CAMS Service Evolution

D6.5 Impact of the uncertainty of the boundary conditions on regional forecasts

Due date of deliverable	31 December 2025
Submission date	16/12/25
File Name	CAMEO-D6-5-V1.0
Work Package /Task	WP6/Task6.4
Organisation Responsible of Deliverable	MET Norway
Author name(s)	Willem van Caspel, Hilde Fagerli, Anna Benedictow, Àlvaro Valdebenito, Charlie Negri, Samuel Rémy
Revision number	V1
Status	ISSUED
Dissemination Level	PUBLIC

Funded by the European Union

The CAMEO project (grant agreement No 101082125) 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

In the models contributing to the regional CAMS forecast services, boundary conditions (BCs, concentrations of chemical species at the borders of the European domain run by the regional models) are based on 12-hourly forecast fields from the global IFS-C model. In CAMEO D6.4, the impact of uncertainties in meteorology, initial conditions (meteorological and chemical), anthropogenic emissions, and model parameterizations were investigated by means of 51-member ensemble perturbation experiments with the global IFS-C model. In the current work (for CAMEO D6.5), results from these perturbation experiments are specified as BCs in diagnostic EMEP model air quality forecast simulations, to quantify the impact of uncertainties in the BCs concentrations on regional CAMS air quality forecasts.

The impact of using the ensemble perturbation experiments as BCs are evaluated for three air pollution episodes in the year 2021, focusing both on ozone (O_3) and particulate matter (PM) with aerodynamic diameters below 10 micrometers (PM_{10}). The model forecast skill during each episode is evaluated based on hourly surface O_3 and PM_{10} observations from around 1400 measurement stations in Europe. Domain-wide maps of the resulting BCs uncertainty impacts are also investigated.

We find that the combined meteorology, initial conditions, and emissions uncertainties (the IDFU ensemble experiment) make a larger impact on the forecast results than the uncertainties associated with model parameterizations (IK1T). However, the IDFU uncertainties make a modest impact on air quality forecasts in Europe as a whole, with the BCs influence having insufficient time to fully propagate across the regional modelling domain over the course of a forecasting period. Nevertheless, for countries close to the regional model boundaries (e.g., Ireland and Spain), the uncertainties associated with the IDFU experiment influence the distributions of O_3 and PM_{10} from the second forecasting day onwards. For these countries, BC uncertainties can induce inter-ensemble normalized mean model bias variations of around 20%.

Table of Contents

2 Introduction

2.1 Background

Monitoring the composition of the atmosphere is a key objective of the European Union's flagship Space programme Copernicus, with the Copernicus Atmosphere Monitoring Service (CAMS) providing free and continuous data and information on atmospheric composition.

The CAMS Service Evolution (CAMEO) project will enhance the quality and efficiency of the CAMS service and help CAMS to better respond to policy needs such as air pollution and greenhouse gases monitoring, the fulfilment of sustainable development goals, and sustainable and clean energy.

CAMEO will help prepare CAMS for the uptake of forthcoming satellite data, including Sentinel-4, -5 and 3MI, and advance the aerosol and trace gas data assimilation methods and inversion capacity of the global and regional CAMS production systems.

CAMEO will develop methods to provide uncertainty information about CAMS products, in particular for emissions, policy, solar radiation and deposition products in response to prominent requests from current CAMS users.

CAMEO will contribute to the medium- to long-term evolution of the CAMS production systems and products.

The transfer of developments from CAMEO into subsequent improvements of CAMS operational service elements is a main driver for the project and is the main pathway to impact for CAMEO.

The CAMEO consortium, led by ECMWF, the entity entrusted to operate CAMS, includes several CAMS partners thus allowing CAMEO developments to be carried out directly within the CAMS production systems and facilitating the transition of CAMEO results to future upgrades of the CAMS service.

This will maximise the impact and outcomes of CAMEO as it can make full use of the existing CAMS infrastructure for data sharing, data delivery and communication, thus supporting policymakers, business and citizens with enhanced atmospheric environmental information.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The objective of this deliverable is to quantify the impact of the uncertainty in the boundary conditions on regional-scale CAMS air quality (ozone and dust) forecasts in Europe. This is achieved using the boundary condition uncertainties derived as part of CAMEO D6.4 in combination with newly performed diagnostic EMEP air quality model simulations.

2.2.2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP6 Task 6.4.4) was performed.

2.2.3 Deviations and counter measures

No deviations have been encountered.

2.2.4 CAMEO Project Partners

ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
MET Norway	METEOROLOGISK INSTITUTT
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMi
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT
BIRA-IASB	INSTITUT ROYAL D'AERONOMIE SPATIALEDE BELGIQUE
HYGEOS	HYGEOS SARL
FMI	ILMATIETEEN LAITOS
DLR	DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV
ARMINES	ASSOCIATION POUR LA RECHERCHE ET LE DEVELOPPEMENT DES METHODES ET PROCESSUS INDUSTRIELS
CNRS	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS
GRASP-SAS	GENERALIZED RETRIEVAL OF ATMOSPHERE AND SURFACE PROPERTIES EN ABREGE GRASP
CU	UNIVERZITA KARLOVA
CEA	COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
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 Background information on boundary conditions and forecasting

CAMS delivers air quality forecasting products on a daily basis. For example for global and European air quality forecasting (https://atmosphere.copernicus.eu/) and European policy support purposes (https://policy.atmosphere.copernicus.eu/). The chemistry-transport model (CTM) developed at the European Monitoring and Evaluation Programme (EMEP) Meteorological Synthesizing Centre-West (MSC-W) is one of the CTMs participating in these forecasting services, with their quality forecast service relying on an ensemble of 11 participating models. The policy support forecasts rely on three participating models.

Of particular relevance are the forecasting services focusing on the European domain. For this domain, CTM boundary conditions (BCs) can exert a strong influence on relatively long-lived pollutants such as ozone (O_3) and particulate matter (PM). For O_3, having a tropospheric lifetime of a week to around a month (Prather and Zhu, 2024), the BCs effectively control the European influx of O_3 generated by North American precursor emissions, and to some extent downward transported stratospheric O_3. For European surface O_3 concentrations, the BCs typically contribute around 50-75%, with its contribution being highest during the winter months. For PM, the BCs control the inflow of natural emission sources such as sea salt and Saharan dust. The chemical composition and source contributions from the BCs (denoted as the 'natural' chemical component and 'hemispheric' source region, respectively) are also included in the CAMS service policy forecasts. These forecasts are in turn used to derive annual source and chemical composition analysis (https://policy.atmosphere.copernicus.eu/yearly/), in addition to air quality exceedance statistics based on the first day of each daily forecast over the course of the year.

In the default CAMS configuration, models participating in the ensemble forecasts employ BCs based on 12-hourly forecast fields provided from the ECMWF global Integrated Forecasting System (IFS) with chemical composition (IFS-C) model. These BCs are used to specify concentrations for sea salt, dust, forest fire, and sulphate (SO_4) aerosols, and for the gaseous O_3, carbon monoxide (CO), nitrogen oxides ($NO + NO_2$), sulphur dioxide (SO_2), and formaldehyde (CH_2O) species. Effectively all of these species either directly or indirectly affect European PM and O_3 concentrations.

Processes controlling the production and inflow of BC species into the regional modelling domain are, for example (but not limited to), generation by surface winds (sea salt, dust), transport by the prevailing winds and synoptic-scale weather patterns, loss through surface deposition (both for aerosols and gaseous species), and photochemical loss and production (O_3). In practically all CAMS models, including the IFS model on which the BCs are based, the aforementioned processes are included based on highly parameterized formulations of the underlying physical and chemical processes. In CAMEO D6.4, The impact of uncertainties in these parameterizations has been investigated by means of a number of ensemble perturbation IFS forecast simulations. Since these results translate also into BCs uncertainties, a selection of ensemble forecast runs are used in the current work in place of the standard IFS BCs, in combination with a diagnostic EMEP CAMS forecasting setup.

4 Experimental setup

4.1 Ensemble IFS perturbation simulations

Work in CAMEO WP6 task 4 (D6.4) focused on the quantification of uncertainty in global IFS-COMPO (Integrated Forecasting System with atmospheric composition extensions, or IFS-C in short) forecasts as propagated from uncertainties in meteorological parameters, emissions of pollutants, forecast initial conditions, and model parameterizations (CAMEO-D6-V1.pdf, last access: October 2025). To this end, the IFS-C was run for a number of ensemble perturbation experiments. Each of these experiments focused on different sets of perturbations to model input and state variables, being run for 50 ensemble members each. Each ensemble member was run for a 5-day forecasting period for every day of the year in a given time period (in the current work falling between January and September 2021). Key to the perturbation experiments is that each ensemble member is initialized from a different random realization of the perturbation fields in question.

The current work considers the IDFU and IK1T experiments. The IK1T experiment applies perturbations only to model parameterizations (referred to as 'MODEL' in D6.4 Table 3). The IDFU experiment includes perturbations to the meteorological tendencies, initial conditions, atmospheric composition, and emissions, but not to the model parameterizations (referred to as 'ALL_NOMODEL' in D6.4 Table 3). In the IDFU experiment, perturbations of the different components (meteorology, emissions, initial conditions) are assumed to be uncorrelated and constant for all forecast times. Details of the different perturbation mechanisms are described below.

