



# Sensitivity of the *CHIMERE* model to changes in model resolution and chemistry over the northwestern Iberian Peninsula

Swen Brands<sup>1</sup>, Guillermo Fernández-García<sup>1</sup>, Marcos Tesouro Montecelo<sup>1</sup>, Nuria Gallego Fernández<sup>2</sup>, Anthony David Saunders Estévez<sup>2</sup>, Pablo Enrique Carracedo García<sup>1</sup>, Anabela Neto Venancio<sup>1</sup>, Pedro Melo Da Costa<sup>1</sup>, Paula Costa Tomé<sup>2</sup>, Christina Otero<sup>2</sup>, María Luz Macho<sup>1</sup>, and Juan Taboada<sup>1</sup> <sup>1</sup>MeteoGalicia - Consellería de Medio Ambiente, Territorio e Vivenda, Xunta de Galicia, Santiago de Compostela, Spain <sup>2</sup>Servicio de Calidad del Aire - Consellería de Medio Ambiente, Territorio e Vivenda, Xunta de Galicia, Santiago de Compostela, Spain

Correspondence: Swen Brands (swen.brands@gmail.com)

**Abstract.** Here, the capability of the chemical weather forecasting model CHIMERE (version 2017r4) to reproduce summertime surface ozone, particulate matter and nitrogen dioxide concentrations in complex terrain is investigated. The study area is the northwestern Iberian Peninsula, where both coastal and mountain climates can be found in direct vicinity and a large fraction of the land area is covered by forests. Fed by lateral boundary conditions from the ECMWF Composition Integrated

- 5 Forecast System, meteorological data from the Weather Research and Forecasting Model (WRF) and the HTAP v2.2 emission inventory, CHIMERE's performance compared to observations is tested with a range of sensitivity experiments, exploring the role of horizontal and vertical resolution and the effects of applying distinct chemistry mechanisms. Using a high horizontal *and* vertical resolution yields the most balanced verification results. If both the daily maximum and minimum values are important for the given application, then the full Melchior mechanism should be used. If, however, the daily maxima are considered more
- 10 important than the minima, SAPRC should be used instead. In any case, model performance for nitrogen dioxide is clearly not satisfactory for our study region, probably indicating deficiencies in the emission inventory.

# 1 Introduction

15

Motivated by the air quality legislation of the European Union (EU, 2008), many governmental air quality departments are currently demanding air quality forecasting schemes based on numerical models (Thunis et al., 2016), and the need for accurate and computationally efficient predictions in this field is perhaps greatest than ever before. For Europe as a whole, the most important real-time prediction system available to date is provided by the Copernicus Atmosphere Monitoring Service (Marécal et al., 2015), comprising an ensemble of currently seven chemical weather forecasting (CWF) models<sup>1</sup> run for the entire continent at a horizontal resolution of  $0.1^{\circ}$ . to  $0.25^{\circ}$  in longitude and  $0.1^{\circ}$  to  $0.2^{\circ}$  in latitude. In addition to this short-term

<sup>&</sup>lt;sup>1</sup>see Kukkonen et al. (2012) for an overview of these models





prediction system, several large research initiatives have been issued during the last two decades in order to assess the *cli-matological* properties of atmospheric composition, including the assessment of long-term tendencies resulting from emission reductions induced by the Convention on Long Range Transboundary Air Pollution (CLRTAP). The final aim of these efforts is to find model configurations, or ensembles thereof, that can be used as surrogates for real observations in order to assess whether emission reductions actually have lead, or would lead, to changes in the atmosphere's composition on climatological

5 whether emission reductions actually have lead, or would lead, to changes in the atmosphere's composition on climatological time-scales (Vautard et al., 2006; Jonson et al., 2006; Colette et al., 2011; Wilson et al., 2012; Banzhaf et al., 2015; Colette et al., 2017; Im et al., 2018b, a; Vivanco et al., 2018; Theobald et al., 2019).

Complementary to the aforementioned large-scale efforts, usually conducted with a single configuration of a given model, small-scale sensitivity tests for particular models are still relevant since they can be run with more sophisticated model con-

- 10 figurations than their large-scale counterparts and are therefore more interesting for regional prediction systems, such as those demanded by national or regional governments (Banzhaf et al., 2012; Beegum et al., 2016; Flamant et al., 2018). Further, following the concept of seamless prediction (Palmer et al., 2008), lessons learned from short-term prediction systems for relatively small geographical areas might as well be important for longer lead-times and larger areas.
- Previous sensitivity studies have identified several factors influencing the models' capability to correctly reproduce observed values, hereafter referred to as "model performance" (Giorgi and Francisco, 2000; Chang and Hanna, 2004). Among these factors, the meteorological input used to drive these models and the accuracy of the underlying emission inventory play a key role and have been assessed in a number of studies (Menut, 2008; Markakis et al., 2015; Colette et al., 2017; Otero et al., 2018; Vivanco et al., 2018). The resolution of the model mesh used to discretize the chemical reactions and atmospheric dynamics is also important and, when it is increased, a trade-off between potential performance gains and computational cost
- 20 must been made in practice. In what concerns the *horizontal* resolution, performance gains have been reported up to a scale of approximately 12 km for a number of models, such as WRF-CHEM and CHIMERE (Valari and Menut, 2008; Schaap et al., 2015; Crippa et al., 2017). However, a further resolution increase does not guarantee further performance gains. Namely, beyond the 12 km threshold, Misenis and Zhang (2010) reported heterogeneous results for WRF-CHEM that strongly depend on the considered time period. For the use of CHIMERE and focussing on surface O<sub>3</sub> concentrations, Valari and Menut (2008)
- 25 even found a performance loss which they attributed to a noise increase in the emission fluxes and meteorological parameters at higher resolution. Regarding the role of *vertical* resolution, an increase therein has been found to to improve the modelled particulate matter (PM) concentrations during desert dust events when using WRF-CHEM (Teixeira et al., 2016). However, CHIMERE's performance was found to be only weakly affected by such an resolution increase (Menut et al., 2013; Markakis et al., 2015).
- 30 Representing the number and complexity of the considered chemical reactions, several *chemistry mechanisms* are usually available for a given model and switching from one mechanism to another can also affect the model's performance. Model sensitivity to this parameter has been found to be appreciable for WRF-CHEM (Balzarini et al., 2015; Karlický et al., 2017). In recent CHIMERE versions, the SAPRC-07A mechanism (hereafter: SAPRC) has been included as an alternative to the full or reduced versions of the Melchior mechanism (Mailler et al., 2017) but, to the authors' knowledge, related sensitivity tests
- 35 are sparse to date.





