations. For Na, the two schemes at 37 Ghz are higher while the neural network scheme at 10 GHz is the lowest. For Cl, the simulations are closer to each other. Regarding the correlations, we observe that the two schemes at 37 GHz have a lower score than the three other cases, this is specially the case for Cl PM10, see Table 4. The FGE and MNMB are slightly better with NN-10 as compared to A16.

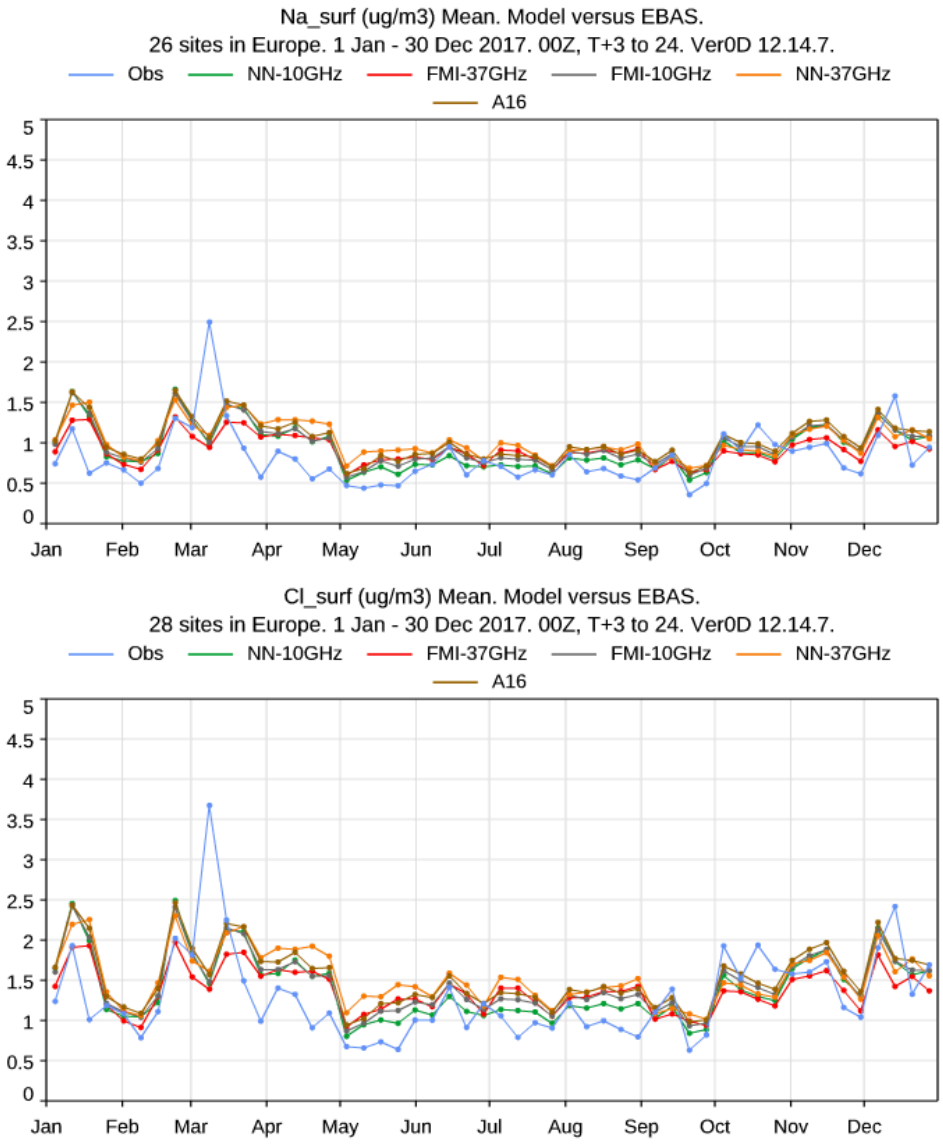


Figure 2.12 : Evaluation of daily Na and Cl at surface in µg/m3 for the full year 2017 compared to EBAS.

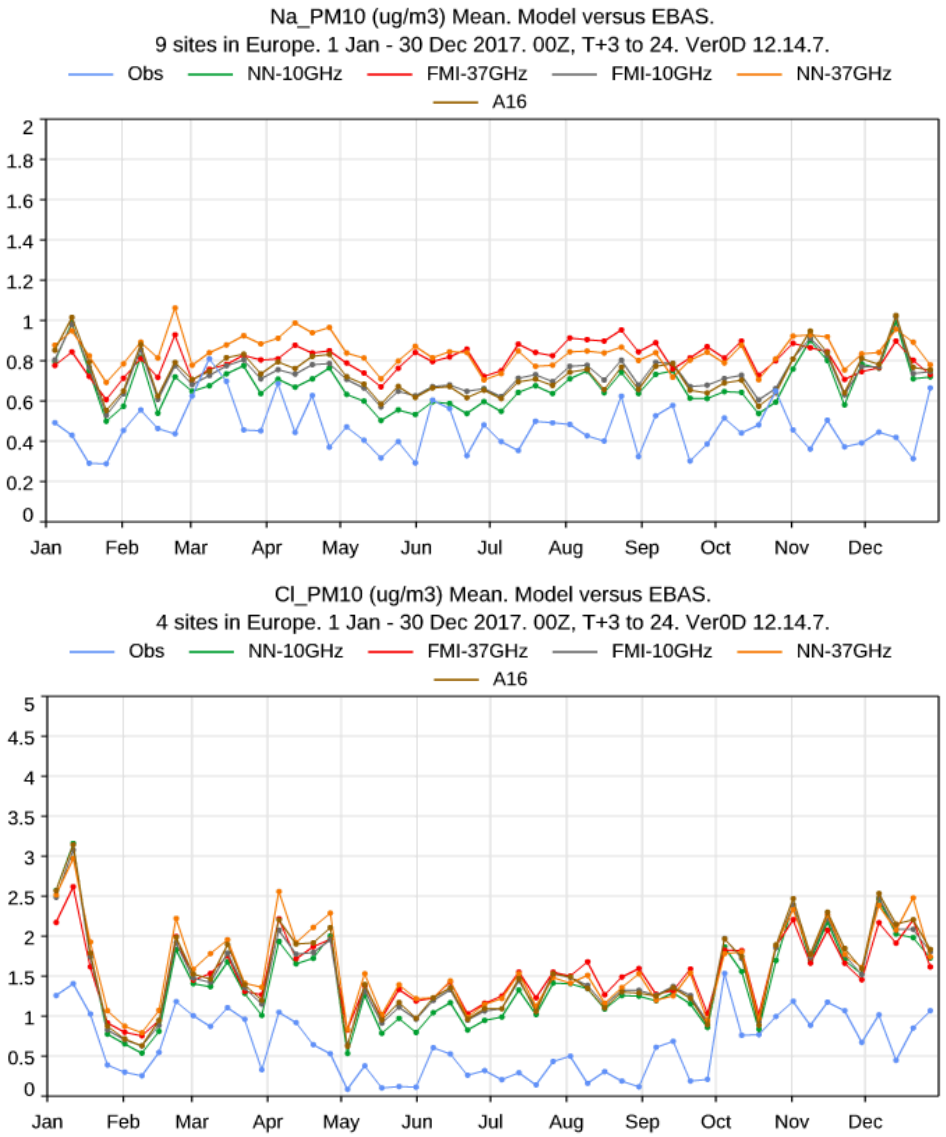


Figure 2.13 : Evaluation of daily Na and Cl as a fraction of PM10 in µg/m3 for the full year 2017 compared to EBAS

Table 2.4 : Correlations between observed and simulated daily Na/Cl in 2017 over EMEP/EBAS stations in Europe, surface and PM10.

	A16	NN-10	FMI-10	NN-37	FMI-37
Na surface	R = 0.55 FGE = 0.67 MNMB = 0.36	R = 0.55 FGE = 0.65 MNMB = 0.29	R = 0.55 FGE = 0.65 MNMB = 0.33	R = 0.51 FGE = 0.69 MNMB = 0.37	R = 0.52 FGE = 0.66 MNMB = 0.29
Cl surface	R = 0.58 FGE = 0.84 MNMB = 0.49	R = 0.58 FGE = 0.82 MNMB = 0.42	R = 0.58 FGE = 0.83 MNMB = 0.45	R = 0.54 FGE = 0.86 MNMB = 0.48	R = 0.55 FGE = 0.83 MNMB = 0.40
Na PM10	R = 0.71 FGE = 0.66 MNMB = 0.38	R = 0.70 FGE = 0.65 MNMB = 0.70	R = 0.70 FGE = 0.66 MNMB = 0.38	R = 0.67 FGE = 0.71 MNMB = 0.44	R = 0.64 FGE = 0.71 MNMB = 0.41
Cl PM10	R = 0.66 FGE = 1.04 MNMB = 0.92	R = 0.66 FGE = 1.00 MNMB = 0.87	R = 0.64 FGE = 1.02 MNMB = 0.90	R = 0.54 FGE = 1.03 MNMB = 0.91	R = 0.47 FGE = 1.01 MNMB = 0.85

To conclude, we have obtained convincing results for the offline estimation of the whitecap fraction using neural networks. The implementation of the inference models in IFS-COMPO shows that the models have good performances, similar to analytical models, however, we do not see a great improvement. This can be due to the fact that the data used for the training can be noisy, and that we do not have access to a lot of data with high values. In addition, the neural network models might be improved by using other predictors.

5 Gridded NORTRIP (GNORTRIP)

Aim: To provide improved temporal and spatial distribution of PM_{10} non-exhaust emissions in CAMS forecasts.

Background: Emissions of road dust are a significant contributor to PM_{10} concentrations in countries where studded tires are in use. In countries without studded tires non-exhaust PM emission contributions are similar or more than exhaust emissions. Unlike exhaust, non-exhaust emissions are controlled by road surface conditions. Road dust will not be emitted under wet conditions and may build up on the surface during wet winter or frozen conditions. Road dust may then be released when the roads dry up, enhancing emissions in, for example, spring time. This process is not captured by constant non-exhaust emission factors that are currently used in most chemical transport models. This shortcoming may contribute to the underestimation of PM_{10} during road dust emission events seen in the CAMS forecasts.

5.1 Model and input data

5.1.1 Modelling concept

The **NO**n-exhaust **R**oad **T**raffic **I**nduced **P**article emission model (NORTRIP, Denby et al. 2013a,b) calculates emissions, road dust loading and surface conditions for roads. The original model simulates individual road segments, called road links. Information (metadata) from each individual road link (Position, ADT, number of lanes, speed, etc.) and meteorological data are used to calculate surface conditions and non-exhaust emissions. Terrain data surrounding these road links are used for shading and long wave radiation calculations.

In the gridded version of the model the same routines are used for calculating non-exhaust emissions but instead of modelling individual roads the road links are aggregated into 5 different classes, based on Open Street Map (OSM) categories, each having specific metadata values. The total road length of each class, within a grid, is then used for calculating total emissions from that class. Gridded traffic volume data (veh.km) are used and distributed to the different classes based on a weighted distribution of the classes.

The following major differences between the normal NORTRIP and the gridded GNORTRIP calculations are then:

- No terrain data is used in the gridded calculations
- Meteorology is specified per grid, rather than at the individual positions of each road link
- Variability in traffic volume and speeds between individual roads are removed and aggregated ‘average’ values are implemented
- The gridded model makes use of gridded traffic volume data (veh.km), provided by TNO for all of Europe, rather than calculations for individual roads
- No sanding is applied in GNORTRIP. Road salting is applied in both versions.

5.1.2 Input data

In addition to meteorological data (IFS), NORTRIP needs traffic data, road metadata, and winter maintenance information to calculate emissions.

5.1.2.1 Traffic data

The gridded traffic data used is provided by TNO (mill veh.km for different vehicle types. Based on 2016 data from Open Transport Map, Kuenen et al., 2022). To estimate traffic volume per gridcell, the road network within a cell is represented by 7 OSM road types (OSM types: trunk, motorway, primary, secondary, tertiary, residential, unclassified). We reduce these to 5 different road types for each grid since unknown/residential and trunk/highway have the same weighting and metadata values. The OSM weighting is presented in Figure 3.1.

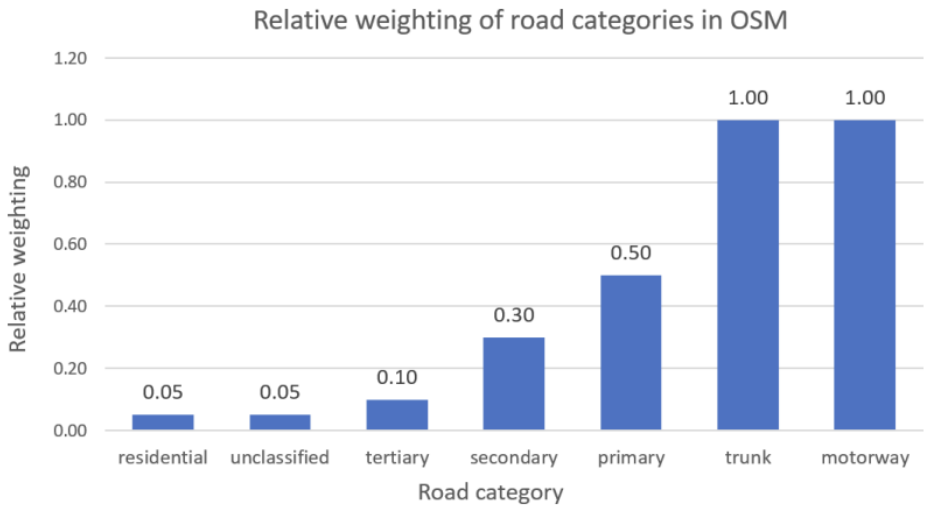


Figure 3.1. Weighting used in the redistribution of traffic volume data to OSM road categories. Taken from Mu et al. (2022).

We calculate an ADT for each of the road types based on the length of the road type in OSM, its weighting and the annual vehicle km from the TNO traffic grid. i.e. ADT = veh_km/OSM_km/365. The traffic data is based on 2016 numbers, and may have changed significantly in 2023, which is the year for which GNORTRIP is run. Therefore the traffic volume has been normalized with reported national totals from 2022 when these have been found available, see Table 3.1 and Figure 3.2. Many countries have a scaling close to 1 but some countries, such as Spain, significantly overpredict the traffic volume in the TNO gridded dataset.

In the TNO dataset the following vehicle types were included: mopeds, motorcycles, passenger cars, buses, heavy duty trucks and light commercial vehicles. In regard to non-exhaust emissions these different types were further aggregated into ‘Light vehicles’ and ‘Heavy vehicles’. Mopeds and motorcycles were counted as 1/5 ‘th of a light vehicle and aggregated together with passenger vehicles and light commercial vehicles into the ‘Light vehicle category’. Buses and heavy duty tracks were aggregated into the ‘Heavy vehicle’ category.

Table 3.1. Comparison of TNO 2016 and national reported 2022 total vehicle kilometers. The major source of traffic data is Eurostat, but for some countries other sources were used. In total 20 out of 51 countries were normalised.

	National totals 2022	TNO 2016	Scaling	Comments
Austria	92451	100114	0,923	
Croatia	24890	45109	0,552	
Czechia	58664	106536	0,551	
Denmark	50000	56013	0,893	Uncertain source
Estonia	11696	13517	0,865	
Finland	48856	77932	0,627	
France	628145	887074	0,708	Reference
Germany	650100	964739	0,674	PC's *110%, Reference
Hungary	48677	90184	0,540	
Ireland	47604	43257	1,101	
Lithuania	13114	31938	0,411	
Netherlands	134061	122223	1,097	
North Macedonia	5454	13017	0,419	
Norway	46907	49174	0,954	
Romania	64051	128534	0,498	
Slovenia	20511	19389	1,058	
Spain	254459	767464	0,332	
Sweden	76116	72460	1,050	
Switzerland	58800	115596	0,509	
United Kingdom	521093	634544	0,821	

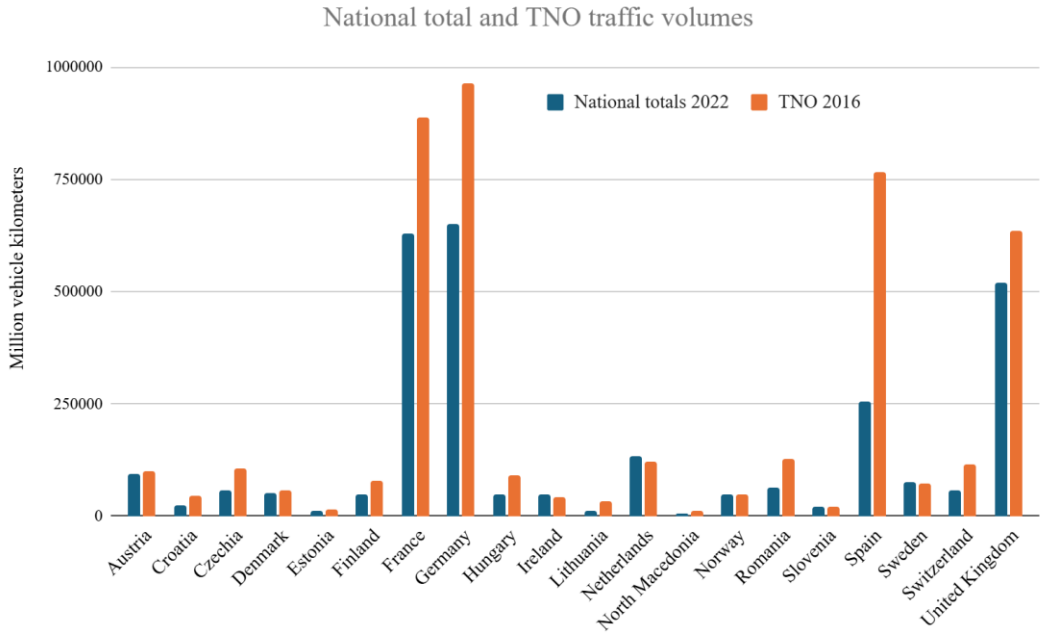


Figure 3.2. Visual representation of Table 3.1 showing the national traffic volume from Eurostat and from the TNO gridded traffic volume data.