- For the meteorological fields, perturbations were applied to the initial conditions and to the meteorological time tendencies following the usual procedure for the NWP ensemble, as described in the documentation (ifs-documentation-cy49r1). The atmospheric composition perturbations relate only to the perturbation of their initial conditions.
- For emission perturbations, 500 km length scale perturbations were applied on a per-sector basis (e.g., for the industry, residential heating, traffic sectors). The standard deviations of the perturbations were based on the maximum value of their respective diurnal variations, as determined in CAMEO work package 5. These were found to range from 0.063 for industry to 1.233 for wood burning (relative to a baseline emission intensity of 1), taken to be representative of their temporal uncertainties. Uncertainties related to emission activity data, spatial distribution and conversion factors (e.g., NO to NO_2 ratios) are not taken into account.
- Perturbations to the model parameterizations include perturbations to the dry and wet deposition rates, online (based on instantaneous meteorological fields) production rates of dust and sea salt aerosol, secondary organic and inorganic formation processes, tropospheric chemical (kinetic) reaction rates, and photolysis (*J*-value) reaction rates.

The perturbations applied (except for meteorology) were introduced as Gaussian variations in spherical coordinates with a horizontal correlation length scale between 250 and 2000 km. In the current work, perturbations with a length scale of 500 km are used, with the choice of length scale having little impact on the resulting ensemble spread. These perturbations are constant during the forecast time.

We note that the results of the perturbation experiments are only specified along the EMEP lateral boundaries, as the IFS-C concentrations would be in the operational forecasting setup. The meteorological perturbations therefore do not affect the IFS forecast simulation on which the meteorology used by the EMEP model is based.

4.2 EMEP MSC-W forecast configuration for CAMS products

The EMEP model configuration used in the current work largely follows that of the setup employed for the CAMS forecasting products. In this configuration, the model is run over the CAMS domain ([30°N-76°N] x [30°W-45°E]), employing 12-hourly chemical BCs from IFS-C in addition to 3-hourly meteorological fields (winds, temperature, clouds, etc.) from the IFS model (the latter not being involved in the perturbation experiments).

One difference is that the current work uses IFS meteorological data on a 0.1° by 0.1° latitude-longitude horizontal resolution rather than the 0.2° by 0.1° resolution used for CAMS forecasting. This choice is based on the availability of historical meteorological data, with the higher resolution not being expected to have any meaningful impact on the results discussed in this work. In addition, while the CAMS forecasting suite employs anthropogenic emissions from the CAMS-REG APv5.1 dataset, the current work employs default EMEP reported emissions. However, since the focus lies on the impact on BC uncertainties, the choice of emission dataset is not expected to impact the analysis.

By default, the EMEP model employs 20 vertical levels up to an altitude of 100 hPa, with a lowest layer approximately 45 meters thick. Annual air pollutant emission totals are distributed in time based on the CAMS-TEMPO timefactor dataset (Guevara et al., 2021; 2020). Natural emissions include BVOC, forest fires, volcanic eruptions, and soil-NO_x while sea salt and dust aerosols are generated online based on meteorological soil moisture and wind fields. Further details about the EMEP model are given in Simpson et al. (2012) and in the annual EMEP status reports (EMEP report 2025, last access: October 2025).

Each simulation (i.e., 5-day forecast) presented in the following is preceded by a 1.5 month (45 days) spin-up simulation using BCs from the unperturbed (baseline) ensemble member. As a result, each ensemble forecast begins with zero spread. The BCs uncertainties then propagate throughout the modelling domain as the forecast simulation progresses. The zero-spread start of each ensemble forecast mimics the behavior of the original ensemble simulations from CAMEO D6.4, in which each forecast also starts from the same initial condition (with the exception of those experiments in which the initial conditions are perturbed). Starting each ensemble forecast (both in the current work and in D6.4) from the same initial condition also relates to the operational CAMS European air quality forecast setup, where each forecast is initialized based on a data-assimilated model state. The latter includes both data-assimilation in the IFS model supplying the BCs and meteorological driver, and data-assimilation of the EMEP (and other participating regional model) states within the modelling domain (e.g., confluence.ecmwf.int/forecast-documentation). For the chemical species within the regional modelling domain, the data-assimilation includes surface observations of O_3, NO2, $PM_{2.5}$, $PM_{10,}$ CO and SO_2 from stations representative of measuring background air pollution levels. While data-assimilated models are by no means perfect with respect to observations, the data-assimilation does effectively suppress model error and uncertainty. The zero-spread start in each of the ensemble forecasts is therefore taken to be similar to the data-assimilated start of the CAMS European air quality forecast product.

However, for the CAMS policy support service forecasts, the initial conditions within the EMEP modelling domain are initialized based on the end-point of the first forecasting day from the previous day forecast. In this service, the sensitivities to emission reductions from individual source regions are tracked, with each being associated with their own initial conditions, making data-assimilation impossible. In this setup, BC uncertainties from the first ensemble perturbation BCs time-slice (12 hr ensemble forecast) enter the regional modelling domain and propagate throughout the entire analysis period (attributed to the ‘hemispheric’ source contribution). We note that the first ensemble BC perturbation time-slice is associated with no uncertainty, except for the perturbation simulations where the initial conditions are perturbed. As a result of the zero-spread start of the BCs, the impact of the BC uncertainties on the initial conditions of the hemispheric source contribution is effectively halved with respect to the BC uncertainties that would arise over time. Since one of the primary products of the CAMS policy support service is to provide country-and-city source contributions, the impact of BC uncertainties on the initial conditions of the hemispheric source contributions are not considered in the current work. However, the impact of BC uncertainties on the forecasts as discussed in the following, in principle applies to the uncertainty of the hemispheric source-contribution. The discussed uncertainty impacts likewise apply to the total simulated air pollutant concentrations calculated as part of the CAMS policy support service.

4.3 Ensemble evaluation

Following also the analysis of D6.4, the standard deviation (ensemble spread) at any given time and location within the modelling domain is taken as a measure of the uncertainty. The resulting standard deviation may additionally be divided by the ensemble mean to provide a measure of the relative spread.

Rank histograms are often used for the verification of ensemble forecasts (as discussed in detail in D6.4). Alternatively, model evaluation guidelines for regional CAMS air quality forecasting products were developed as part of CAMS2_83 based on observational data from the EEA-UTD (European Environment Agency - Up-To-Date) network. These data included measurements from stations classified as urban, suburban, and rural (Joly and Peuch, 2012). The guidelines and their implementation are described in more detail in D83.4.1.1 (last access: October 2025).

4.3.1 Aeroval performance statistics

The D83.4.1.1 guideline document forms the basis of the model ensemble evaluations shown on the Aeroval website (e.g., copernicus.eu/evaluation, last access: October 2025). In the context of the Aeroval evaluation, the ensemble refers to the suite of 11 models participating in the CAMS2_83 air quality forecasts. The focal points of these evaluations revolve around the construction of time series, Taylor diagrams, and median scores in relation to the EEA-UTD verification data. The number of EEA-UTD stations within the CAMS modelling domain typically amounts to around 1400. In the Aeroval interface, performance statistics can furthermore be selected for individual countries within the domain. Taking Germany as an example, the ensemble evaluation would then be based on around 235 stations within the country.

The time series and Taylor diagram analysis are used as measures of the model skill for the first forecast day (day-zero) over the course of any given time period under consideration. Since the ensemble simulations considered in the current work are based on the same CTM employing the same meteorological driver, and since it takes at least a day for the influence of BC perturbations to reach mainland Europe, the time series and Taylor diagram evaluations show identical results for the baseline and perturbation runs (i.e., there is no inter-model spread for the ensemble simulations based on their respective day-zero forecasts). The time series and Taylor diagram evaluation metrics are therefore not discussed in this report.

However, for the median scores metric, the model normalized mean bias, modified normalized mean bias, Pearson correlation statistic, fractional gross error, and root mean squared error are calculated for each forecast hour up to forecast hour 95 (i.e., first 4 forecasting days), such that the influence of the BCs uncertainties does become apparent. In the current work, the median score metrics for the often employed Pearson correlation (r) and normalized mean bias (NMB) statistics are discussed. The method behind the median score calculations is described below.

In the operational CAMS forecasting setup – and in the current work – each forecast starts at 00:00 UTC and is performed once daily. Considering a single EEA-UTD observational site and a single model (or ensemble member) for which 10 forecasts have been performed (for example for a time period spanning 10 days, for which a forecast was performed each day), results from each forecasting hour from each of the 10 forecasts can be compared to observations. Here, if the selected forecasting hour corresponds to 12 UTC, the observations would correspond to noon-time values sampled over the course of 10 days. The correlation and NMB are then calculated based on the co-located 10 forecast and observational data points, with the statistical scores reflecting the skill of the 12th forecasting hour in reproducing observations. This analysis is repeated for each forecasting hour between forecast hours 0 to 95. If for example the resulting correlation and NMB statistics generally perform better for night-time than for day-time forecasting hours, it reflects that the forecasts are more accurate for night-time than for day-time hours. And if the calculated statistics generally perform worse for later forecasting hours, it reflects that the forecasting skill degrades (or that the forecast becomes more uncertain) the further ahead the forecast is. The time series of the statistical scores are thereby taken to be representative of the temporal variations of the forecasting skill, as a function of forecasting hour.

While illustrated for a single observational site in the above, the analysis can be repeated for each EEA-UTD station within the modelling domain. The median of all individual station results is then what corresponds to the median score metric for any given model (or ensemble member). In the current work, median score evaluations are performed for the 51-member IDFU and IK1T experiments. These time periods for these evaluations focus on the date-ranges corresponding to the air pollution episodes described in the following.

4.4 Pollution episodes

While it is interesting to quantify the impact of the uncertainty in the BCs for any given time period, the uncertainties are arguably most important during air pollution episodes. Not in the least because such events are often marked by high numbers of air quality exceedances, to which hemispheric transport may be an important contributing factor. However, the results for air pollution episodes are expected to be relevant also for other time periods, with the episodes discussed in this section spanning both summer and winter time atmospheric conditions. Another reason for focusing on pollution episodes is that performing the diagnostic ensemble simulations is highly resource demanding.

An overview of air pollution events for the year 2021 is given in the Interim Assessment Report (IAR) policy support document that is published annually as part of the CAMS2_71 project (CAMS2_71_IAR_2021, last access: October 2025).

The above IAR highlights that 2021 was marked by an O_3 episode between the 14th to 21st of June, PM_{10} episodes between the 19th and 27th of February and the 12th and 24th of December, and a combined O_3 and PM_{10} episode between the 8th and 17th of August. The report further demonstrates that the February PM_{10} episode was caused by an inflow of Saharan air containing significant amounts of dust, contributing to the majority of exceedances of daily mean PM_{10} in winter over central Europe. The combined O_3 and PM_{10} episode in August was a more drawn out event, being marked by both wildfire contributions and Saharan dust intrusions over the European continent. The PM_{10} episode in December was to a large extent marked by stagnant and dry meteorological conditions in combination with high (local) residential fuel consumption (i.e., wood and coal combustion). Since the December episode was marked largely by local meteorological conditions and local emission sources rather than hemispheric transport, this episode is not considered here. For the other episodes, combinations of the IDFU and IK1T experiments are performed. An overview of the pollution episodes and the ensemble perturbation runs is given in Table 1. Note that not all episodes are covered by each of the ensemble runs in the interest of saving computational resources, although the simulations have been chosen such that each ensemble covers both a PM_{10} and O_3 episode.

Table 1. Air pollution episodes and IFS ensemble perturbation experiments for the year 2021.			
	February 19-27 PM_{10}	June 14-21 O_3	August 8-17 PM_{10} and O_3
IDFU (meteorology, initial conditions, temporal emissions)	**X**		**X**
IK1T (model parameterizations)	**X**	**X**	

5 General results

5.1 IDFU

This section highlights some of the general features of the ensemble experiments. To this end, Figure 1 illustrates the variability in surface PM_{10} concentrations for forecast day 1 to 4, here based on daily averaged results from the IDFU experiment for a forecast initialized for the 19th of February. The first day of a forecast is typically referred to as day zero. Day 1 shown in Figure 1 thereby refers to the daily average concentration over the second day of the forecasting period.

For PM_{10} the variations are largest near to large dust sources in northern Africa, with the uncertainty of sea salt (along the western boundary) having a much smaller impact. Figure 1 further shows that a north African dust plume is dispersed northwards between day 1 and 4, with uncertainties (standard deviations) of around 2-4 µg m^{-3} over parts of France, Germany, and Italy. Given the likeliness with the dust-plume transport pattern, the PM_{10} uncertainty appears to be dominated by dust (driven by initial conditions and meteorology), with the impact of (temporal) emission perturbations having a much smaller impact. Diagnostic analysis finds that these results also apply to forecasts in the summer period. For dust transport into the modeling domain, the northward dispersion patterns are strongly influenced by the geographical blocking caused by central European mountain regions (e.g., the Alps). Countries lying northward of the Alps, such as Poland and Czechia, therefore experience comparatively little PM_{10} uncertainty from BC dust. Similarly, the presence of the Pyrenees can limit the dust inflow into France, even though it is not impossible for high-altitude dust plumes to traverse the Pyrenees.

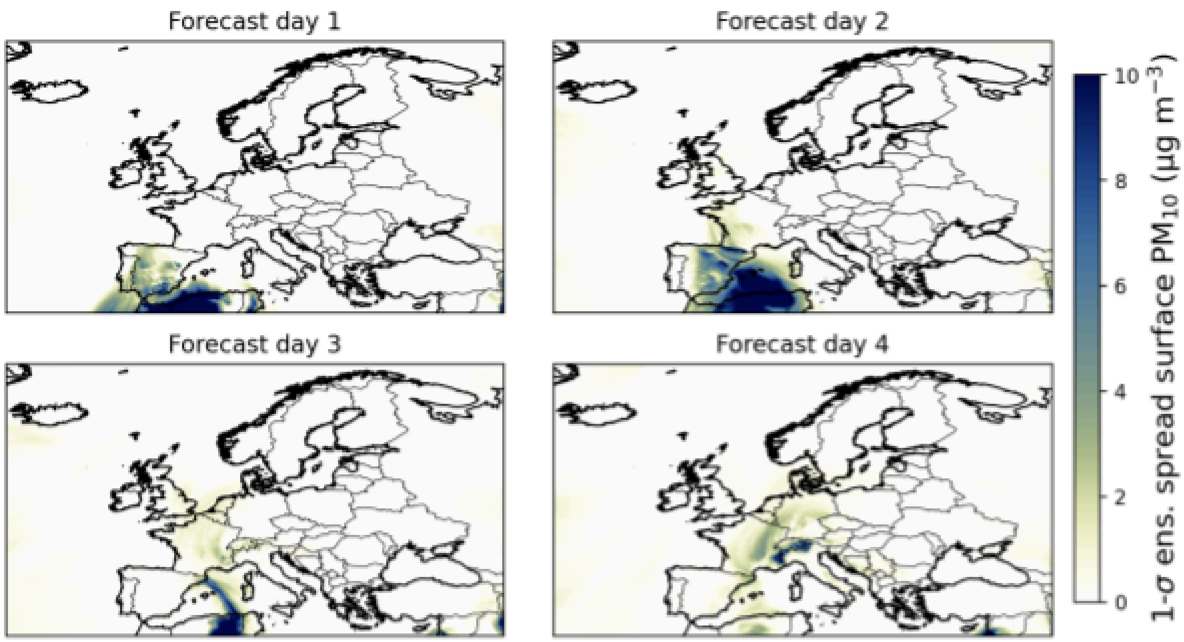

Figure 1. Ensemble spread in daily mean surface PM_{10} concentrations for the IDFU simulation, starting on the 19th of February with 1 to 4 days forecast times.

In Figure 2, the IDFU results for the 19th of February forecast are repeated for surface O_3 concentrations. Here the influence of the BC perturbations has by forecast day 1 reached the shores of western Europe (Ireland, parts of the United Kingdom, and Portugal). In addition, perturbations along the northern and eastern boundary sway out over northern Europe and Russia. These transport patterns are also common for other forecast days, being driven by the surface westerlies and high pressure systems over continental Europe deflecting air masses in a clock-wise manner. Owing to this land-sea pressure contrast, Atlantic transport patterns are often deflected northward, thereby acting as a barrier for easterly transport into central Europe. This transport pattern is further enforced by topographic features such as the Alps. As a result, the influence of the BC perturbations are more modest in central Europe than they are in the British Isles, at least on the time-scales relevant to the current work. While cold-air westerly intrusions are more common during wintertime, bringing in air masses from along the western model boundaries, O_3 inflow concentrations are typically lower from this direction (being influenced by the relatively more pristine Siberian air). In contrast, inflow from the eastern boundary conditions is comparatively higher due to the influence of O_3 produced downwind from North America.

Figure 2 demonstrates that the uncertainty (standard deviation) resulting from the meteorological, initial condition, and emission perturbations amounts to 15-20 μg m^{-3} for O_3 along the western coast of Europe and approximately 5 μg m^{-3} over more centrally located countries such as France and Germany. Further taking Ireland as an example, the uncertainty is highest on forecast day 2 rather than day 4. The latter highlights the importance of variations in transport patterns for countries closer to the boundary, being affected by mesoscale weather systems more so than by 'mixing' of the BC impacts within the modelling domain over time. While the February episode was identified as a dust episode, the O_3 results shown here are qualitatively similar to the results for the O_3 episode in August shown in Figure A1 of the appendix.