5.1.2.2 Studded tyre and winter road maintenance data

In some countries the use of studded tyres on vehicles in winter contributes to enhanced road wear. GNORTRIP takes this into account by specifying the fraction of light and heavy vehicles that use studded tyres and applying appropriate road wear factors for these. For Norway, Sweden and Finland studded tyre shares are applied per region. In Norway this is per municipality (356) but in Sweden and Finland this is per region (~20 regions each). The other countries using studded tyres, Table 3.2, use a single value for the whole country. The start and stop of the studded tyre season is based on Norwegian data. Salting is applied under particular meteorological conditions and to roads classified as 'saltable'. Saltable roads are based on Norwegian ADT's and salting regime. Sanding is not used in these model calculations.

Table 3.2. Studded tyre share (%) for light and heavy duty vehicles during the studded tyre season. Various sources have been used where mostly web based searches for information have been used. Only Norway, Sweden and Finland, marked with $^{+}$ have directly provided studded tyre shares on a regional basis. * indicates that no data was available, so estimated.

Country	Light vehicles	Heavy vehicles
Austria	5*	1
Belarus	25*	8
Switzerland	5*	1
Estonia	62	20
Finland	80^{+}	26
Iceland	30	10
Lithuania	8	2
Latvia	8	2
Norway	45^{+}	15
Russian Federation	70	23
Sweden	60^{+}	20

5.1.2.3 Road Metadata

The resulting 5 aggregated road types were assigned particular metadata values, as shown in Table 3.3.

Table 3.3. Metadata parameters assigned to the 5 OSM aggregated road types.

Country	Number of lanes	Signed speed (km/h)
Motorway/Trunk	6	100
Primary	4	80
Secondary	2	60
Tertiary	2	50
Residential/Unclassified	2	30

5.1.2.4 Emission factors for non-studded tyre contributions

Wear factors for road (non-studded), tyre and brake currently used in NORTRIP tend to be higher than those reported in the EMEP/EEA emission handbook (EMEP/EEA, 2023). To try to be consistent with non-exhaust emission factors being used across Europe we have implemented the EMEP/EEA wear rates, but not size distributions, in the gridded European calculations of NORTRIP.

In addition to non-exhaust emission factors, emission factors for NO_X and exhaust particles (EP) are also included in the calculations for reference, Table 3.4. These are based on average Norwegian emission factors and applied to every country.

Table 3.4. Emission factors applied in the modelling for NO_X and exhaust particles (EP). Units are g/km.

EP Light vehicles	**EP Heavy vehicles**	**NOX Light vehicles**	**NOX Heavy vehicles**
0,0103	0,047	0,502	2,11

5.2 Results

5.2.1 Model output

Calculated emissions, on the CAMS 0.1º x 0.05º grid, are provided in netcdf format for each day of the year at 1 hour temporal resolution. The fields provided are listed in Table 3.5 and are available at:

https://thredds.met.no/thredds/catalog/data/fou-kl/uEMEP/CAMAERA/EU_emissions_v2/catalog.html

Table 3.5. Variable fields provided in the GNORTRIP netcdf files.

	PM_{co}	$PM_{2.5}$	PM_{10}	NO_x
Road wear	X	X		
Tyre wear	X	X		
Brake wear	X	X		
Sand	X	X		
Salt	X	X		
Exhaust		X		X
Total	X	X	X	

5.2.2 Assessment in Norway

Before applying the model to all of Europe a number of test calculations were made in Norway where a complete national road network is available (NVDB) and where higher resolution meteorology is also available (MEPS/AROME 2.5 km, Müller et al., 2017). In these tests we applied the following calculations:

1. Individual road calculations using national traffic network data and 2.5 km MEPS meteorology (NVDB-NORTRIP-MEPS), aggregating road link emissions to grid afterwards
2. The national traffic road network is first gridded and MEPS meteorology applied to the gridded data (NVDB-GNORTRIP-MEPS)
3. Gridded TNO traffic in Norway using MEPS meteorology (TNO-GNORTRIP-MEPS)
4. Gridded TNO traffic in Norway using IFS meteorology (TNO-GNORTRIP-IFS)

A comparison of calculations 1 and 2 will show how the gridding will affect the calculations, compared to individual roads, i.e. a test of the methodology. A comparison of calculations 2 and 3 will show the difference between the national Norwegian and TNO datasets for traffic, i.e. a test of the traffic data input.

In Figure 3.3 we compare calculations 1 and 2. For annual total emissions there is a very good correspondence between the individual road calculation and the gridded calculation. There is also a good correspondence for daily totals, but there is more scatter, indicating the influence of the meteorology and/or shading on the road surface conditions on a daily basis.

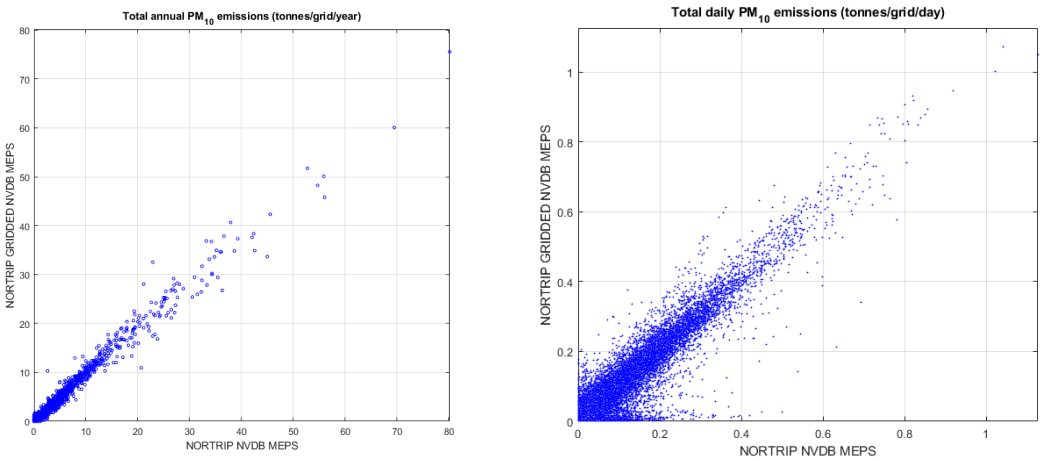


Figure 3.3. Comparison of a gridded emission calculation (NORTRIP GRIDDED NVDB MEPS) with the individual road calculation (NORTRIP NVDB MEPS) in Norway, both using MEPS meteorology. Left is the total annual emissions of PM_{10} in tonnes/grid/year and right the daily total emissions of PM_{10} tonnes/grid/day.

In Figure 3.4 we compare calculations 2 and 3. This will indicate the difference between the TNO traffic dataset and the NVDB Norwegian national dataset. These are significantly different, indicating there can be poor correspondence between the TNO dataset and the Norwegian national dataset. In Figure 3.5 we visualise the difference between the national and TNO traffic data emission results. There are significant discrepancies between the two,

particularly in the region around Hamar, around 1º North of Oslo. Daily total national emissions for Norway are also shown in Figure 3.6 for calculations 1, 2 and 3. These indicate that all applications of the model produce very similar temporal variability, on a national scale.

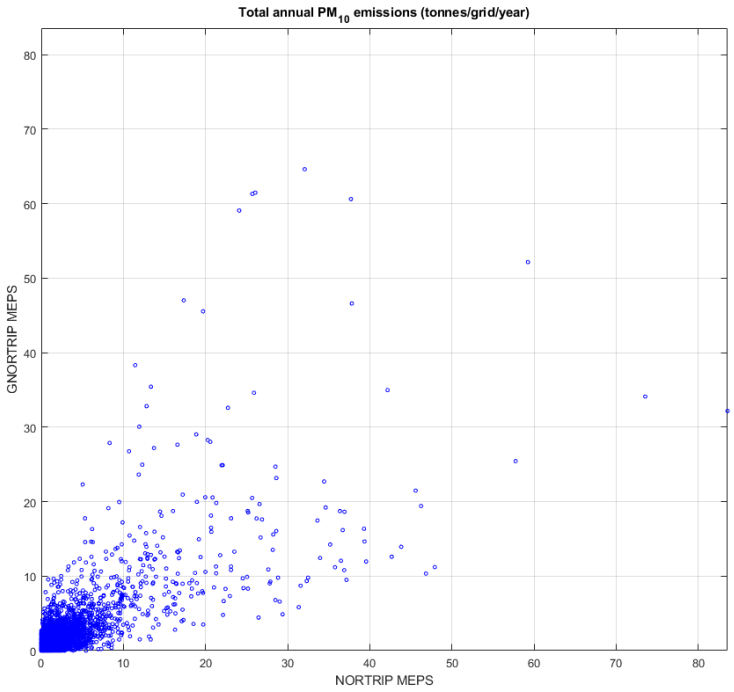


Figure 3.4. Comparison of a gridded emission calculation with national traffic data (NORTRIP MEPS) with the gridded TNO (GNORTRIP MEPS) in Norway using MEPS meteorology.

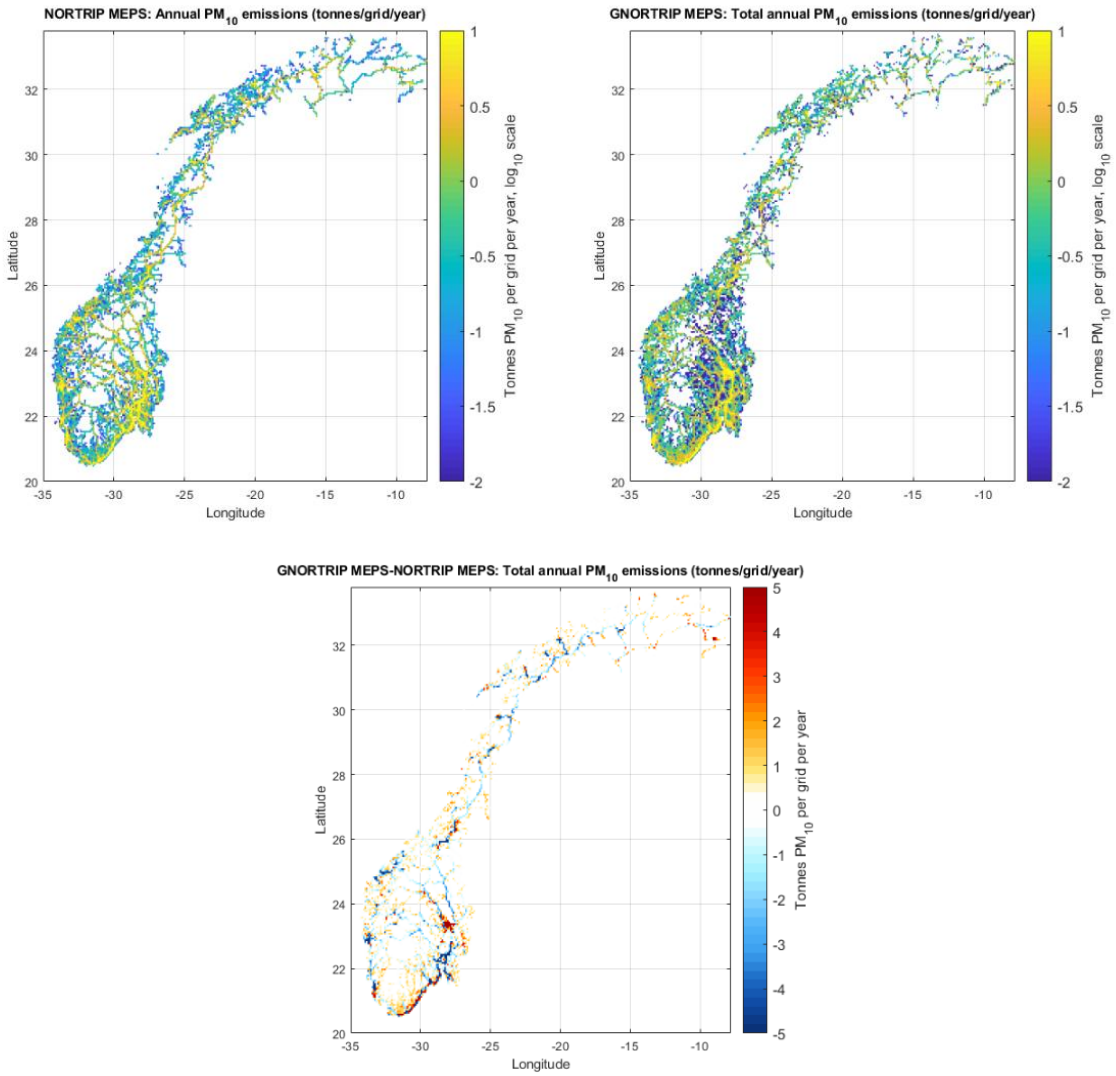


Figure 3.5. Top left, annual total PM_{10} non-exhaust emissions using NVDB traffic data. Top right, annual total PM_{10} non-exhaust emissions using TNO traffic data. Bottom, difference between these two. MEPS meteorology is used for both calculations. Note colour scales are log_{10} in the two top plots.

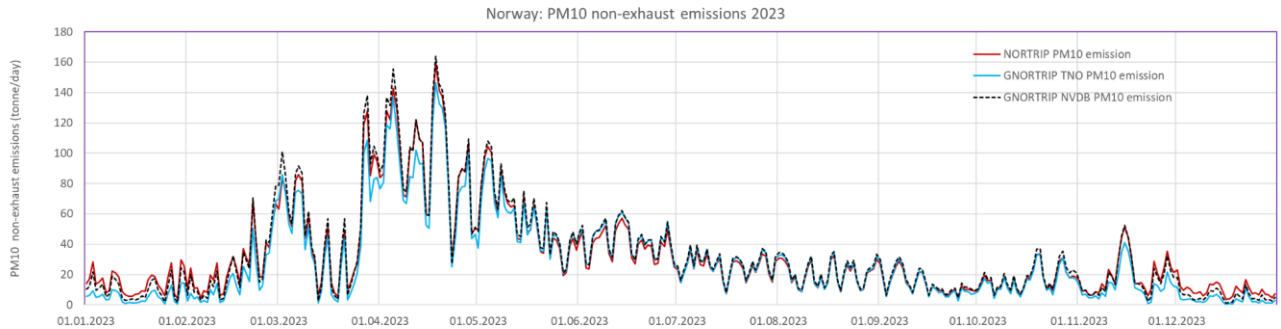
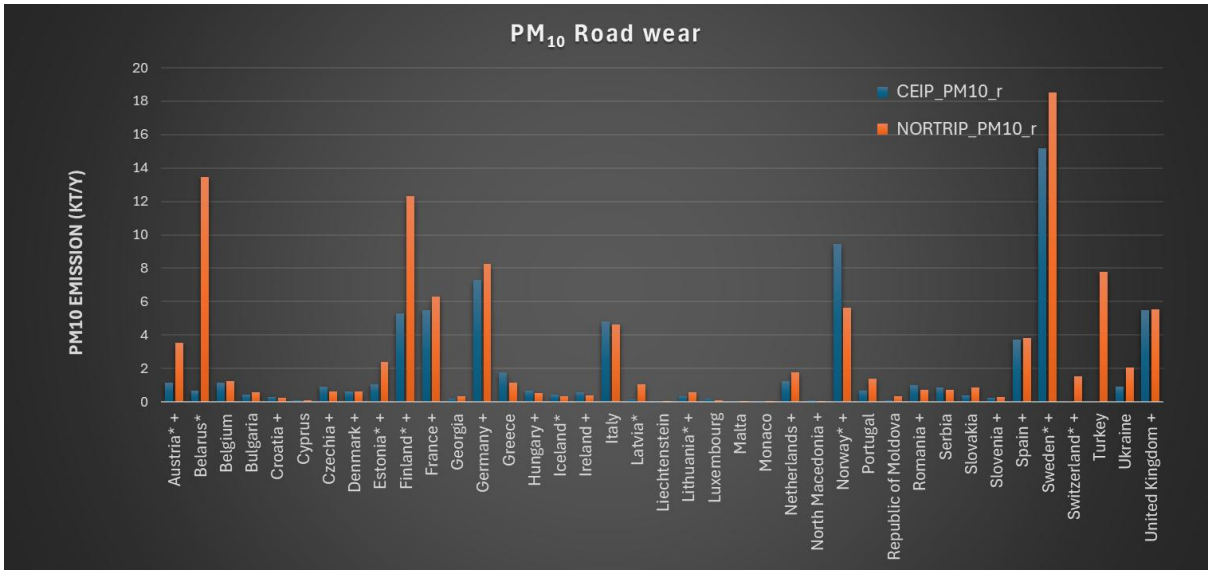


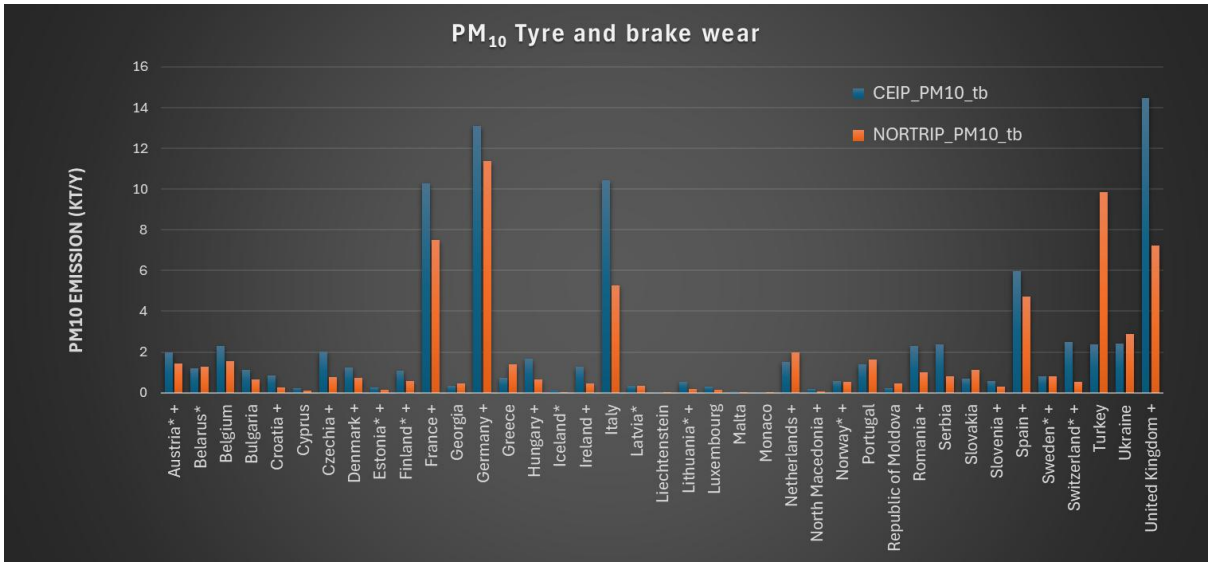
Figure 3.6. National total daily PM_{10} non-exhaust emissions for 2023 in Norway. Shown are the calculations 1 (NORTRIP), 2 (GNORTRIP NVDB) and 3 (GNORTRIP TNO).