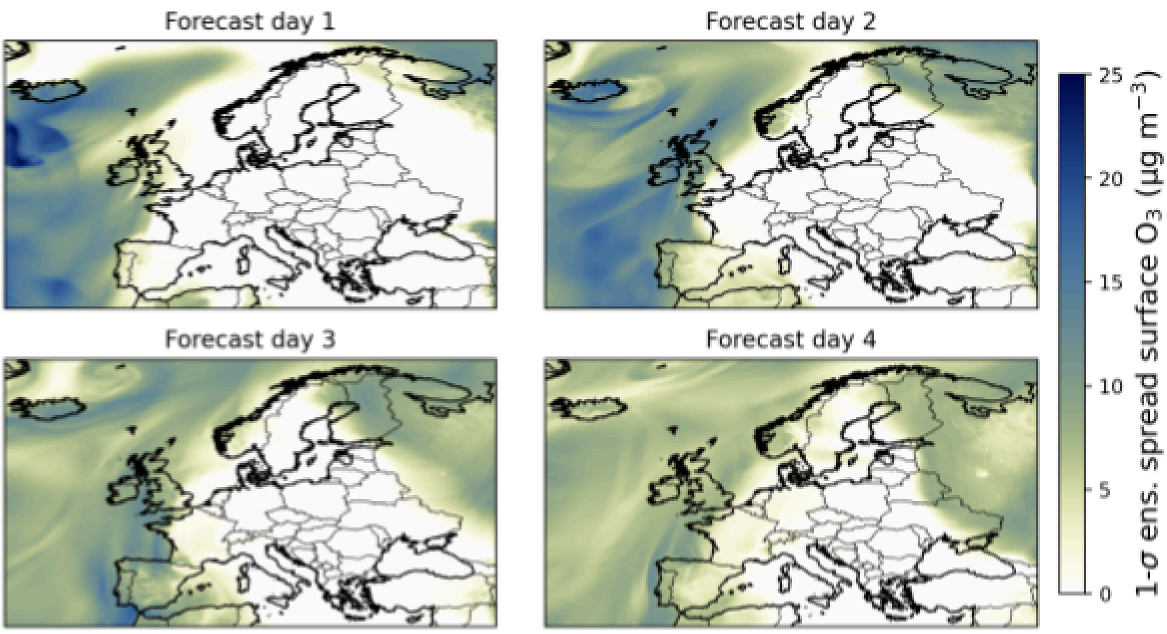

Figure 2. Ensemble spread in daily mean surface O_3 concentrations for the IDFU simulation, starting on the 19th of February with 1 to 4 days forecast times.

While the absolute uncertainties for PM_{10} concentrations are smaller than those for O_3 (at least over western Europe), Figure 3 shows that the relative uncertainties are often of a comparable magnitude over land areas. Thereby indicating that baseline concentrations for PM_{10} are lower than for O_3. The relative ensemble spread for O_3 can exceed 0.20 over the oceans and countries along the west coast of Europe. For more central countries, the O_3 relative spread reaches around 0.05-0.10. For PM_{10} the relative spread in central countries is slightly higher, being in the range of 0.10-0.15 in France and Germany.

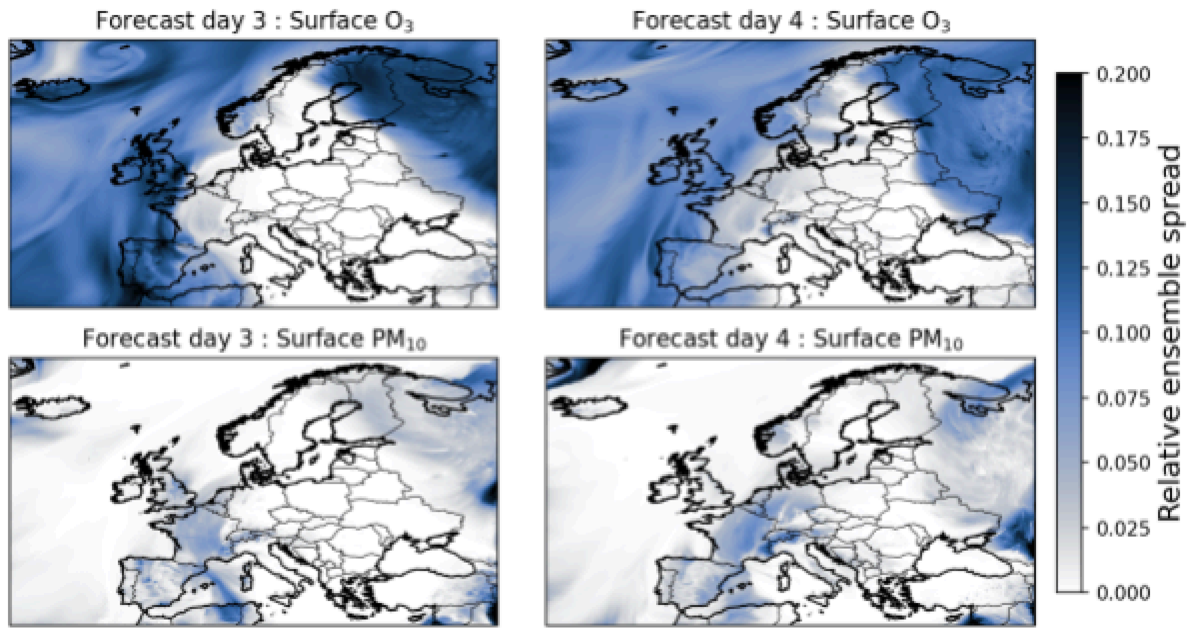

Figure 3. Relative uncertainty ensemble spread in daily mean surface O_3 (top panels) and PM_{10} (bottom panels) concentrations for the IDFU simulation, starting on the 19th of February with 3 and 4 days forecast times.

5.2 IK1T

This section highlights the general features of the IK1T experiment, focusing on the results for the 19th of February forecast as before. To that end, the relative uncertainties shown in Figure 4 can be directly compared with those shown in Figure 3 for the IDFU experiment. Figure 4 demonstrates that the IK1T uncertainties are generally smaller than those in the IDFU experiment. Namely, that the relative ensemble spread amounts to around 0.05 and 0.01 for PM_{10} and O_3 while for the IDFU experiment these fell in the range of 0.15-0.20 and 0.05-0.10, respectively.

For PM_{10}, the IK1T experiment highlights that the uncertainty in dust generation from its online production parameterizations is smaller than the uncertainty in its production resulting from meteorological variability (by comparison to the IDFU results). However, the difference between the IK1T and IDFU central European relative ensemble spread is smaller than one might expect based on the differences between their perturbations along the southern model boundary, which are considerably larger in the IDFU experiment. The latter may be indicative of the effects of changes in wet and dry deposition rates of dust (and PM_{10} as a whole) in the IK1T experiment. The relative spread in the IK1T experiment nevertheless remains modest, indicating that uncertainties in the model parameterizations have little impact on the forecast results. For reference, the relative ensemble spread (following Figure 4) is shown for the first forecast of the June episode in Figure A2 of the appendix. The latter finds similar results, in that the relative spread is low over Europe, being less than 0.05 for both O_3 and PM_{10}. However, while the relative spread is comparatively higher for PM_{10} over parts of Ukraine and Russia, the absolute spread is modest (3-5 µg m^{-3}).

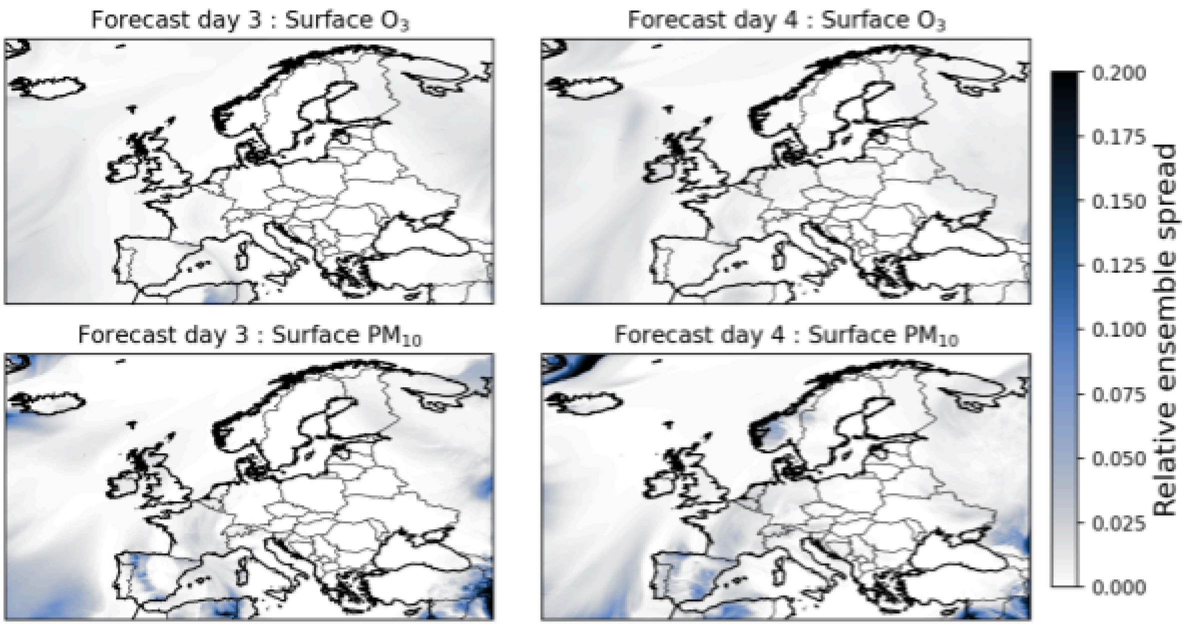

Figure 4. Relative uncertainty ensemble spread in daily mean surface O_3 (top panels) and PM_{10} (bottom panels) concentrations for the IK1T simulation, starting on the 19th of February with 3 and 4 days forecast times.