5.2.3 Application in Europe

For this work, hourly emissions of $PM_{2.5}$ and PM_{coarse} (as well as NO_x and EP exhaust for comparison) for 2023 are calculated. In Figure 3.7 the total emissions per country (kt/yr) calculated by GNORTRIP for 2023 are compared to reported CEIP emissions from 2022. Shown are PM_{10} road wear, PM_{10} tyre and brake wear, NO_X and EP.

For road wear many countries give similar emissions, though not all countries report this emission source. Significant differences can be seen, for example in Finland. The largest emitter is Sweden, since this country has the highest traffic volume of all the countries using studded tyres. The similarities between the CEIP and NORTRIP data for countries not using studded tyres probably reflects the common use of non-exhaust emission factors from the EMEP/EEA handbook.





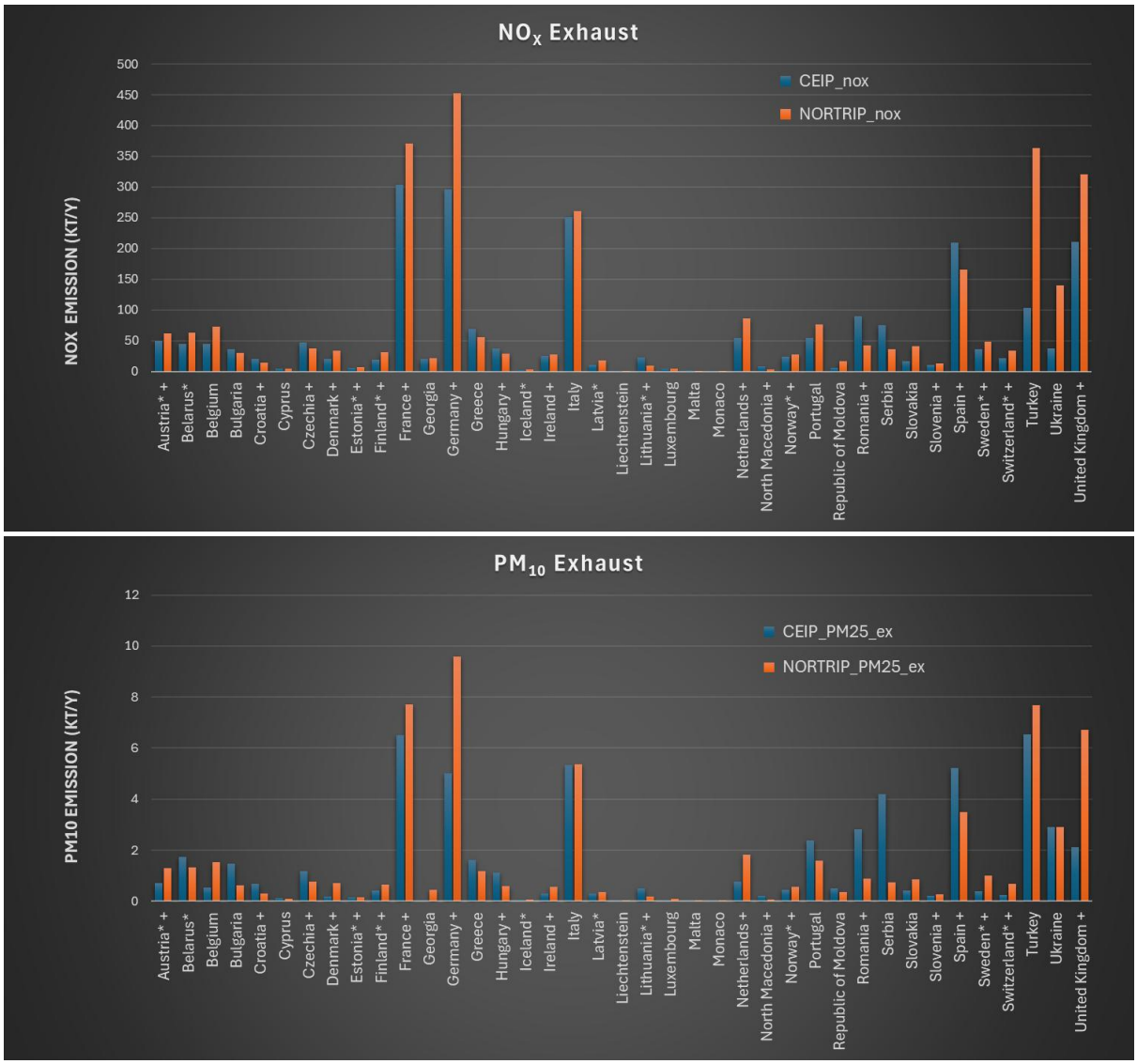


Figure 3.7 Country total emissions for PM_{10} road wear, PM_{10} tyre and brake wear, NO_X and EP. Comparison of CEIP reported emission data 2022 and the GNORTRIP calculations. Countries where studded tyres have been applied are labelled with '*' and countries where the TNO traffic data has been normalised with national totals for 2022 are shown with a '+'.

5.3 Application of GNORTRIP in EMEP

5.3.1 Method

To evaluate how the new GNORTRIP emissions affect the temporal and spatial correlation of modeled and observed PM_{10} concentration, experiments have been performed using EMEP. Two experiments have been run, one using GNORTRIP for GNFR sector F4 emissions (road transport non-exhaust), and the other using sector F4 emissions from CAMS-REG-AP_v7 yr2022 (as used in the CAMS regional validated reanalysis for 2022). Emissions from all other sectors are from CAMS-REG-AP_v7, yr2022 for both experiments. The simulation setups are summarized in Table 3.6. The simulations otherwise follow the same setup as the latest EMEP Status Report (1/2024). The results are evaluated by comparing with observations from all available EEA stations, including traffic stations, using the Aeroval evaluation tool.

Table 3.6. EMEP simulation configuration.

	CAMS	GNORTRIP
Meteorology	ECMWF-IFS Cycle 48r1, 2023	
Resolution	0.1x0.1 degrees	
EMEP version	v5.5	
Sector F4 emissions	CAMS-REG-AP_v7, 2022	GNORTRIP
All other sectors	CAMS-REG-AP_v7, 2022	

5.3.1.1 Comparison of CAMS, CEIP and GNORTRIP non-exhaust emissions

Though national total emissions from CEIP and from CAMS should be the same in most cases, a comparison was made of the CAMS total non-exhaust PM_{10} emissions used in the EMEP simulations with the GNORTRIP and CEIP emissions in Figure 3.8, as in Figure 3.7a and 3.7b. The CAMS emissions are generally similar to CEIP, as expected, with a few exceptions. For Norway, the CEIP emissions have been updated recently to use NORTRIP generated values, which increased the total CEIP emissions markedly. It is unclear why there are large differences between CAMS and CEIP for e.g. Turkey and the United Kingdom.

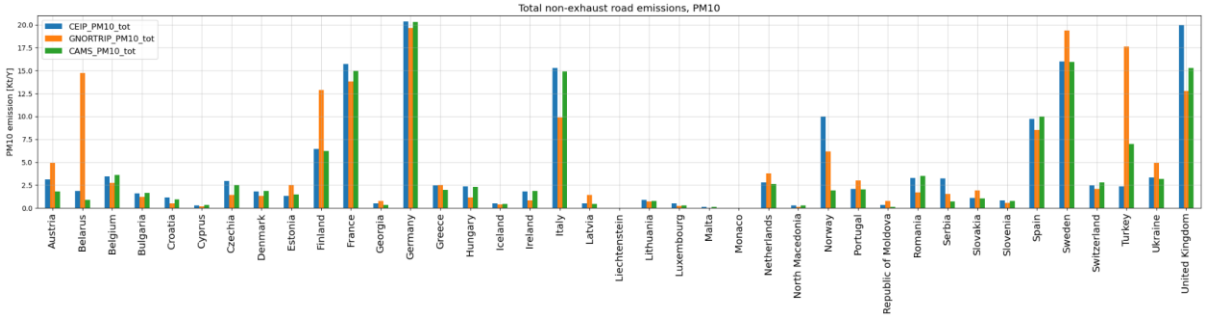


Figure 3.8. As in figure 3.7 but also showing the CAMS non-exhaust PM_{10} emissions.

5.3.2 EMEP results

In this section we present some of the results from the EMEP calculations using CAMS and GNORTRIP non-exhaust emissions. Comparisons to observations are made that include traffic stations from the EEA database. Since EMEP has a spatial resolution of 0.1º x 0.1º we do not expect EMEP to be able to reproduce the absolute concentrations observed at these stations. However, important information on temporal variation is contained in these traffic station data, so these are included.

5.3.2.1 Correlation

For regions that experience winter conditions and use studded tyres, mainly the Nordic and Baltic countries, the improvement in temporal correlation is notable when using GNORTRIP emissions (Figure 3.9a). For the countries that don't experience winter conditions or use studded tyres, there is generally no difference in correlation between the CAMS and the GNORTRIP simulations.

There is limited or no improvement in spatial correlation within countries. As the model resolution is relatively coarse (0.1º x 0.1º), it does not capture local concentration gradients between the roads and the measurement stations that often cause the spatial variability.

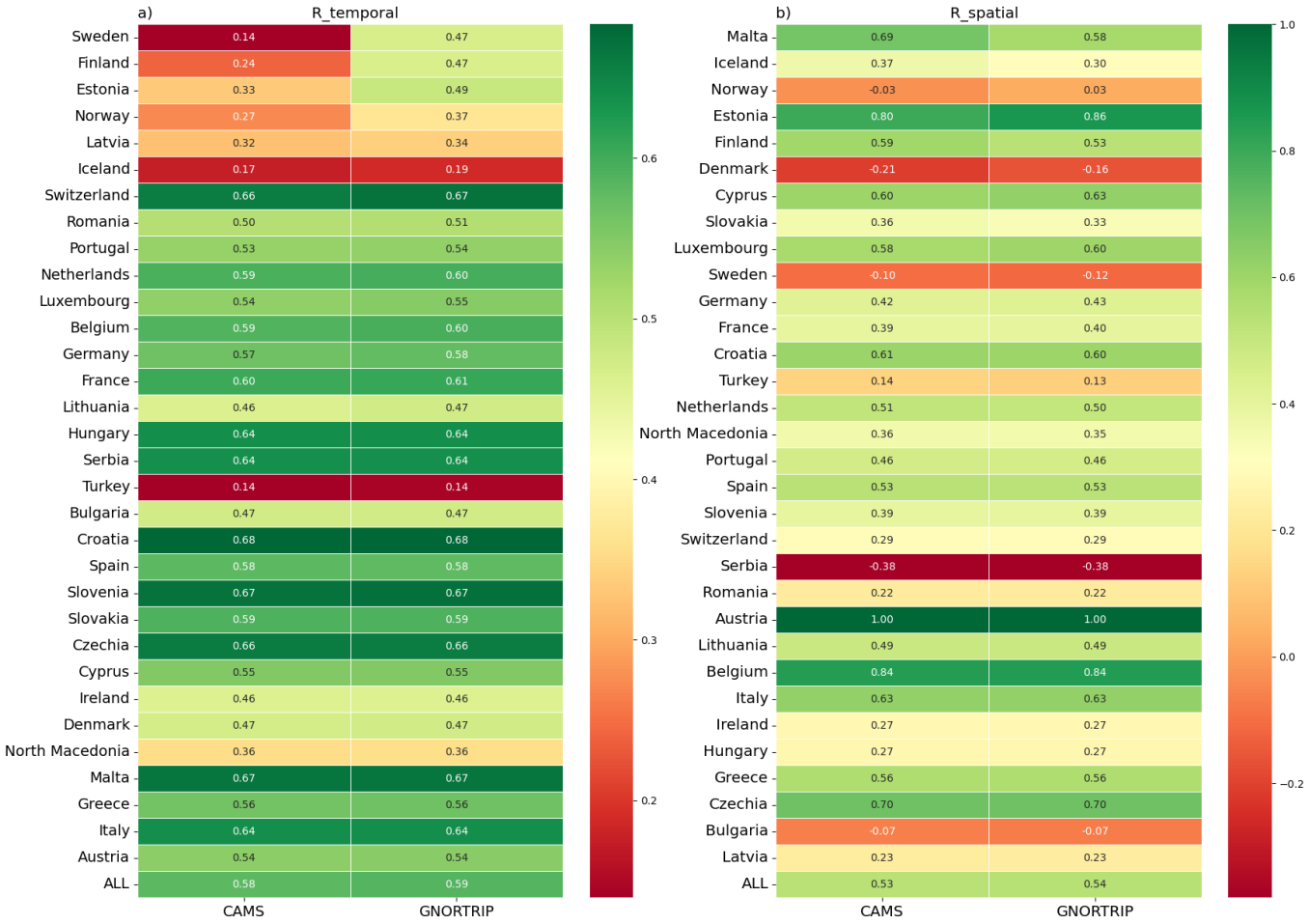


Figure 3.9. Temporal (a) and spatial (b) correlation between modeled and observed (all stations from EAA) PM_{10} concentrations (2023) for European countries, separately and together ("ALL", bottom). The countries are sorted by size of difference between the CAMS and GNORTRIP values, with the largest differences at the top.

5.3.2.2 Seasonal variation difference between GNORTRIP and CAMS

In Figure 3.10 we show the PM_{10} concentration difference between GNORTRIP and CAMS calculated for all of Europe for the four seasons. The largest difference in EMEP results when using GNORTRIP emissions is seen in spring, where PM_{10} concentrations are markedly higher along the road network, especially for the Scandinavian countries (see also Figure 3.6 for the annual temporal variation). This reflects the release of non-exhaust particles that have been building up during winter. During fall and winter, the PM_{10} concentrations are slightly lower in certain areas when using GNORTRIP emissions, especially for Sweden, illustrating how particles accumulate on the wet or frozen surface of the road instead of being emitted immediately. This type of temporal variation is not available in the CAMS non-exhaust emissions.

The difference between the GNORTRIP and CAMS PM_{10} annual mean concentrations is thus small, up to +2 µg/m^3 for some Nordic countries. In some very few grid squares in the Nordic countries CAMS may give slightly higher concentrations. Throughout the rest of Europe, Italy and the United Kingdom are the two countries which show higher non-exhaust contributions from CAMS, see also Figure 3.8 which indicates that national emissions from CAMS non-exhaust are higher than GNORTRIP. Since consistent non-exhaust wear rates have been used in GNORTRIP across Europe then the higher CAMS non-exhaust emissions in these countries are due to different national reporting mehtods.

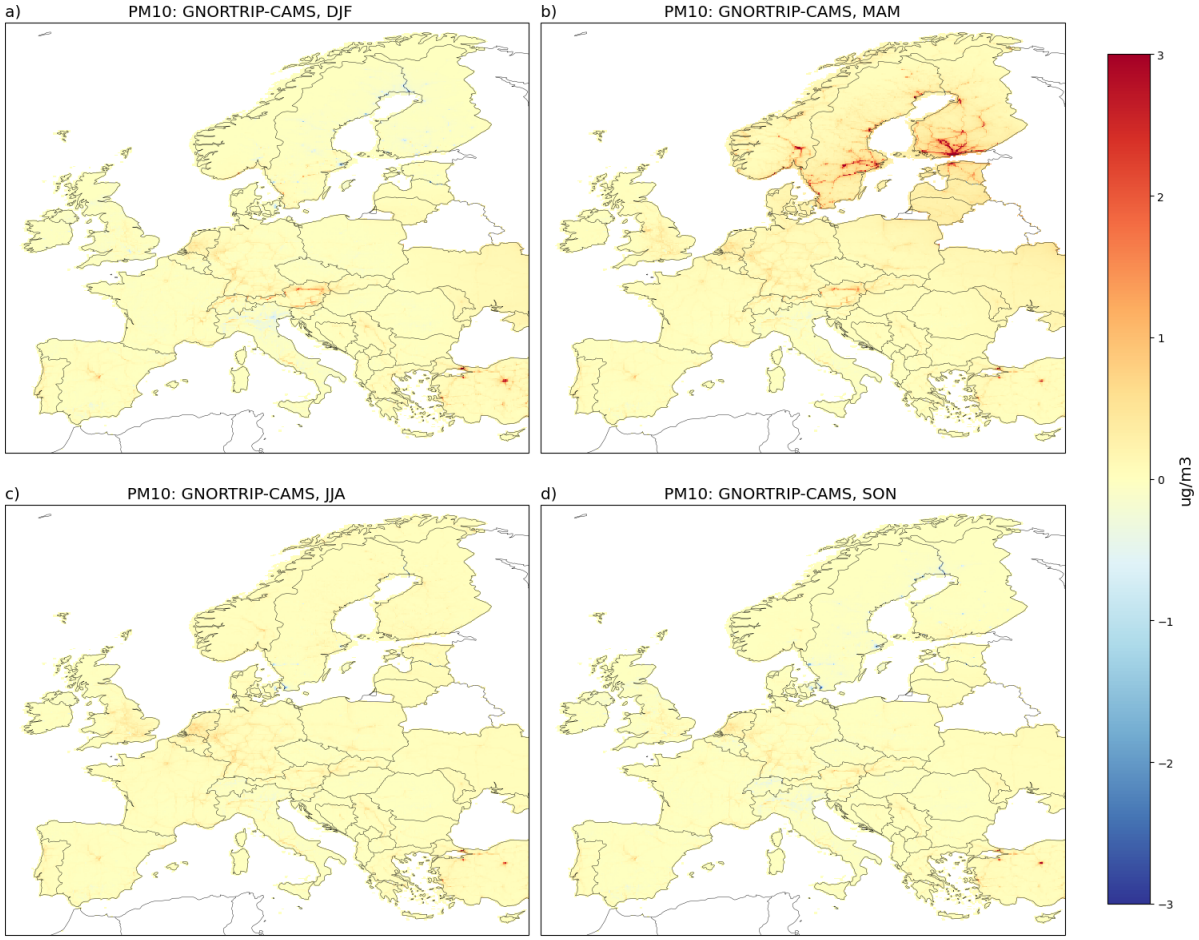


Figure 3.10. Difference in PM_{10} concentration (µg/m^3) between the GNORTRIP and CAMS simulations for winter (a), spring (b), summer (c) and fall (d).

5.3.2.3 *Relative contribution of non-exhaust emissions to PM_{10}*

In Figure 3.11 we show the contribution of non-exhaust emissions to annual mean PM_{10} concentration, as simulated with EMEP using GNORTRIP emissions. For grid cells covering many/large roads, the contribution is substantial, reaching close to 40 % in some Nordic grid cells. Generally, however, we see that along major road networks or in cities in the Nordic countries the non-exhaust contribution to annual mean PM_{10} concentrations is around 5 - 15%. In other European countries, like Germany, this contribution can be around 3 - 6% in grids containing major roads.

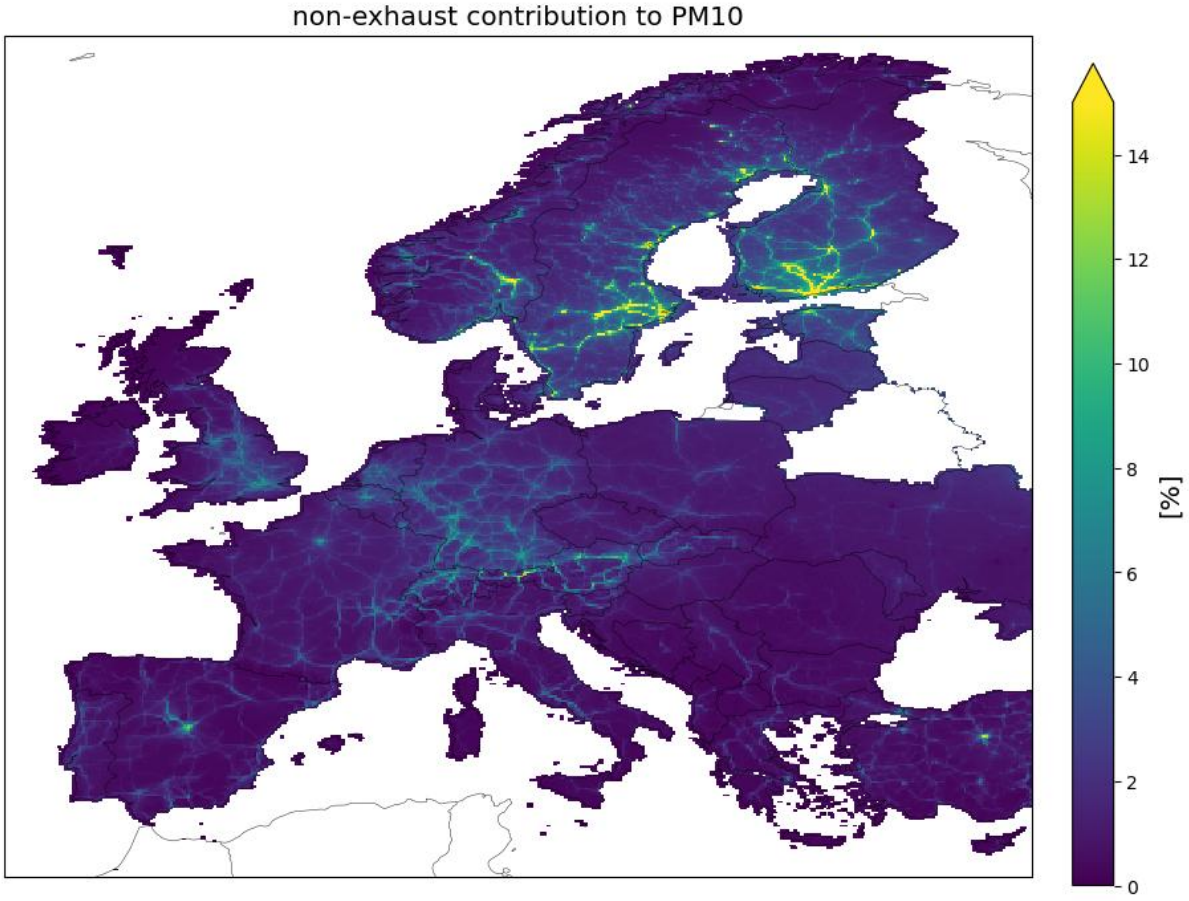


Figure 3.11. The contribution (in percent) of non-exhaust emissions to annual mean PM_{10} concentration. Data are from the GNORTRIP simulation, values averaged over the length of the simulation (year 2023). The highest contribution is 38 %.

5.3.2.4 *Examples at individual stations*

A look at individual stations, exemplified by two traffic stations in Sweden (Fig. 3.12), shows that the size of the improvement varies a lot from station to station. The SE157993 station (Fig. 3.12a) is an urban station in Stockholm with high road density, while the SE0114A station is located close to one, larger highway outside of the city. As such the urban station reflects more the urban background concentrations. This is more spatially representative of the EMEP output.

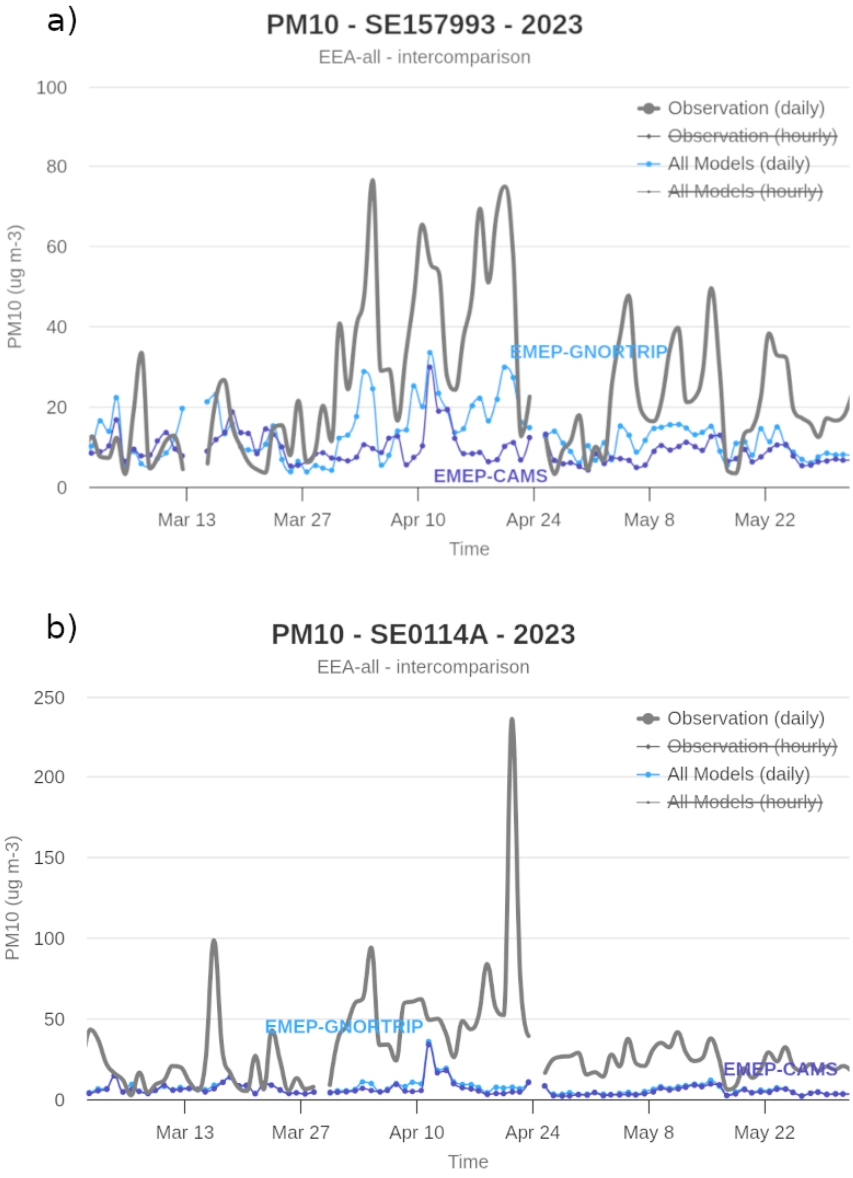


Figure 3.12. Measured and modeled PM_{10} concentration in 2023 for the two Swedish measurement stations SE157993 (a, lat, lon: 59.2013, 17.6348) and SE0114A (b, lat, lon: 59.3133, 18.004). Both of these are traffic stations, SE157993 is within the urban area of Stockholm and SE0114A is on a major highway outside of Stockholm.

5.3.2.5 Comparison to Norwegian high resolution calculations

The total concentration of PM_{10} is still underestimated during spring in Norway (and other countries) with the new emissions (Figure 3.13a) when compared with all station types. Figure 3.13b shows PM_{10} concentration as modeled by the uEMEP model for the same year (2023). uEMEP is a Gaussian dispersion model that downscales EMEP emissions, in this case to 25 m. The EMEP simulation values in Figure 3.13b (grey line 'Original EMEP') uses 2.5 km resolution meteorology, and Norway's local high-resolution emissions. More details can be found in Denby et al., 2020. It is clear that the downscaled uEMEP concentrations fit much better with observations than the GNORTRIP and CAMS concentrations in Fig. 3.13a. Also the EMEP simulation with higher resolution is slightly closer to the observations than the low

resolution simulations in Fig. 3.13a. This indicates the need for further downscaling and/or high resolution simulations to better utilize the extra information in the GNORTRIP emissions.

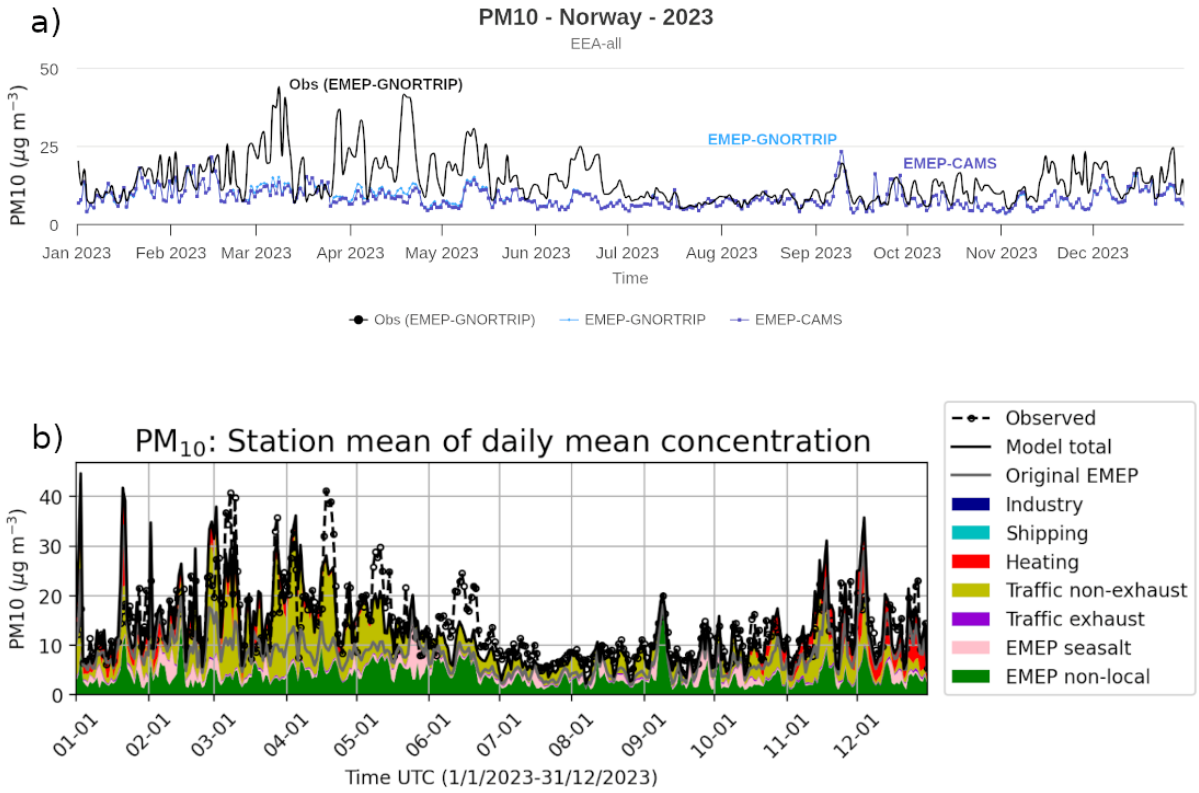


Figure 3.13. Measured and modeled PM_{10} concentrations for 2023. Values are the average over all stations. a) Observations from EAA stations, modeled values from GNORTRIP and CAMS simulations. b) Observations from Norwegian air quality stations, described here. "Model total" refers to the uEMEP model, with colors illustrating the different source contributions. Non-exhaust emissions are indicated with olive green colour. "Original EMEP" is a high resolution EMEP simulation for Norway at 2.5 km resolution (see text for details).

5.4 Conclusions GNORTRIP

The calculations carried out here with GNORTRIP are the first time such a model has been implemented for all of Europe. There remain a number of challenges in implementing such a model. The following conclusions can be drawn from this first implementation.

- The TNO gridded traffic volume dataset shows significant deviation from national totals available on Eurostat and also significant spatial deviations in a comparison in Norway with the Norwegian national traffic dataset. The TNO gridded dataset is very useful for all traffic related emission applications, but there appears to be significant uncertainty in this. An update of this dataset is desirable.
- A comparison of the gridding methodology with individual road link calculations in Norway using the same traffic data showed that the gridding methodology gives almost the same results for annual totals, in every grid, and also comparable results for daily totals. This indicates that the GNORTRIP methodology works well, given the right input data.
- Total annual non-exhaust emissions per country from GNORTRIP are often very similar to reported emissions. As a result mean concentrations do not change significantly with the implementation of GNORTRIP.
- For countries using studded tyres there can be a significant increase in the temporal correlation of daily mean PM_{10} concentrations with the implementation of GNORTRIP. In countries without studded tyres there is little change with the implementation of GNORTRIP. As a result the application of GNORTRIP could be limited to countries using studded tyres.
- The contribution of non-exhaust emissions to total PM_{10} EMEP concentrations varies significantly from country to country and grid to grid. In Nordic countries where studded tires are used this can be up to 38 % in non-exhaust emitting grids but generally around 5 - 15%. In other countries the maximum contribution of non-exhaust is less in emitting grids, at around 3 - 6%.
- Non-exhaust traffic emissions contribute significantly to PM_{10} concentrations near roads. Regional scale resolutions do not reflect this well. When high resolution modelling is applied in Norway a very good correspondence of measurements and models is found. To extract the most out of the GNORTRIP emission calculations then downscaling to appropriate scales is required.

References

Denby, B. R., Gauss, M., Wind, P., Mu, Q., Grøtting Wærsted, E., Fagerli, H., Valdebenito, A., and Klein, H.: Description of the uEMEP_v5 downscaling approach for the EMEP MSC-W chemistry transport model, Geosci. Model Dev., 13, 6303–6323, https://doi.org/10.5194/gmd-13-6303-2020 , 2020.

Denby, B.R., Sundvor, I., Johansson, C., Pirjola, L., Ketzel, M., Norman, M., Kupiainen, K., Gustafsson, M., Blomqvist, G. and Omstedt, G., 2013a. A coupled road dust and surface moisture model to predict non-exhaust road traffic induced particle emissions (NORTRIP). Part 1: road dust loading and suspension modelling. Atmos. Environ. 77, 283-300. DOI: http://dx.doi.org/10.1016/j.atmosenv.2013.04.069

Denby, B.R., Sundvor, I., Johansson, C., Pirjola, L., Ketzel, M., Norman, M., Kupiainen, K., Gustafsson, M., Blomqvist, G., Kauhaniemi, M. and Omstedt, G., 2013b. A coupled road dust and surface moisture model to predict non-exhaust road traffic induced particle emissions (NORTRIP). Part 2: surface moisture and salt impact modelling. Atmos. Environ. 81, 485-503. DOI: http://dx.doi.org/10.1016/j.atmosenv.2013.09.003

EMEP/EEA air pollutant emission inventory guidebook 2023, EEA Report 06/2023. URL: https://www.eea.europa.eu/en/analysis/publications/emep-eea-guidebook-2023. Road, tyre and brake wear URL: https://www.eea.europa.eu/en/analysis/publications/emep-eea-guidebook-2023/part-b-sectoral-guidance-chapters/1-energy/1-a-combustion/1-a-3-b-vi/@@download/file

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, Earth Syst. Sci. Data, 14, 491–515, https://doi.org/10.5194/essd-14-491-2022 , 2022.