6 Ensemble performance evaluations

In this section the ensemble members are evaluated against EEA-UTD observations using the Aeroval evaluation interface. As before, the ensemble includes 50 perturbation members and 1 baseline simulation (the baseline simulation being labelled as member 00). However, in the previous section the IK1T BC perturbations were found to have a considerably smaller effect than those of the IDFU experiment. The following analysis therefore focuses on the results of the IDFU experiment. Note that in the following, hour zero corresponds to the start of the forecasts, i.e. the start of day-zero.

6.1 Surface O_3

6.1.1 European domain

Figure 5 shows the median IDFU forecasting skill for the Pearson correlation statistic and NMB (%) for the forecasts between the 8th and 17th of August 2021 (O_3 and PM_{10} episode; Table 1). The results shown here are based on all EEA-UTD data within the CAMS modelling domain. This figure illustrates that the ensemble spread is practically zero for the first two days of forecasting, indicative of the time that it takes for the BC perturbations to reach mainland Europe. By the 95th forecasting hour, the spread in the correlation between the best and worst performing ensemble members amounts to 0.09, representing a modest variation in the model skill. The correlation skill is further marked by pronounced diurnal variations with correlations being greatest during daytime, although the BCs have little impact on this. The effect of the BC is greatest during the night hours, probably due to the reduced influence of photochemical O_3 production by local emission sources. While the NMB reaches values as high as 40% during night-time hours, the overall ensemble spread remains low, with the largest NMB difference between ensemble members at forecast hour 95 amounting to 4%.

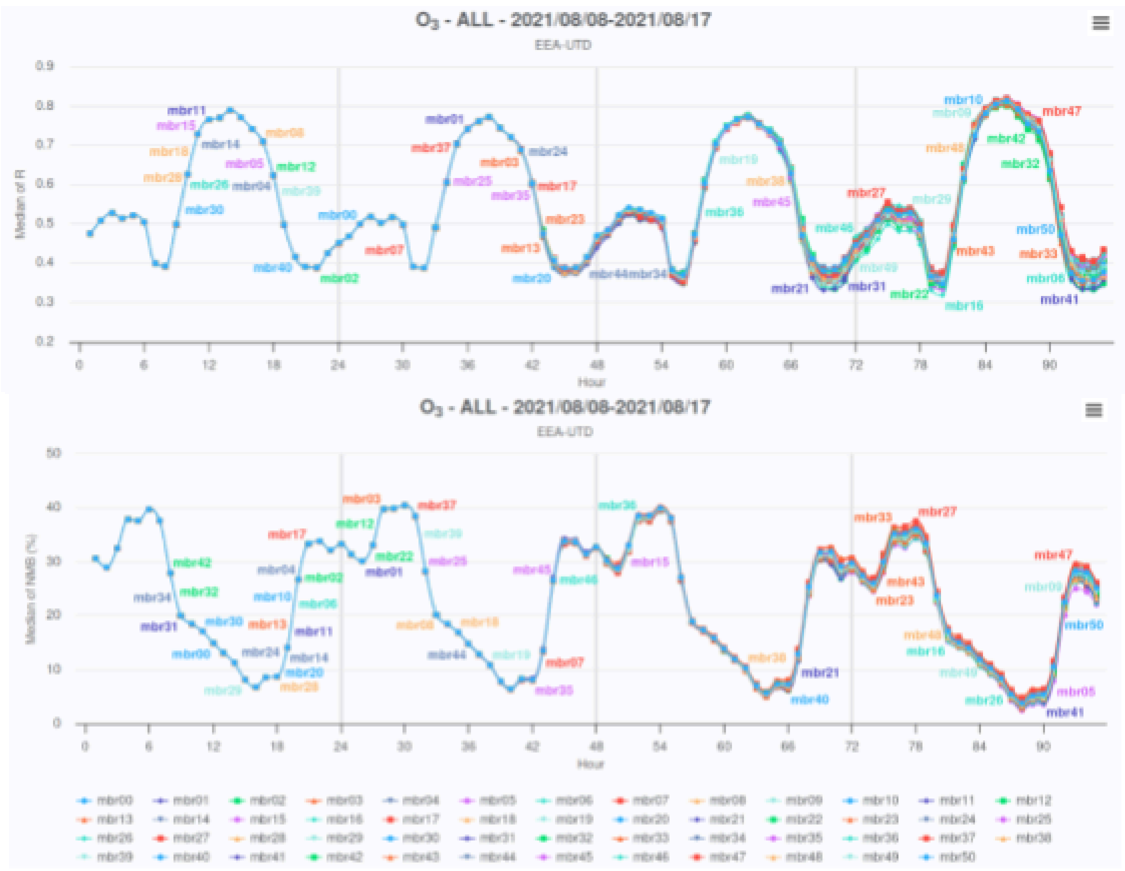

Figure 5. IDFU O_3 Pearson correlation (top) and NMB (bottom) forecasting skill for the August episode.

To illustrate the difference between episodes, and to some extent winter and summertime conditions, Figure 6 shows the IDFU correlation forecast skill for the February episode (although this is technically a PM_{10} episode). In this case, the spread in correlation is bigger, especially around hr 90 of the forecast. For the latter, the maximum spread amounts to 0.29. Interestingly, the diurnal cycle in correlation scores is effectively inverted between the February and August episodes. That is, for the February episodes the correlations are highest during the night and morning rather than during the day as in August. While a detailed analysis of why this occurs is beyond the scope of the current work, it can be noted that the inverted cycle happens mostly because day-time correlations are much poorer in February than they are in August (night-time correlations hover around 0.5-0.6 in both periods). For the February episode, the impact of the BCs on the correlations does not appear to be stronger or weaker depending on the time of day.

The generally greater ensemble-spread in the February results compared to the August results may be indicative of the geographical distribution of BC O_3 inflow playing a more important role in wintertime compared to in summer, when the spatial distribution of O_3 is to a larger extent driven by local photochemical production. Results for the NMB forecasting skill during the February episode further find a smaller ensemble spread than for the August episode (not shown here). Consistent with the results for O_3 shown in Figures 4, the impact on O_3 in the IK1T experiment is highly limited (negligible correlation and NMB impacts for both June and August).

Figure 6. IDFU O_3 Pearson correlation forecasting skill for the February episode.

6.1.2 Western Europe (Ireland)

In Section 5.1, the O_3 ensemble spread was found to be by far the greatest along the western coast of Europe (Ireland, UK, Portugal, Spain). To highlight the impact of the BCs for this region, the forecast performance for Ireland is discussed here. The results for Ireland are qualitatively similar to those for other countries along the western coast. For Ireland, the ensemble evaluation is based on 11 EEA-UTD stations.

Figure 7 shows the correlation and NMB statistics for the IDFU experiment for the February episode. This shows that there is a considerably greater ensemble spread in Ireland than in the rest of Europe (by comparison to Figure 5). While the variations in the NMB are still modest, with variations amounting to a roughly 20% spread by the 95th forecast hour, variations in the correlation are much greater. For the correlations, scores rapidly diverge from around the 24th forecast hour onward (indicative of the close proximity of Ireland to the western lateral model boundary). For the 48th forecast hour, 24 members have correlations below 0.5, while for the 72th forecast hour 7 members even show negative correlations. The correlation and NMB results together indicate that the ensemble perturbations impact the total amount of O_3 inflow into Ireland relatively little, but that the spatial distribution of the inflow is markedly different from member to member. This is likely the result of changes in weather pattern positions in the different ensemble members. Such misaligned transport patterns (compared to the baseline simulation, whose correlation stays above 0.52 for all forecast hours, with an average of 0.75) in turn lead to a severe degradation of the correlation scores (the correlation of the large majority of the ensemble members falls below that of the baseline simulation).

Figure A3 shows the evaluation for Ireland for the August episode, noting that the results shown in this figure are qualitatively similar to those shown in Figure 7, even though here the baseline simulation performs markedly worse for the first 24 hrs. For the August (summer) and February (winter) episodes, both results indicate that the ensemble uncertainty for total O_3 concentrations is modest. That is, the IDFU uncertainties redistribute O_3 more so than they increase or decrease the total amount of BC inflow.