Mu, Q., Denby, B. R., Wærsted, E. G., and Fagerli, H.: Downscaling of air pollutants in Europe using uEMEP_v6, Geosci. Model Dev., 15, 449–465, https://doi.org/10.5194/gmd-15-449-2022 , 2022.

Müller, M., Homleid, M., Ivarsson, K., Køltzow, M. A., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M., and Vignes, O.: AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction, Model, Weather Forecast., 32, 609–627, https://doi.org/10.1175/WAF-D-16-0099.1 , 2017

6 New dust emission scheme in IFS-COMPO and application to high latitude dust forecasting

6.1 New dust emission scheme in IFS-COMPO

The dust emission in SILAM is expressed as:

$$F_{dust} = A\, r^{-½}\, max(U_{10} - U_{10,min})^3 L_f\, max(\,(sd_t - sd)/sd_t, 0)\, max((LAI_t - LAI)/LAI_t, 0)$$

where:

- U_{10} is the 10 m wind speed,
- $U_{10,min} = U_t + 50 * VSM\ (m/s)$ with VSM the volumetric soil moisture fraction,
- U_t is the threshold velocity of 5 m/s
- L_f is the land fraction of the cell,
- r is the roughness length,

- sd is the water equivalent snow depth,
- sd_t is a water equivalent snow depth threshold, currently 0.02 m,
- LAI is the leaf area index,
- LAI_t is a leaf area index threshold, currently $1.5 m^2/m^2$,
- A is a scaling constant that also sets the correct dimension.

An adapted version has been adapted into IFS-COMPO, with the following expression:

$$F_{dust} = PLSM\ A\, r^{-½}\, max(WG_{10} - U_{10,min})^3 L_f\, max(\,(sd_t - sd)/sd_t, 0)\, max((LAI_t - LAI)/LAI_t, 0)$$

where:

- $PLSM$ is the fraction of the grid cell composed of land
- WG_{10} is the 10 m wind gusts,
- $U_{10,min} = U_t + 50 * VSM\ (m/s)$ with VSM the volumetric soil moisture fraction,
- U_t is the threshold velocity as provided by a monthly input from Pu et al (2020)
- L_f is the land fraction of the cell,
- r is the roughness length,

- sd is the water equivalent snow depth,
- sd_t is a water equivalent snow depth threshold, currently 0.02 m,
- LAI is the leaf area index,
- LAI_t is a leaf area index threshold, currently $1.5 m^2/m^2$,
- A is a scaling constant that also sets the correct dimension.

The soil wetness used in the computation of the volumetric soil moisture fraction is normalized by the fraction of the grid cell composed of land. This is because the soil wetness is null over ocean grid cell, which can artificially increase emissions over grid cell that are composed of both ocean and continent.

This new emission scheme doesn't use two inputs : the roughness length and the wind erosion threshold velocity. For the roughness length, three specific inputs have been tested, as provided by FMI:

- roughness length from the scatterometer onboard the European Remote Sensing (ERS) (Prigent et al 2005)
- roughness length from the Advanced Scatterometer (ASCAT) onboard METOp
- roughness length from the ASCAT modulated by orography

The best results in terms of skill versus AERONET have been obtained using the roughness length from Prigent et al (2005), so only results with this version are shown. The term $r^{-½}$ is shown in Figure 4.1, which highlights the areas with high desert dust emissions : parts of Sahara, the Southern Arabian Peninsula, the Taklimakan and Gobi areas.

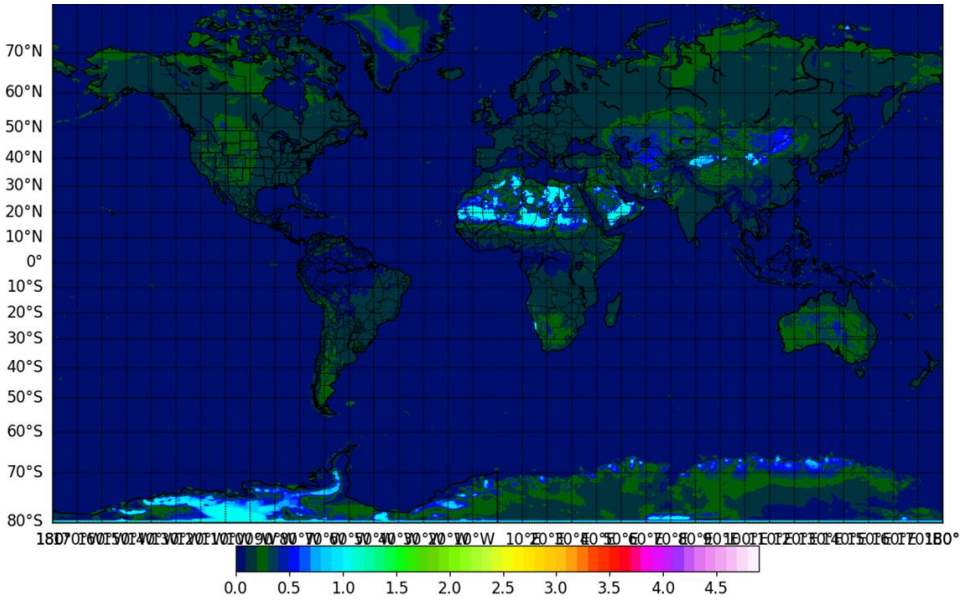


Figure 4.1 : surface roughness from Prigent et al (2005) at the power -0.5. The higher values correspond to areas with preferential dust emissions.

Mineral dust particles are lifted from dry and bare soils into the atmosphere by saltation and sandblasting. This process is initiated when surface winds reach a threshold velocity of wind erosion. The value of this wind erosion threshold depends on soil and surface characteristics, including soil moisture, soil texture, and particle size, and the presence of pebbles, rocks, and vegetation residue, and thus it varies spatially and temporally. However, due to the lack of information, many models use a fixed value, often between 5 and 6.5 m/s. Pu et al (2020) implement a methodology to derive the threshold velocity from MODIS AOD observations and wind speed from reanalysis. The result is a monthly gridded (0.5x0.5°) dataset, which we fill with values of 5 m/s where no values are provided. The threshold velocities for January and May are shown in Figure 4.2. Lower values than 5 m/s over parts of Southern and Eastern Sahara will mean higher desert dust emissions, while values higher than 5 m/s over Taklimakan/Gobi and other areas will result in lower desert dust emissions. Further adjustments of the inputs of the new dust emission scheme has been tested to improve the prediction of high latitude dust. These adjustments are described and evaluated in section 6.3.

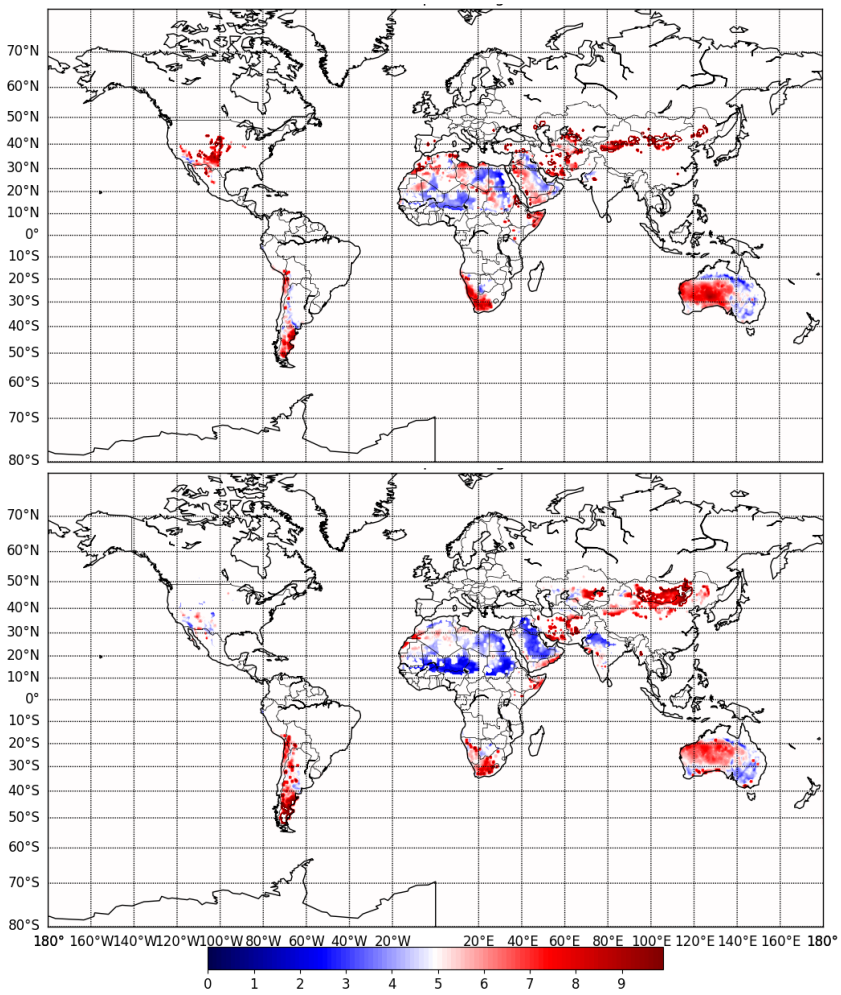


Figure 4.2 : Threshold velocity from Pu et al (2020), gap-filled with 5 m/s values; January (top) and May (bottom).

6.2 **Evaluation**

In this subsection, the impact of the implementation in IFS-COMPO of the new dust emission scheme described in section 6.1 are evaluated against a wide range of observational datasets. The new dust emission scheme is evaluated and compared to the operational dust emission scheme through a series of forecast only experiments (without AOD data assimilation) using the usual resolution of global CAMS products : TL511 (40km grid cell) and 137 vertical levels. The forecast only simulations cover the years 2017 and 2019 with a one month spinup time, and additionally the spring 2020, including the June 2020 exceptional “Godzilla” dust event. The experiments are denoted “REF” for the operational scheme, and “NEW” for the new dust emission scheme.

First, we derive a reference dust AOD at 550nm dataset computed from the merged AOD at 550nm product from FMI (Sogatcheva et al. 2020), and using the ratio of dust AOD/AOD provided by the MIDAS dataset (Gkikas et al. 2020) to compute monthly dust AOD at 550nm over the 1995-2017 period. Figure 4.3 shows a comparison of observed and simulated dust AOD at 550nm by REF and NEW, along with the RMSE between monthly observed and simulated dust AOD at 550nm for the year 2017. The NEW experiment improves the simulated dust AOD at 550nm over most of Sahara and over the Atlantic, with a decreased RMSE. However, over the Arabic peninsula and the Taklimakan/Gobi areas, dust AOD at 550nm simulated by NEW is often lower than REF, which brings a degradation of the RMSE over these regions.

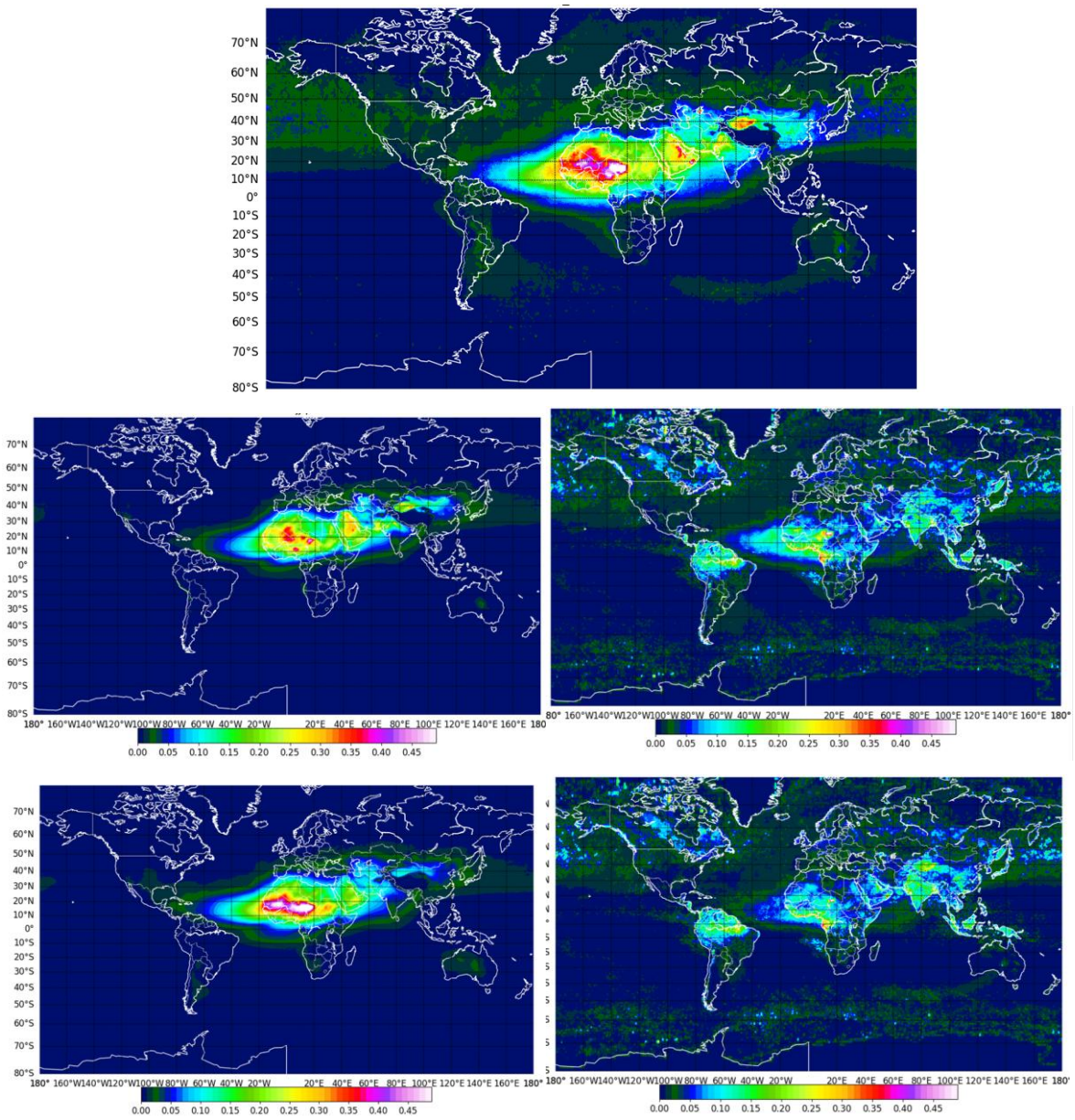


Figure 4.3 : 2017 average. Top, dust AOD at 550nm derived from the merged AOD dataset from Sogacheva et al (2020) and the MIDAS dataset. Left middle and bottom: simulated dust AOD at 550nm by REF and NEW. Right middle and bottom : RMSE of monthly simulated versus observed dust AOD at 550nm.

Figure 4.4 shows an evaluation of simulated AOD at 500 and 1020nm against AERONET data, averaged over all stations (421 sites globally) in 2017. In order to assess the skill of the forecast, FGE is also shown. Desert dust has a relatively larger contribution to AOD at 1020nm than at 500nm. REF shows a low bias at 1020 nm, which is partly reduced by NEW. At 500nm, the impact is smaller and generally positive. The impact on global FGE is small at 500nm and significant and positive at 1020nm, particularly outside of spring months.

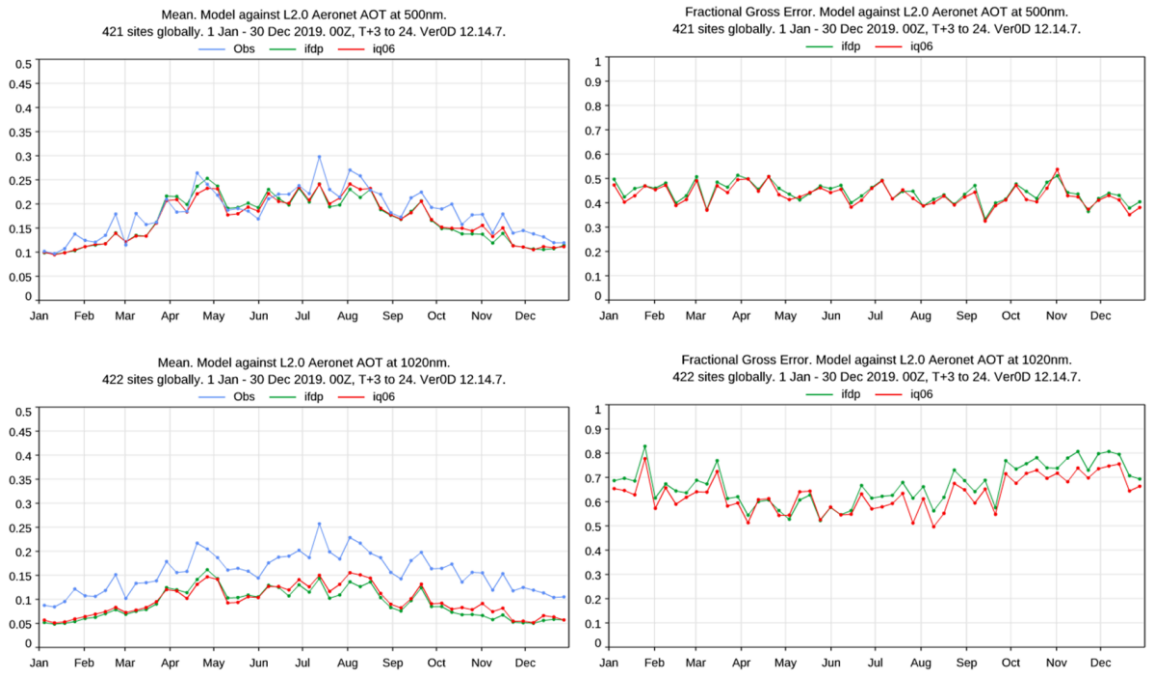


Figure 4.4 : 2017; left observed (blue line) and simulated (**green, REF**; **red, NEW**)weekly AOD at 500 (top) and 1020 (bottom) nm. Right, fractional gross error (FGE) of simulated AOD at 550nm (top) and 1020nm (bottom). Average over all AERONET sites

Figure 4.5 shows another evaluation versus AERONET AOD, focusing on the correlation between simulated observed daily AOD at 500 and 1020nm, and applying a filter on the observed Angstrom exponent : only observations with values below 0.3 are selected, which are representative of dusty observed situations over all AERONET stations. The NEW experiment brings the simulated mean closer to the observed mean at both 500 and 1020nm, although the impact is larger at 1020nm. The positive impact of NEW is also shown in the mean correlation factor between observed and simulated values, which is 0.03 higher at both wavelengths. Although this may look like a small improvement, this represents an average improvement over all AERONET stations for a whole year, so it is quite significant.

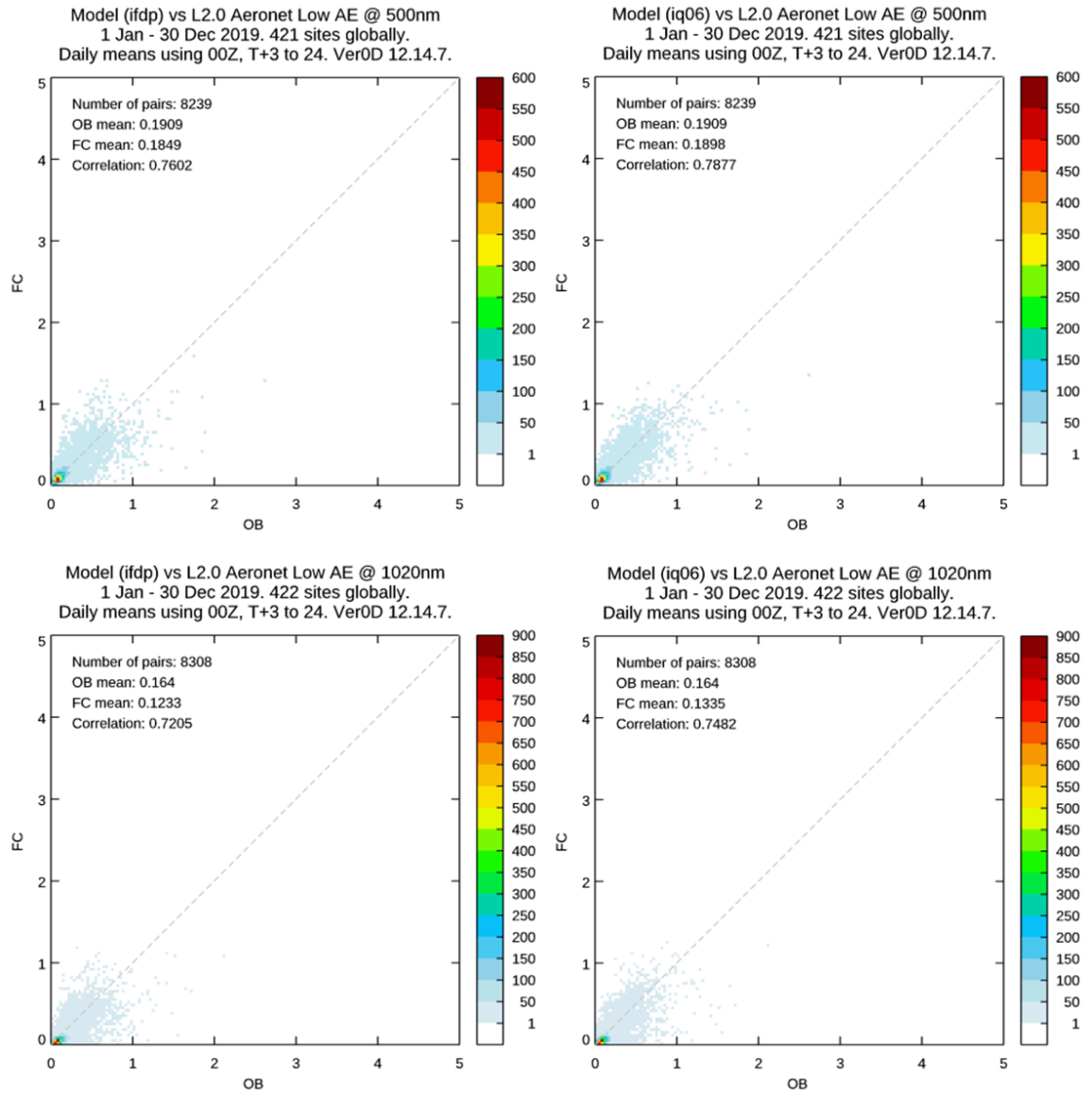


Figure 4.5 : 2017, density scatterplots of observed versus simulated daily AOD at 500nm (top) and 1020nm (bottom). REF experiment (left) and NEW (right). Average over all AERONET stations; only observations with an Angstrom exponent below 0.3 are included, representative of observed dusty situations.

Figure 4.6 shows the weekly simulated and observed Angstrom exponent as well as its FGE aggregated over all AERONET stations. REF shows a small positive bias all year long except in July/August, meaning that the simulated size distribution is skewed towards finer particles. NEW is lower than REF by around 0.05, thus reducing this low bias. The FGE is also improved by NEW over most of the year. Also, the correlation factor (not shown) is improved by NEW as compared to REF, with a value of 0.67 versus 0.64 for REF.

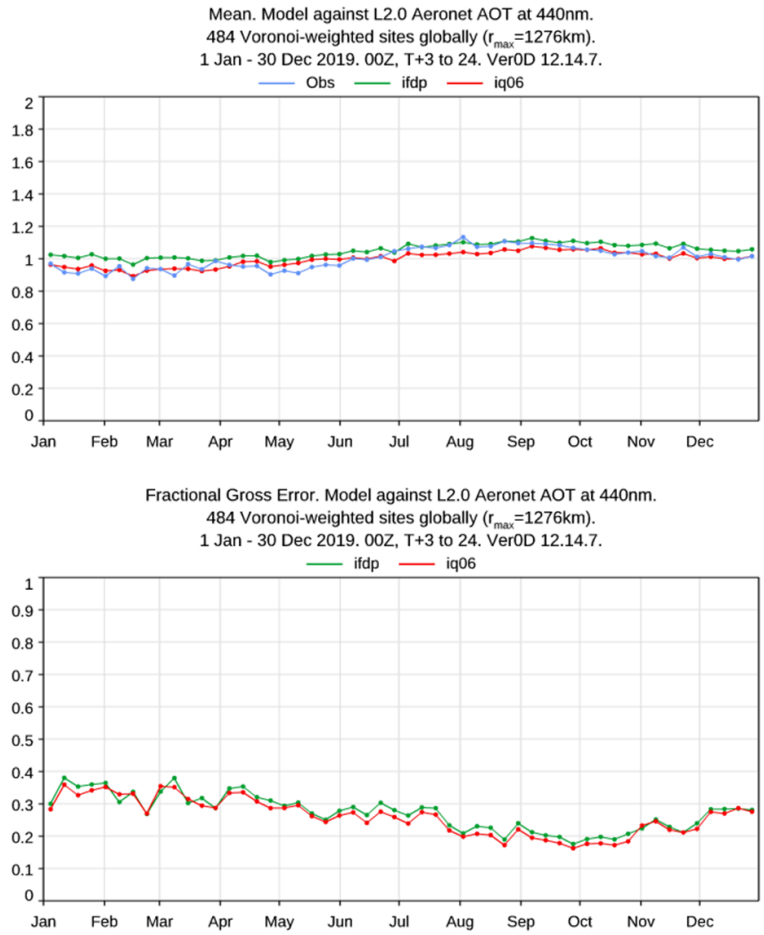


Figure 4.6 : 2017, observed (blue line) and simulated (**green, REF**; **red, NEW**)weekly Angstrom exponent. Average over all AERONET sites

Finally, Figure 4.7 shows the impact of the NEW experiment on simulated PM10 over background rural stations in Europe, China and U.S. Over Europe, PM10 is often underestimated by REF. The changes brought by NEW are in general small; the simulated PM10 is slightly higher, which helps somewhat the low bias. Over China, REF shows a significant low bias in winter/autumn seasons, and a small positive bias during summertime. NEW brings an increase of simulated PM10 in winter and autumn, and a decrease in summer, which clearly brings an improvement of the simulated PM10 over China. This is confirmed in Figure 4.8, which shows a density scatterplot of simulated and observed daily PM10 over background rural stations in China for the two REF and NEW experiments : the correlation factor of NEW is 0.62 against 0.4 for REF, a very nice improvement. Over US, PM10 shows a persistent underestimation over background rural stations, with a small improvement with NEW.

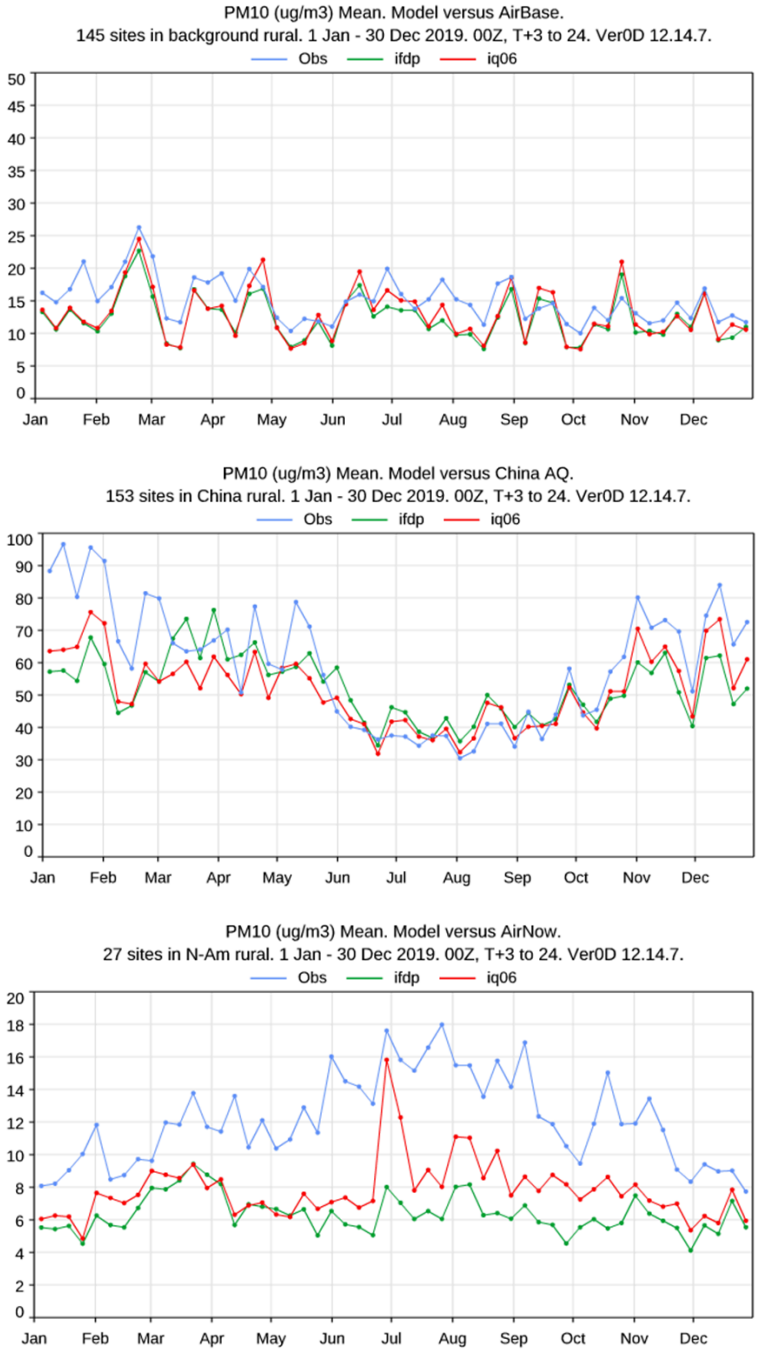


Figure 4.7 : 2017, observed (blue line) and simulated (**green, REF**; **red, NEW**) weekly PM10 over background rural stations in Europe (top), China (middle) and U.S. (bottom).

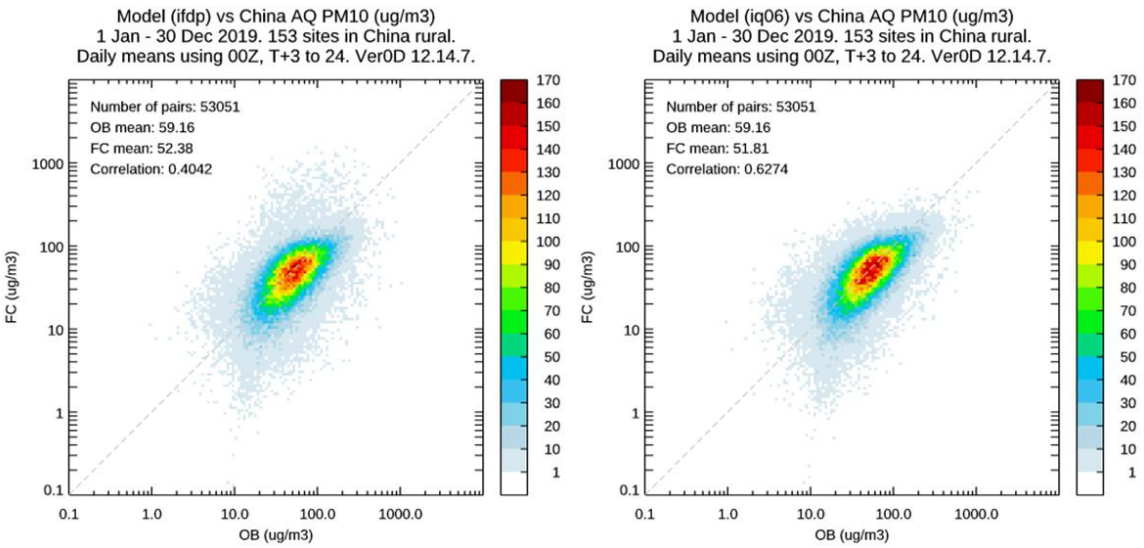


Figure 4.8 : 2017, density scatterplots of observed vs simulated daily PM10 over background rural stations in China. REF (left) and NEW (right).

In Figure 4.9, the NEW and REF experiments are compared for PM10 simulated over Texas and Florida at the end of June 2020. An exceptional dust event originating from Western Sahara struck most of the Caribbean islands and then Texas at the end of June 2020, which resulted in exceptionally high observed PM10 over Texas and to a lesser extent over Florida. This is very apparent in Figure 4.9, with observed values above 100 µg/m3 over Texas on 27th and 28th of June 2020. REF underestimated PM10 for this event, by a wide margin over Texas, and by a smaller one over Florida. The NEW experiment significantly improved simulated PM10 over Texas on 27/28th of June but also for the following days, but brought a large overestimation over Florida.

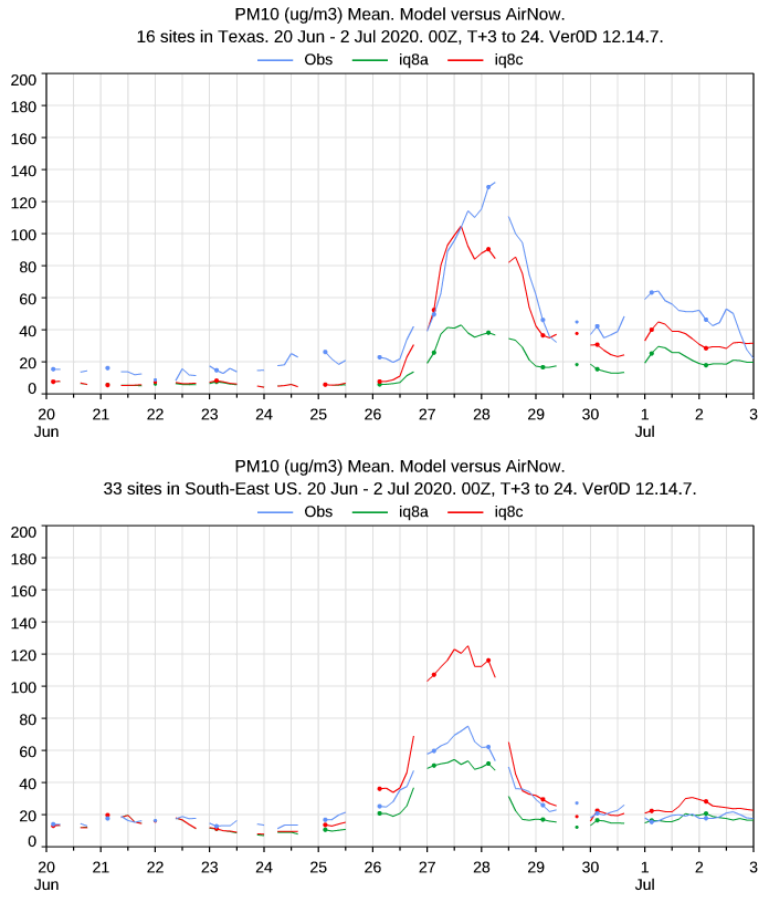


Figure 4.9 : 20/6 to 3/7/2020, observed (blue line) and simulated (**green, REF**; **red, NEW**) 3-hourly PM10 over stations in Texas (top) and Florida (bottom).