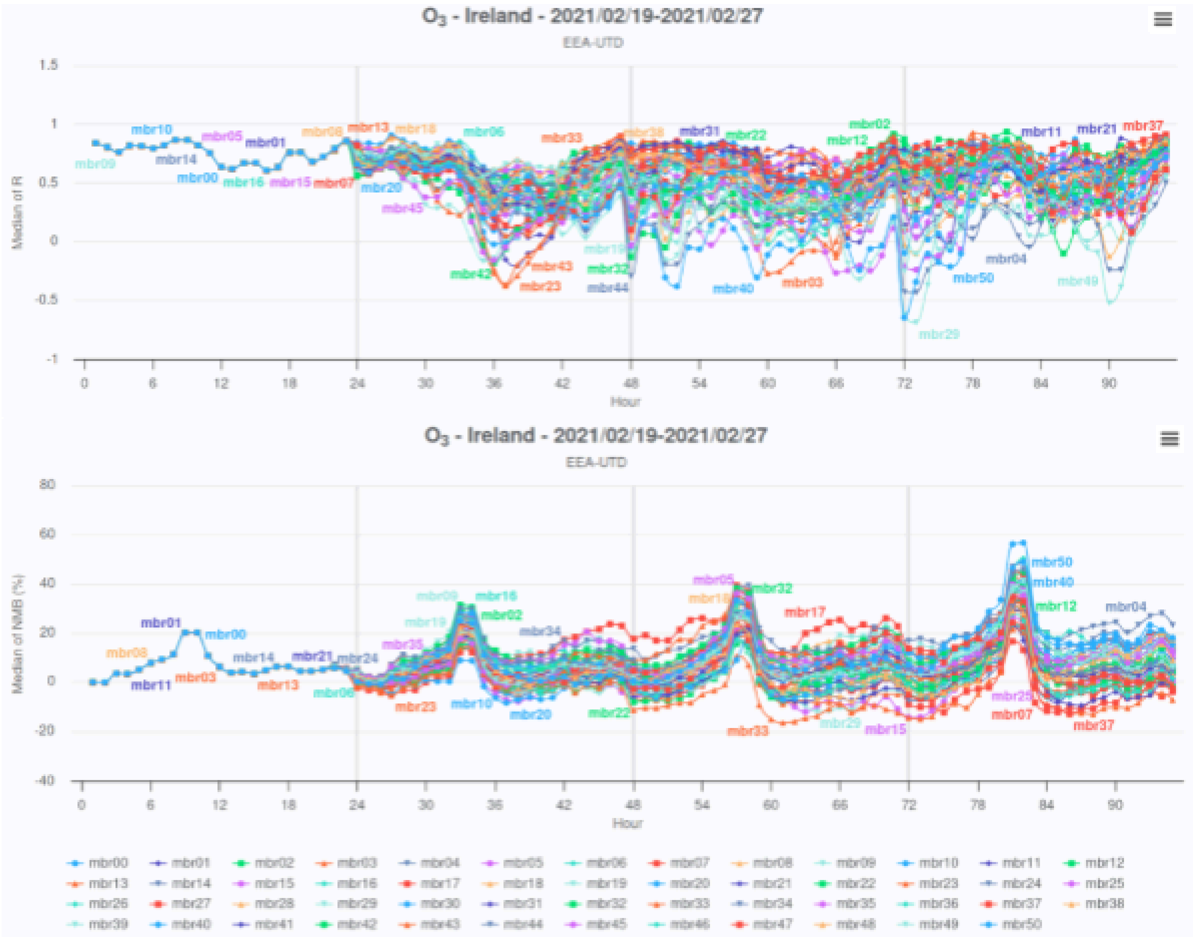

Figure 7. IDFU O_3 Pearson correlation (top) and NMB (bottom) forecasting skill for the February episode in Ireland (11 stations).

6.2 Surface PM_{10}

6.2.1 European domain

This section focuses on the PM_{10} episodes, coinciding with Saharan dust intrusions. Following the analysis of the previous section, Figure 8 shows the PM_{10} evaluation for the August episode. In this evaluation, being based on all EEA-UTD station data within the modelling domain, the ensemble spread in both the correlation and NMB statistics is small. This is in line with the ensemble spread shown in Figure A4 for four day-4 forecasts spanning the period of the August episode, indicating that the effects on PM_{10} is mostly contained to regions along the southern model boundary (consistent with the results shown in Figure 1 for the February episode). In particular, the most pronounced PM_{10} variations for the August episode are found in Spain, resulting from Saharan dust intrusions. While not shown explicitly, the overall (European) PM_{10} evaluation for the February episode also finds only small variations in the correlation and NMB statistics.

Figure 8. IDFU PM_{10} Pearson correlation (top) and NMB (bottom) forecasting skill for the August episode.

6.2.2 Southern Europe (Spain)

Given that the BC impacts are strongest along the southern edge of the modelling domain, in this section the ensemble evaluation focuses on the results for Spain. For Spain, the evaluation is based on 60 EEA-UTD stations. Figure 9 shows the corresponding evaluation for the correlation and NMB statistics. During the dust episode, the correlation performance sees an overall downward trend, driven mostly by poor night-time model performance. The ensemble perturbation experiments effectively follow this trend, representing variations around the baseline simulation with no particular predisposition for influencing either the night-time or day-time model performance. The overall model spread in the correlation statistic also remains modest, with the difference between the poorest and best performing ensemble members being around 0.1. Likewise, the NMB statistic is influenced relatively little, showing an inter-ensemble spread of about 10%.

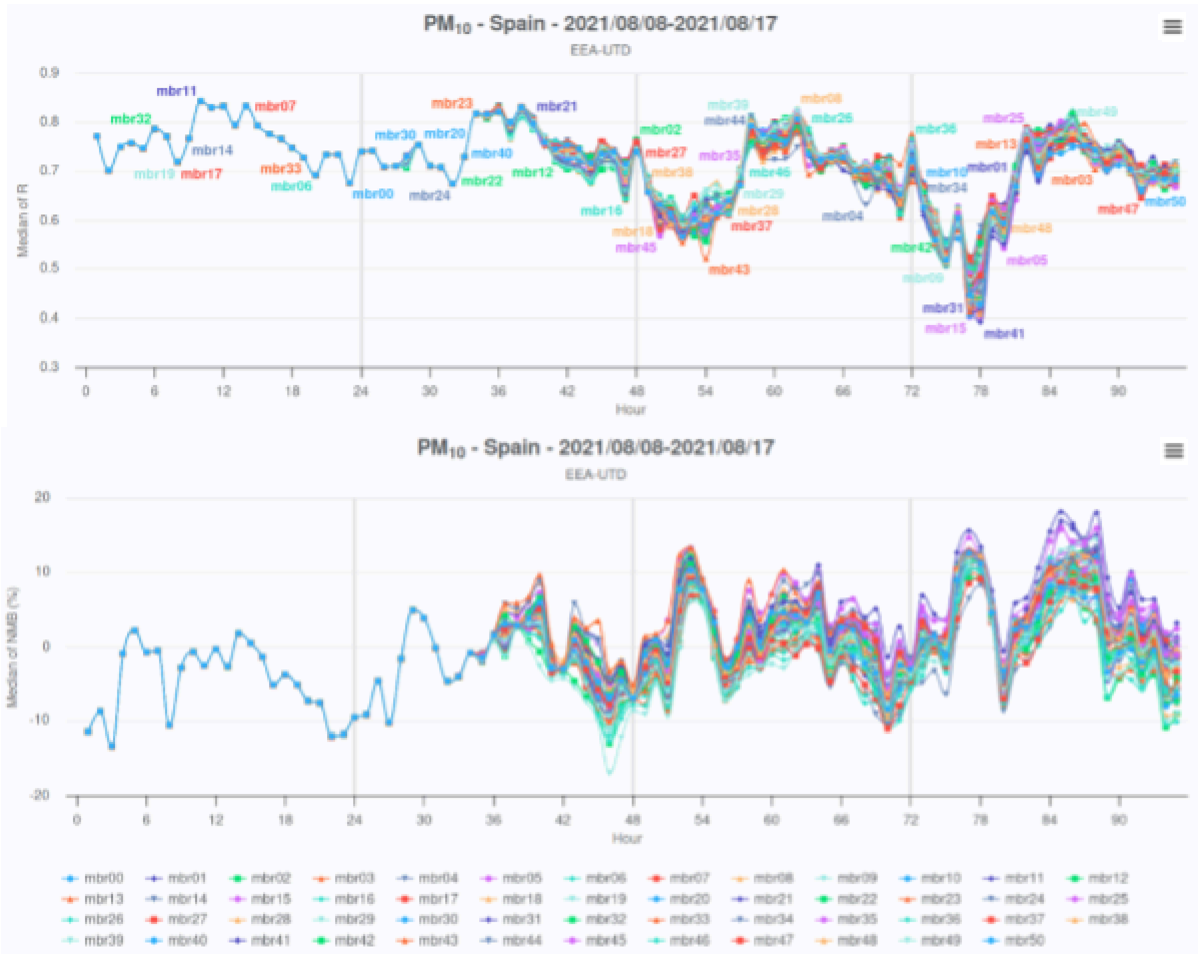

Figure 9. IDFU PM_{10} Pearson correlation (top) and NMB (bottom) forecasting skill for the August episode in Spain (60 stations).

To highlight that the results for the August episode are qualitatively similar also to other dust episodes and (winter) time periods, Figure 10 shows the correlation evaluation for the February dust episode. In this evaluation, the influence of BC perturbations is more pronounced from an early forecasting hour onwards (from around hr 30 compared to around hr 36 for the August episode). However, the BC influence remains relatively stable over time, also here amounting to an inter-ensemble variation in the correlation statistics of around 0.1 at most.

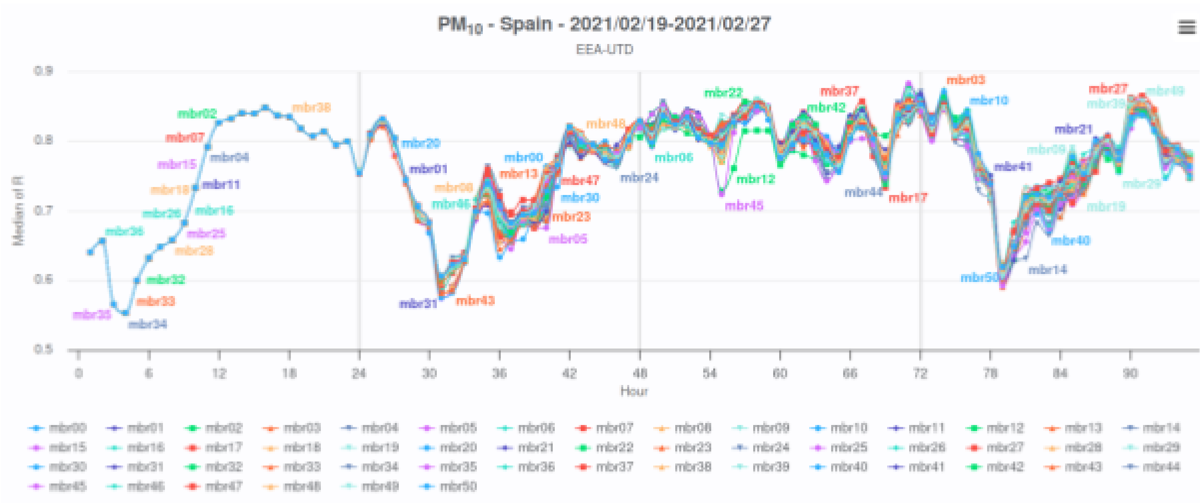

Figure 10. IDFU PM_{10} Pearson correlation forecasting skill for the February episode in Spain (60 stations).

7 Conclusion

This work investigated the effects of BC uncertainties on European air quality forecasts provided by CAMS services. To this end, results from global IFS ensemble perturbation experiments focusing on the impact of meteorological, initial condition, and chemical composition uncertainties (IDFU) and model parameterization uncertainties (IK1T) were specified along the model boundary edges of the regional CAMS forecasting domain in the EMEP model. Diagnostic EMEP model forecast simulations were performed to investigate the resulting impact on surface O_3 and PM_{10} within the modelling domain, focusing on the results for three air pollution episodes in the year 2021.

In general, the impacts of BC uncertainties were greater for the IDFU-based experiments, where the meteorology, initial conditions, and chemical composition were perturbed, when compared to the IK1T experiment, where chemical parameterizations (e.g., photolysis and deposition rates) were perturbed. The impact of BC uncertainties was further found to become most apparent around the final diagnostic forecasting days, based upon forecasting skill evaluations using EEA-UTD surface observations. The modest impacts of the BC uncertainties prior to the final forecasting days (day 4 onwards) effectively results from uncertainties that are introduced at the model boundaries not fully reaching (central) Europe over the course of a forecasting period. As a result, the forecasting skill as evaluated in Europe as a whole finds an overall small impact of BC uncertainties, expressed as minor variations in the correlation and NMB forecasting skill.

However, the influence of BC uncertainties can exert a stronger influence for countries closer to the model boundaries. This is most strongly expressed for O_3 for countries along the western coast of Europe and for PM_{10} (dust) for countries along the southern border of Europe. For both O_3 and PM_{10} within these respective regions, the BC uncertainties most strongly affect the evaluation of the correlation statistic, with inter-ensemble variations in the NMB forecasting skill remaining below around 20% (while being considerably less for countries not along the border of the modelling domain). The latter indicates that uncertainties in the inflow patterns play a more prominent role than uncertainties in the total amount of air pollutant inflow. Especially for O_3 in countries such as Ireland, the BC uncertainty in the final forecasting days can strongly affect the forecasting correlation skill, with median correlation scores ranging from roughly -0.5 to 0.8. Nevertheless, also for Ireland and for other countries along the western coast of Europe, inter-ensemble variations in the NMB forecasting skill due to BC uncertainties remain modest. If the IFS-C model would therefore be set up to provide ensemble uncertainty estimates of the BCs in an operational setting, we do not expect this to provide a significant improvement to the quality of the CAMS air quality forecasts.

It should be noted, however, that this does not mean that uncertainties in BCs in general are entirely unimportant for European air quality forecasting. The setup we have used here reset uncertainties in the European domain to zero at the start of every forecast, which in reality is exactly not the case. Even if the forecast start from the chemical data assimilation (DA) run, providing the best possible initial state, the DA is mostly affecting the European mainland (as that is where the surface stations are located), and for instance the air pollution concentrations for sea areas west of European mainland would be close to the concentrations from the deterministic regional scale model run. However, it would require a much more extensive setup to take the uncertainties in the areas that are not corrected by DA into account. It would in theory be possible to assess how large this uncertainty is by doing DA for each of the regional scale European runs forced with BCs from one ensemble member, and combining those runs with ensemble BCs, but it would be very CPU costly and far from something that could be implemented operationally.

8 References

Prather, M. J., & Zhu, X. Lifetimes and timescales of tropospheric ozone. *Elem Sci Anth*, *12*(1), 00112. https://doi.org/10.1525/elementa.2023.00112, 2024.

Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C., and Pérez García-Pando, C.: Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-TEMPO): global and European emission temporal profile maps for atmospheric chemistry modelling, Earth Syst. Sci. Data, 13, 367–404, https://doi.org/10.5194/essd-13-367-2021, 2021.

Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C., and Pérez García-Pando, C. Copernicus Atmosphere Monitoring Service TEMPOral profiles for the Global domain version 2.1 (CAMS-GLOB-TEMPOv2.1) **Copernicus Atmosphere Monitoring Service [publisher]** ECCAD [distributor], *https://doi.org/10.24380/ks45-9147,* 2020.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys., 12, 7825–7865, https://doi.org/10.5194/acp-12-7825-2012, 2012.

Joly, M., & Peuch, V. H. (2012). Objective classification of air quality monitoring sites over Europe. *Atmospheric Environment*, *47*, 111-123. https://doi.org/10.1016/j.atmosenv.2011.11.025

Appendix

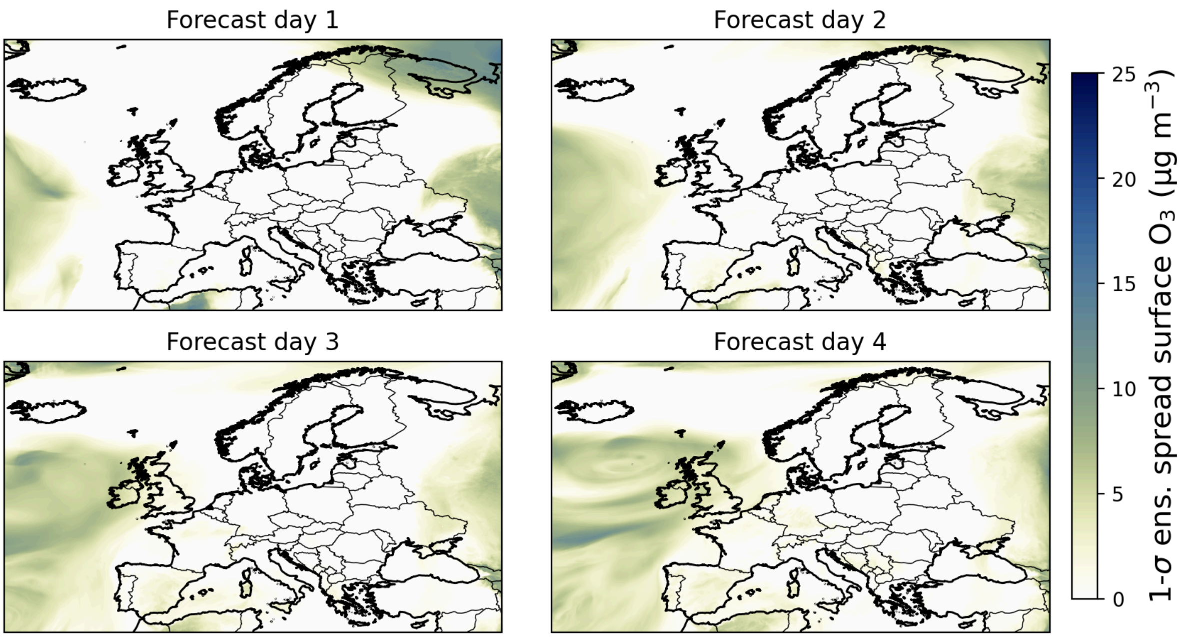

Figure A1. Ensemble spread in daily mean surface O_3 concentrations for the IDFU simulation, starting on the 8th of August with 1 to 4 days forecast times.