6.3 High latitude dust

While most of the mineral dust production occurs over the well known hot deserts around the world such as the Sahara, Taklimakan, etc., a significant fraction of dust emissions occurs over bare areas in the Arctic and Antarctic, in regions such as Iceland, Greenland, Svalbard, Alaska, Canada, Antarctica, New Zealand, and Patagonia. High latitude dust has been the focus on increasing research in recent years ((Arnalds et al., 2016; Bullard et al., 2016; Groot Zwaafting et al., 2016, 2017; Meinander et al, 2022). Estimates of the contribution of high latitude dust to the total dust emissions vary from 3 to 5% (Bullard et al. 2016, Groot Zwaafting et al, 2016).The particles that are being lifted are often from volcanic origin over Iceland, or provided by glacial processes, past or present : the so-called “glacial flour”.

The representation of high latitude dust sources is often a challenge in global models, as is its evaluation because of the lack of observations in the Arctic and Antarctic. High latitude dust sources are often localized and episodic, and the signal on the AOD observations from the AERONET network or from remote sensing is weak. High latitude dust shows a clear signal in PM10 observations; however, the stations are also quite scarce in the Arctic. In this section, we’ll focus mostly on Iceland, which has a good record of high latitude dust studies (Dagsson-Waldhauserová et al. 2014, 2015, 2016) and some observations of PM10, even if mostly clustered around Reykjavik. Also, an in-depth study of the dust emitting areas of Iceland has been carried out by Arnalds et al (2014, 2016), which produced a map of areas vulnerable to

erosion shown in Figure 4.10. The main dust emitting regions are a thin strip along the Southern coast, and volcanic areas in center Iceland.

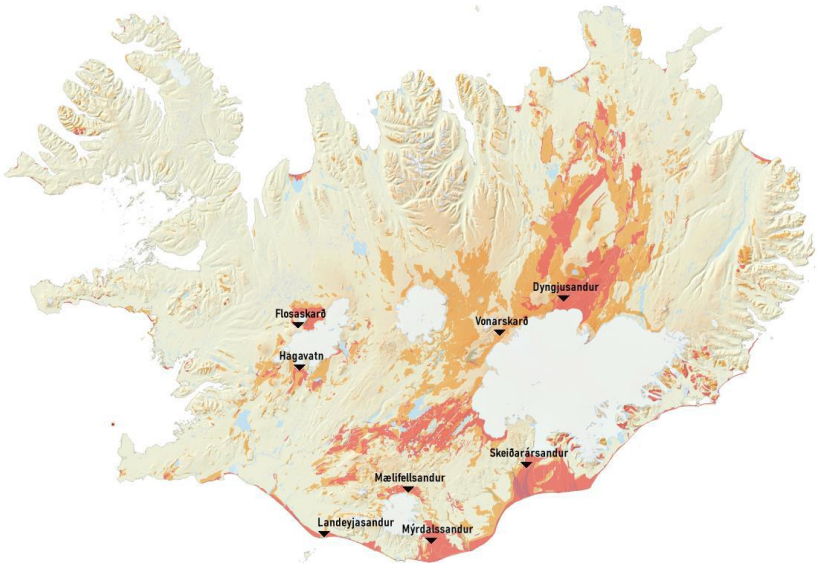


Figure 4.10 : Areas vulnerable to erosion from Arnalds et al. (2014); extreme areas are shown in red, severe in orange. Hot spots of dust emissions are shown with black triangles.

6.3.1 High latitude dust developments in IFS-COMPO

The current operational dust emission scheme in IFS-COMPO is ill equipped to represent high latitude dust emissions. It relies on a dust source function which originates on an observed AOD frequency of occurrence dataset : high latitude sources are totally absent from it (see Figure 4.11).

In order to better represent high latitude dust sources with the operational scheme, a dust source function (DSF) that includes high latitude sources from Groot Zwaafting et al., 2017 has been merged with the current operational dust source function. The resulting dust source function clearly accounts for the polar sources as shown in Figure 4.11. Dust sources over Iceland, Greenland, Spitzberg, parts of Canada and Alaska and Northern Siberia are clearly visible. However, while the operational DFS is monthly, the high latitude information is the same for every month, which can be the cause of too frequent high latitude dust emissions. Over Iceland (Figure 4.12), no dust sources are present with the operational DSF, while areas with potential dust emissions along the Southern coast and the interior appear with the new DSF.

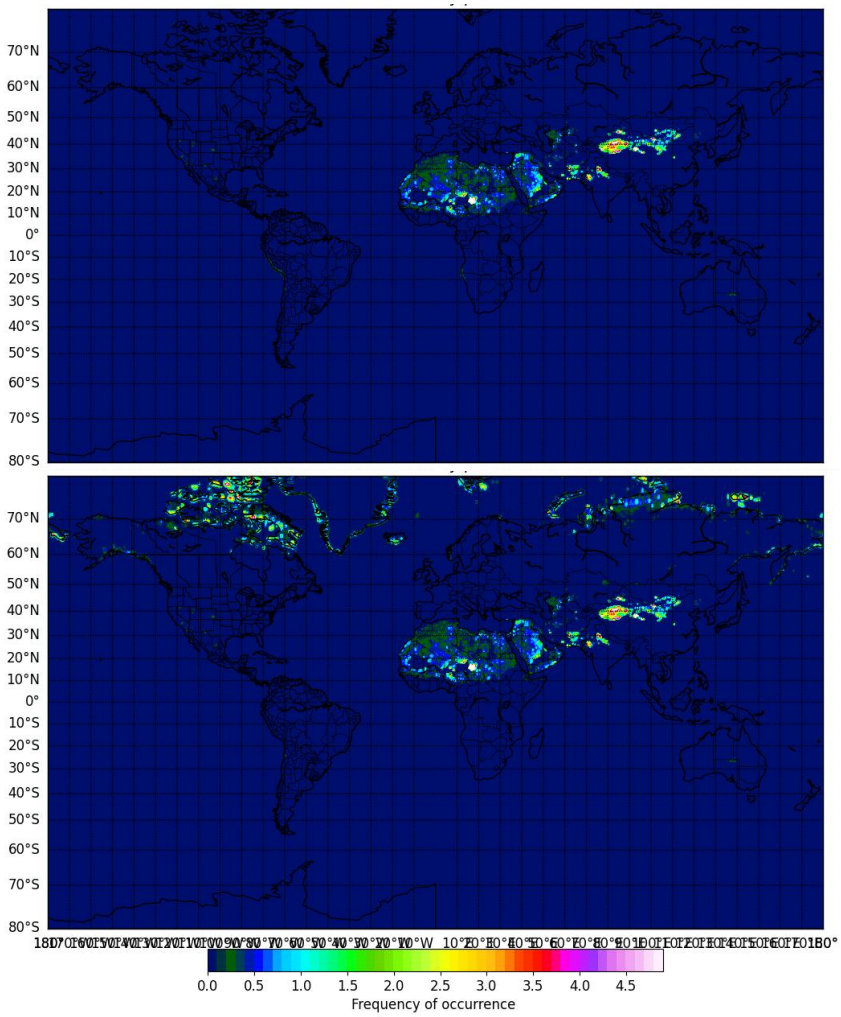


Figure 4.11 : Dust source function used in cycle 49R1 of IFS-COMPO for May (top), and merged with high latitude dust sources information from Groot Zwaafting et al, 2017 (bottom).

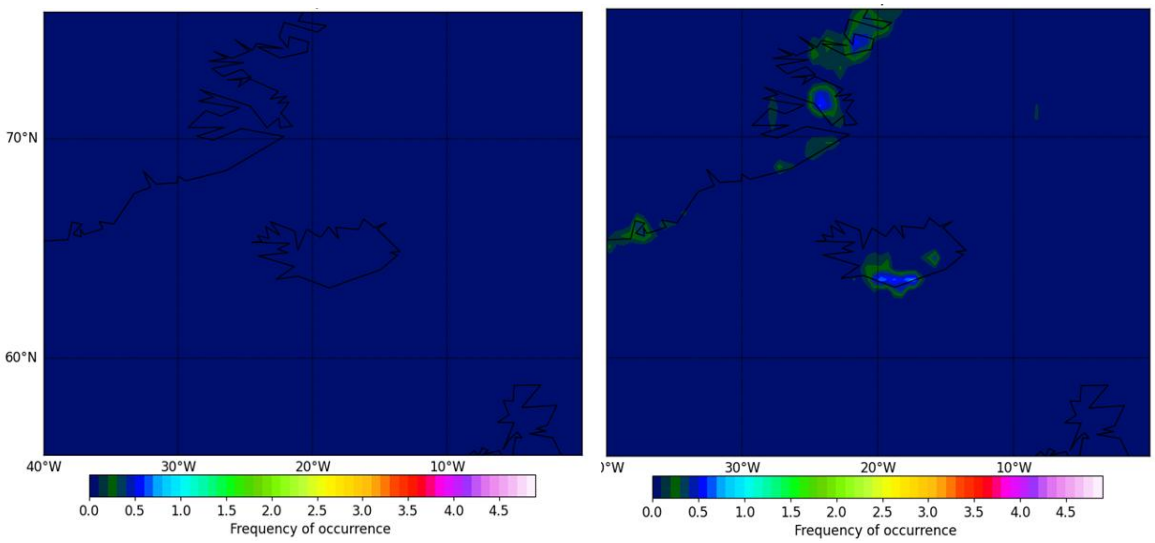


Figure 4.12 : Same as Figure 4.11, zoomed over Iceland.

Figure 4.13 shows the impact of using the new dust source function in the operational scheme in terms of simulated PM10 in July 2019. The simulated values increase by 5 to 10 µg/m3 on average, which represents a more than doubling over large swathes of Northern Canada and Siberia. Over Iceland, the PM10 values are increase by more than 50% on average for this particular month.

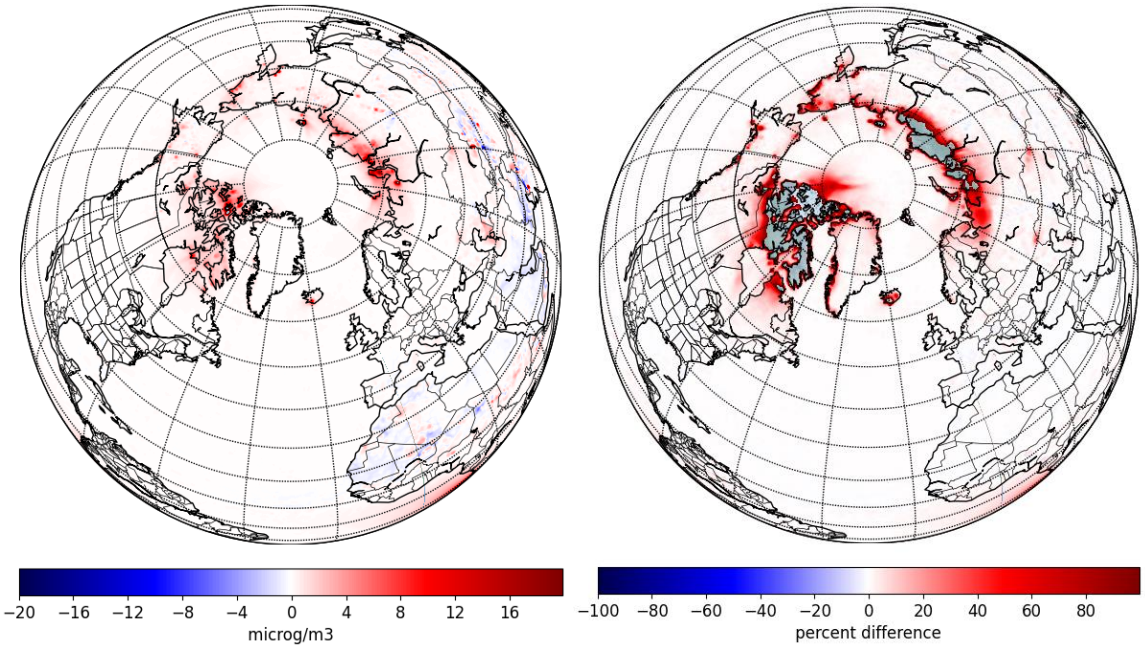


Figure 4.13 : July 2019 simulated PM10 with the operational emission scheme : absolute (left) and relative, in percentage (right) difference of the specific high latitude developments as compared to the version that doesn't include these developments.

The new dust emission scheme that has been implemented and tested in CAMAERA, described in section 6.1, doesn't rely on dust source function. Because of that, it is better positioned to represent high latitude dust sources. In order to improve the representation of high latitude dust, the information about high latitude dust sources from Groot Zwaafting et al., 2017 have been combined in the roughness length input of the dust emission scheme. Additionally, the threshold velocity for dust emissions has been lowered from 5 to 2.5 m/s for areas above 50°N and below 50°S. Figure 4.14 shows the impact of these developments on simulated PM10 : the same areas as in Figure 4.13 are easily recognizable. The high latitude specific developments of the new dust emission scheme increase simulated PM10 in July 2019 by 40 to 60% over many areas in Canada, Greenland, Iceland and Siberia.

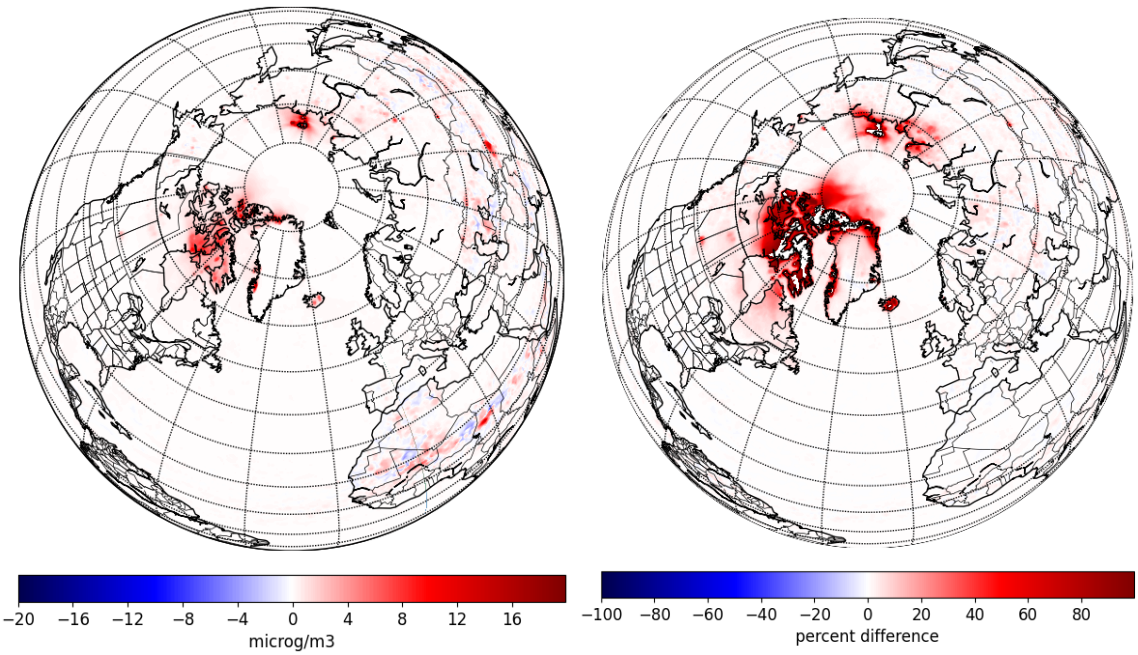


Figure 4.14 : July 2019 simulated PM10 with the new dust emission scheme : absolute (left) and relative, in percentage (right) difference of the specific high latitude developments as compared to the version that doesn't include these developments.

6.3.2 Evaluation focusing on Iceland

Two series of forecast only experiments (ie without AOD data assimilation) with cycle 49R1 IFS-COMPO at the usual CAMS resolution (TL511 meaning around 40km horizontal resolution and 137 levels over the vertical). were carried out for the summer 2019, listed in Table 4.1.

Table 4.1 : Experiments carried out for the summer of 2019

experiment name	experiment id	specifics
REF	imqf	Reference experiment using the operational dust emission scheme
REF_HLD	imq4	Experiments using the operational dust emission scheme with a dust source function taking into account HLD sources.
NEW	im1q	Experiment with the new dust emission scheme
NEW_HLD	ipc3	Experiment with the new dust emission scheme with HLD specific developments

The mean simulated PM10 in June 2019 over Iceland is shown in Figure 4.15 for all four experiments. Mean PM10 from sea-salt aerosols are simulated around 12-20 µg/m3 in the Northern Atlantic. With REF, over central Iceland, the simulated PM10 is very low, below 3 µg/m3, as no source of desert dust is simulated. REF_HLD see an increase of simulated PM10 along the South coast, with mean values around 15 µg/m3. This corresponds to the high erosion area along the Icelandic South coast as shown in Figure 4.9. The NEW experiment increases simulated PM10 over two areas of the central Plateau, West and North of the Vatnajökull icecap, which all correspond to areas of high dust erosion in Figure 4.9. The simulated values reach 15 to 20 µg/m3 close to the source areas. No desert dust emissions are represented along the South coast with NEW. Finally, NEW_HLD increases the desert dust emissions over the same areas as NEW, up to more than 30 µg/m3.