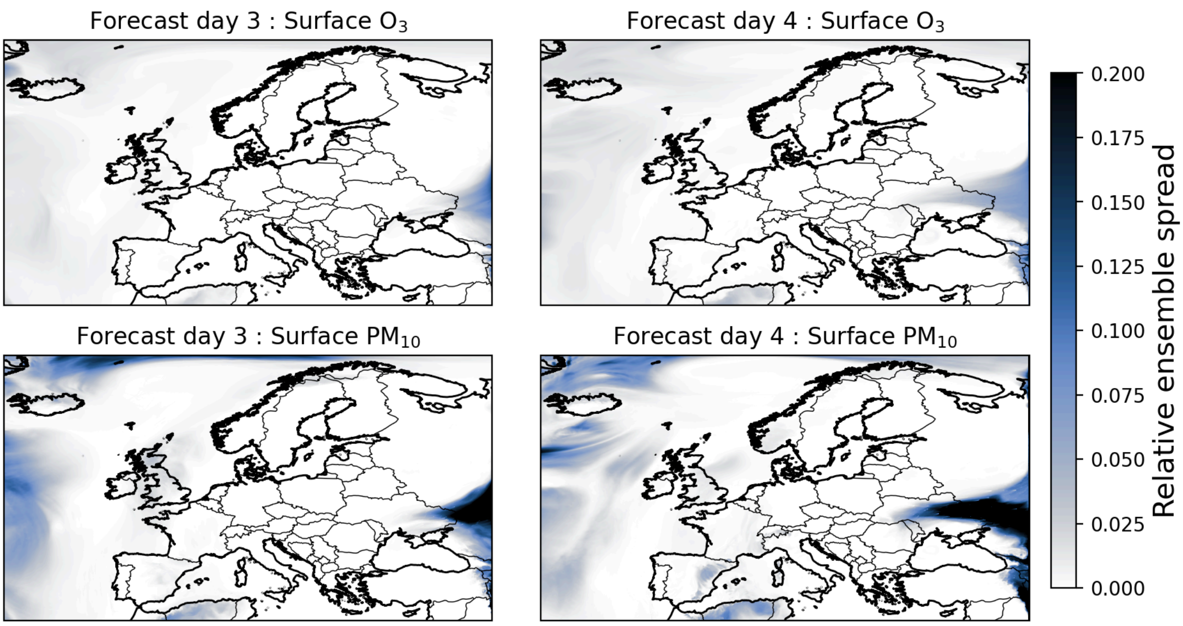

Figure A2. Relative uncertainty ensemble spread in daily mean surface O_3 (top panels) and PM_{10} (bottom panels) concentrations for the IK1T simulation, starting on the 14th of June with 3 and 4 days forecast times.

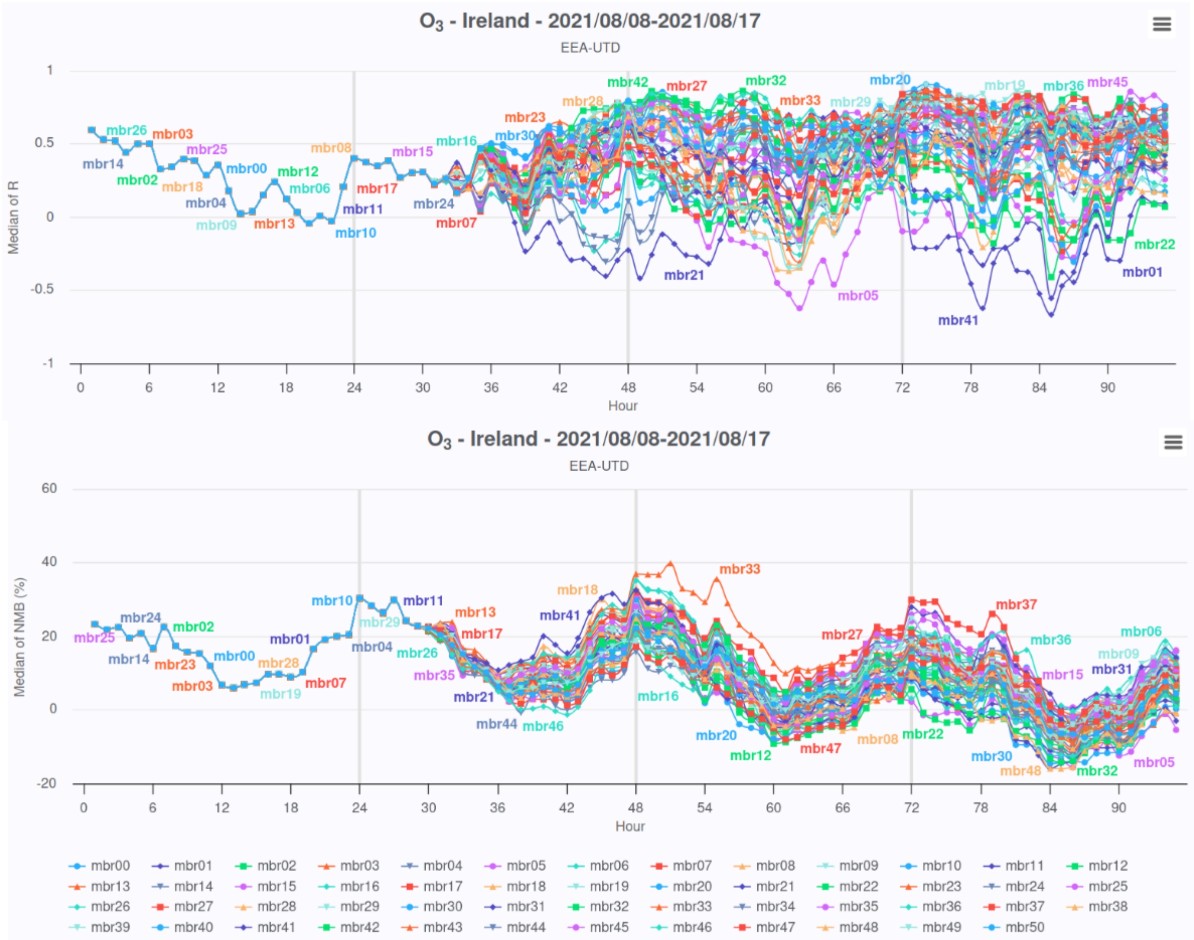

Figure A3. IDFU O_3 Pearson correlation (top) and NMB (bottom) forecasting skill for the August episode in Ireland (11 stations).

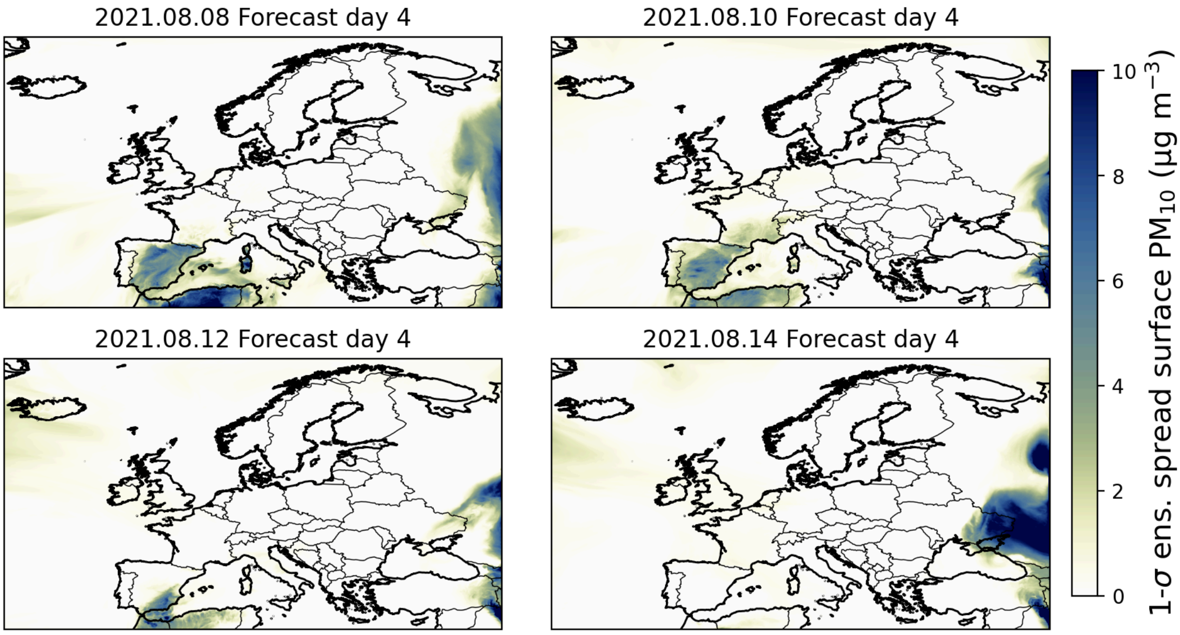

Figure A4. 4th forecasting day ensemble spread in daily mean surface PM_{10} concentrations for the IDFU simulation, for forecast days spanning the August air pollution episode (dates in panel titles).

Document History

Version	Author(s)	Date	Changes
0.1	Willem van Caspel, Hilde Fagerli, Anna Benedictow, Àlvaro Valdebenito, Charlie Negri, Samuel Rémy	05.12.2025	Initial version
1.0	Willem van Caspel, Hilde Fagerli, Anna Benedictow, Àlvaro Valdebenito, Charlie Negri, Samuel Rémy	Dec 2025	Updated and issued after internal review.

Internal Review History

Internal Reviewers	Date	Comments
Zhuyun Ye (AU)	11.12.2025	The report is well-structured and achieves its objectives with appropriately modest conclusions. Some minor edits and comments are suggested.
Noureddine Semane (AU)	12.12.2025	This report clearly analyses how boundary-condition uncertainties influence CAMS regional air-quality forecasts. The methods are robust, indicating mainly minor impacts across Europe, with larger effects near the edges of the domain, especially for O_3 in the west and PM_{10} in the south. The discussion of limitations and practical implications is clear and well-balanced.