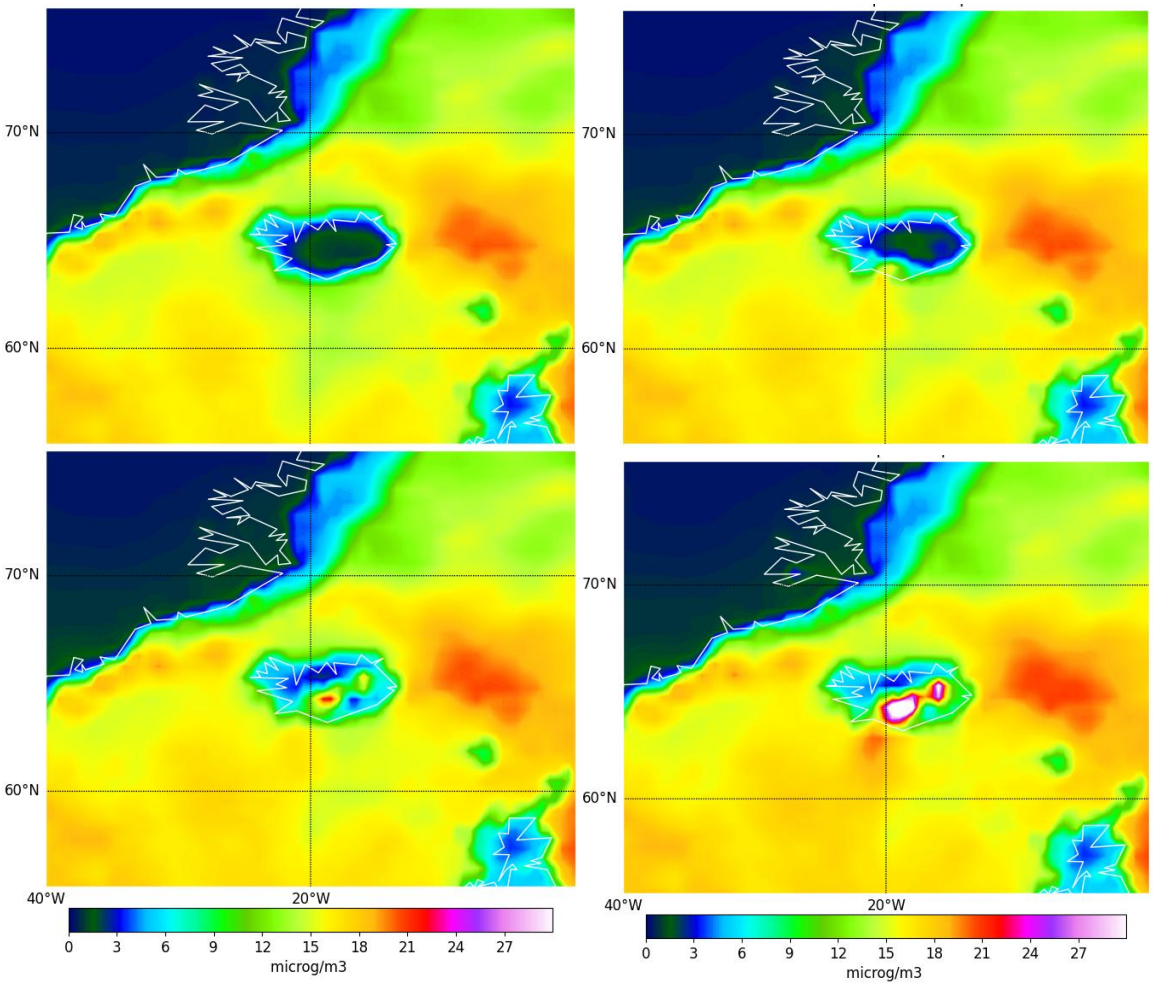


Figure 4.15 : June 2019 simulated PM10 with the following experiments, from left to right and top to bottom : REF, REF_HLD, NEW and NEW_HLD.

Figure 4.16 shows a comparison of simulated and observed daily PM10 over Icelandisc aribase/EEA stations. Out of 6 observing sites, 5 are located close to the capital city, Reykjavik, so the results are representative mainly over this area. The observed PM10 show several spikes above 15 µg/m3 in June, around 10th of July, and in early August, which are caused by local desert-dust sources, as the spikes are much less prominent in the PM2.5 plots (not shown), indicating a majority of coarse particles, and no dust intrusion over Reykjavik from sources outside of Iceland was simulated or observed in the studies period (there was one at the end of April, correctly simulated by the REF and NEW experiments). Those spikes are not simulated at all by REF. REF_HLD simulates them much better, in particular for the early August event. The June and July events are underestimated by REF_HLD. The NEW

experiments shows a better skill than REF in simulating the June dust events, but improves little on REF for the July and August events. The NEW_HLD improves further for the June events, and has a small and positive impact for the July and August events. The differences between REF_HLD one one hand and NEW/NEW_HLD on the other hand are caused by the differences in simulating the Iceland dust emission areas. NEW_HLD seems to simulated quite well the dust sources along the Southern coast, but much less well sources from the central plateau, while this is the reverse for NEW and NEW_HLD. It is probably the the dust event from early August originated mostly from the Southern coast, while the early June dust events were more from the central Plateau.

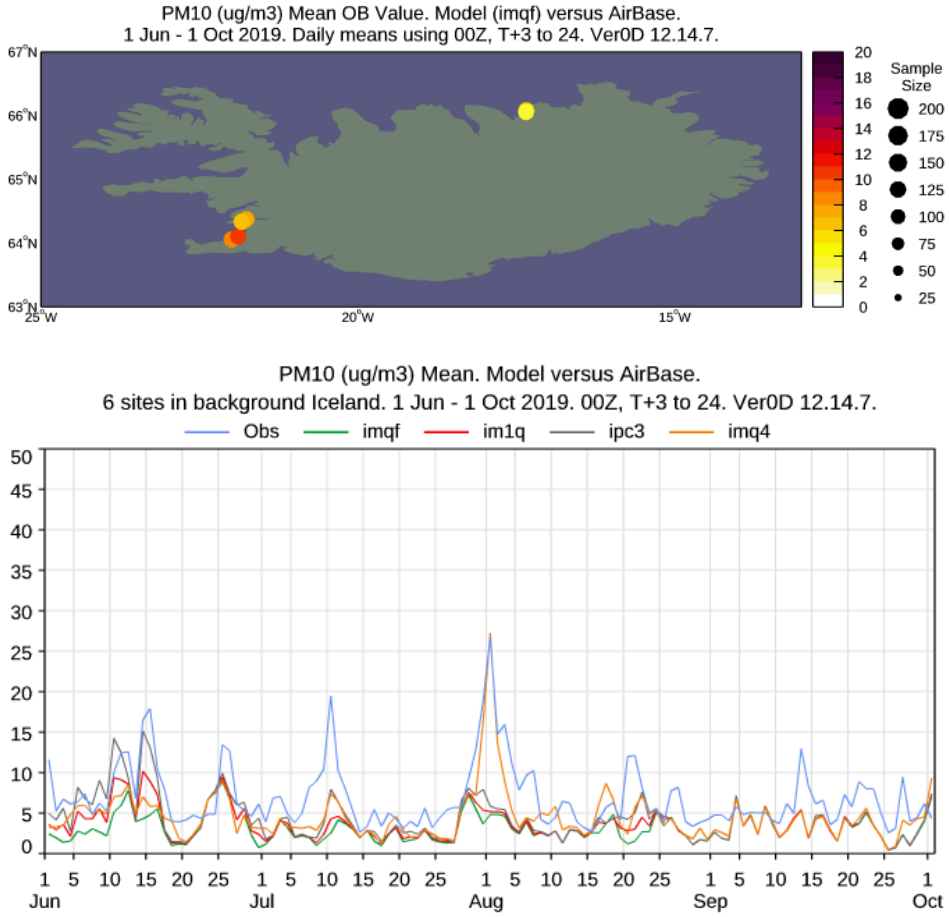


Figure 4.16: Summer 2019, observed (blue line) and simulated (**green, REF**; **red, REF_HLD; gray, NEW; orange, NEW_HLD**) daily PM10 over stations in Iceland. Top, location of the 6 Icelandic stations from EEA/Airbase.

Figure 4.17 shows how the observed and simulated PM10 correlates over Icelandic EEA/Airbase stations. For REF, the correlation factor is low, at 0.28. REF_HLD improves a lot on the correlation and bias: the correlation factor increases up to 0.59. NEW increases the correlation factor as compared to REF to 0.39, and NEW_HLD to 0.43.

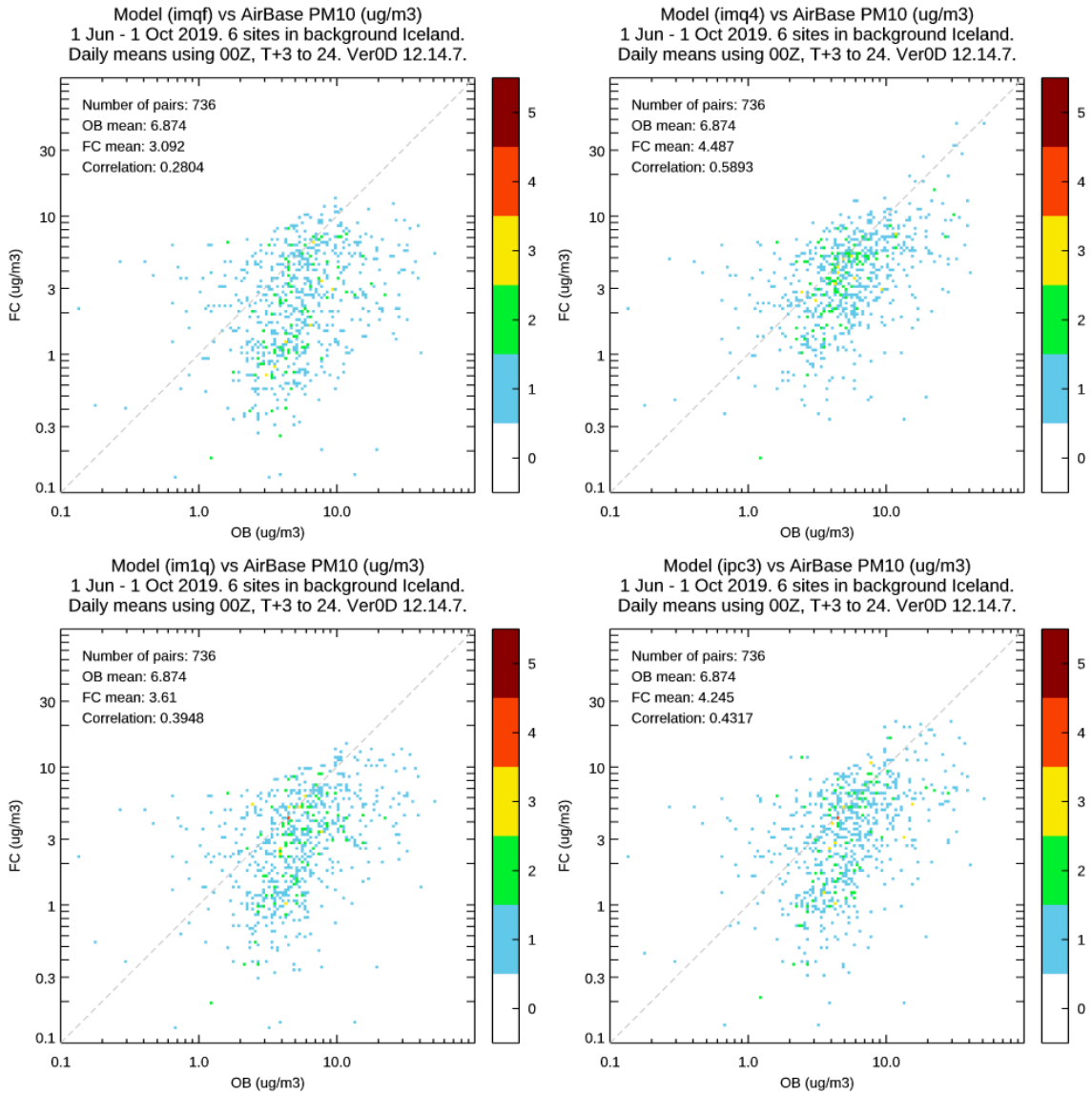


Figure 4.17 : Summer 2019, density scatterplots of observed versus simulated daily PM10 over all Iceland EEA/Airbase stations. From left to right and top to bottom : REF, REF_HLD, NEW and NEW_HLD.

6.4 Conclusion for IFS-COMPO dust emission schemes and high latitude dust

A new dust emission scheme, adapted from the SILAM dust emission scheme, has been implemented into IFS-COMPO and evaluated. The impact on most of the skill scores of simulated AOD and PM10 are positive over a wide range of evaluation periods and regions, despite some local degradations. The new dust emission scheme, together with high latitude specific developments, also improves the simulation of Icelandic dust sources, and implements a first representation of high latitude dust sources. However, high latitude sources that are located along coastlines, particularly over Iceland, are not well represented with the new dust emission scheme.

References:

Arnalds, O., Olafsson, H., and Dagsson-Waldhauserova, P.: Quantification of iron-rich volcanogenic dust emissions and deposition over the ocean from Icelandic dust sources, Biogeosciences, 11, 6623–6632, https://doi.org/10.5194/bg-11-6623-2014, 2014.

Arnalds, O., Dagsson-Waldhauserová, P., and Olafsson, H.: The Icelandic volcanic aeolian environment: Processes and impacts – A review, Aeolian Res., 20, 176–195, https://doi.org/10.1016/j.aeolia.2016.01.004, 2016.

Bullard, J. E.: The distribution and biogeochemical importance of high-latitude dust in the Arctic and Southern Ocean-Antarctic regions, J. Geophys. Res.-Atmos., 122, 3098–3103, https://doi.org/10.1002/2016jd026363, 2016.

Bullard, J. E., Baddock, M., Bradwell, T., Crusius, J., Darlington, E., Gaiero, D., Gassó, S., Gisladottir, G., Hodgkins, R., McCulloch, R., McKenna-Neuman, C., Mockford, T., Stewart, H., and Thorsteinsson, T.: High-latitude dust in the Earth system, Rev. Geophys., 54, 447–485, 2016.

Dagsson-Waldhauserová, P., Arnalds, O., Ólafsson, H., Skrabalova, L., Sigurðardóttir, G., Branis, M., Hladil, J., Skala, R., Navratil, T., Chadimova, L., Löwis, S., Thorsteinsson, T. Carlsen, H., and Jónsdóttir, I.: Physical properties of suspended dust during moist and low wind conditions in Iceland, Iceland. Agr. Sci., 27, 25–39, 2014.

Dagsson-Waldhauserová, P., O. Arnalds, H. Olafsson, J. Hladil, R. Skala, T. Navratil, L. Chadimova, O. Meinander: Snow–Dust Storm: Unique case study from Iceland, March 6–7, 2013, Aeolian Res., 16, 69–74, 2015.

Dagsson-Waldhauserova, P., Magnusdottir, A. Ö., Olafsson, H., and Arnalds, O.: The spatial variation of dust particulate matter concentrations during two Icelandic dust storms in 2015, Atmosphere, 7, 77, https://doi.org/10.3390/atmos7060077, 2016.

Groot Zwaaftink, C. D., Grythe, H., Skov, H., and Stohl, A.: Substantial contribution of northern high-latitude sources to mineral dust in the Arctic, J. Geophys. Res.-Atmos., 121, 13678–13697, https://doi.org/10.1002/2016jd025482, 2016.

Groot Zwaaftink, C. D., Arnalds, Ó., Dagsson-Waldhauserova, P., Eckhardt, S., Prospero, J. M., and Stohl, A.: Temporal and spatial variability of Icelandic dust emissions and atmospheric transport, Atmos. Chem. Phys., 17, 10865–10878, https://doi.org/10.5194/acp-17-10865-2017, 2017.

Meinander, O., Dagsson-Waldhauserova, P., Amosov, P., Aseyeva, E., Atkins, C., Baklanov, A., Baldo, C., Barr, S. L., Barzycka, B., Benning, L. G., Cvetkovic, B., Enchilik, P., Frolov, D.,

Gassó, S., Kandler, K., Kasimov, N., Kavan, J., King, J., Koroleva, T., Krupskaya, V., Kulmala, M., Kusiak, M., Lappalainen, H. K., Laska, M., Lasne, J., Lewandowski, M., Luks, B., McQuaid, J. B., Moroni, B., Murray, B., Möhler, O., Nawrot, A., Nickovic, S., O'Neill, N. T., Pejanovic, G., Popovicheva, O., Ranjbar, K., Romanias, M., Samonova, O., Sanchez-Marroquin, A., Schepanski, K., Semenkov, I., Sharapova, A., Shevnina, E., Shi, Z., Sofiev, M., Thevenet, F., Thorsteinsson, T., Timofeev, M., Umo, N. S., Uppstu, A., Urupina, D., Varga, G., Werner, T., Arnalds, O., and Vukovic Vimic, A.: Newly identified climatically and environmentally significant high-latitude dust sources, Atmos. Chem. Phys., 22, 11889–11930, https://doi.org/10.5194/acp-22-11889-2022, 2022.

7 Conclusion

In this deliverable report, numerous actions to improve online emissions of primary aerosols (desert dust, road dust, sea-salt aerosols) have been described. Machine learning techniques have been implemented for the first time as a part of the IFS-COMPO global CAMS system, which paves the way for the future work planned to represent dust emissions through ML processes. A new gridded dataset of non exhaust road emissions over Europe has been produced : the GNORTRIP dataset, and disseminated to CAMAERA partners. 4 models are currently testing the impact of this dataset on simulated PM; results with the EMEP model are encouraging in terms of improvement of the temporal correlation of PM10 over countries most impacted by road dust emissions. A new dust emission has been implemented into IFS-COMPO, adapted from the SILAM dust emission schemes, which brings significant improvements on the forecast of many dust-impacted variables. Specific developments have been tested on both the new and the operational dust emission schemes in order to improve the representation of high latitude dust sources. Evaluation is hard because of missing observations; a first assessment over Iceland shows a beneficial impact of these developments on the representation of local dust sources.

Document History

Version	Author(s)	Date	Changes
1.0	Rose-Cloé Meyer, Nathan Capon, Samuel Remy, Risto Hänninen, Elin Aas, Bruce R. Denby	27/06/2025	Initial version

Internal Review History

Internal Reviewers	Date	Comments