A common shortcoming of small-scale sensitivity studies is that the conclusions therein only hold for specific regions, time periods or seasons of the year. In this context, most of the aforementioned conclusions for CHIMERE (the model applied here) have been drawn for the *Île de France* region, which is densely populated, relatively flat and not directly influenced by sea-salt emissions. The model has been applied for a number of other regions, but the map is still incomplete and sensitivity testing is not the main focus of the corresponding studies (Mazzeo et al., 2018; Menut et al., 2018; Monteiro et al., 2018; Brasseur et al.,

5

2019; Deroubaix et al., 2019).

This is where the present study comes into play: A series of sensitivity test has been run with CHIMERE over the *northwestern Iberian Peninsula*, a region characterized by a complex coastline, forested mountain terrain and the advection of sea-salt from the surrounding Atlantic Ocean. Thus, our study region is quite different from the *Île de France* region. The applied tests

- 10 will isolate the effects of an increase in the model's horizontal and/or vertical resolution, as well the effects arising from a the use of distinct chemical mechanisms (full Melchior or SAPRC). To this end, version 2017r4 of the CHIMERE model is used (Mailler et al., 2017) in combination with the HTAP emission inventory version 2.2 (Janssens-Maenhout et al., 2015). Saharan dust intrusions are *not* accounted for by running CHIMERE on a large domain covering all relevant dust sources (Bessagnet et al., 2017), but by using a far smaller domain ingesting the global forecasts provided by the European Centre for Medium-
- 15 Range Weather Forecasts (ECMWF) Composition Integrated Forecasting system (C-IFS) at its lateral boundaries (Flemming et al., 2015). This strategy largely reduces the computational costs and is an interesting alternative to simulating remote mineral dust emissions, e.g. originating in the Sahara desert, with the CHIMERE model itself (Bessagnet et al., 2017).

In Section 2, the applied data and model configurations used for sensitivity testing are described. Results are presented in Section 3 and a discussion and some general conclusions are provided in Section 4.

# 20 2 Data and Methods

This section opens with a description of the meteorological input data and the general characteristics of the CHIMERE experiments. Then, the particularities of the individual experiments are presented and the in-situ station network used for as reference for verification is introduced. The section closes with a description of the verification measures used to estimate CHIMERE's performance.

# 25 2.1 Meteorological Input and General Characteristics of the CHIMERE Experiments

The meteorological input data for the CHIMERE experiments are provided by the Weather Research and Forecasting (WRF) model version 3.5 (Skamarock et al., 2008), driven by Global Forecast System (GFS) forecasts initialized at 00 UTC (Caplan et al., 1997). WRF is run on three domains, a continental-scale domain having a resolution of 36km, followed by a regional domain covering southwestern Europe at a resolution of 12km and, finally, a 4km domain covering our study region, the

30 northwestern Iberian Peninsula. For these domains, WRF is executed with a minimum time step of 216, 72 and 24 seconds and a maximum time step of 360, 180 and 60 seconds, respectively. All domains comprise 33 vertical layers with a model top at 10 hPa. A detailed overview of the WRF physics can be found in Table 1.





**Table 1.** WRF physics common to all sensitivity tests

Parameter	Option
Microphysics	WRF single-moment 6-class scheme
Longwave radiation	Rapid Radiative Transfer Model
Shortwave radiation	Dudhia scheme
Surface layer	MM5 similarity
Land surface	5-layer thermal diffusion
Planetary boundary layer	Yonsei University scheme
Cumulus	Kain-Fritsch scheme

With this meteorological input, version 2017r4 of the CHIMERE model (Mailler et al., 2017) is run in combination with the HTAP v2.2 inventory, representing anthropogenic emissions of the year 2010 (Janssens-Maenhout et al., 2015). Two domains are used for the CHIMERE model integrations: a coarse domain having a horizontal resolution of  $0.15^{\circ} \times 0.15^{\circ}$  (*longitude* × *latitude*), and a fine domain, nested into the former, having a resolution of  $0.05^{\circ} \times 0.04^{\circ}$  (see Figure 1a). Note that the terms

- 5 "coarse" and "fine" shall hereafter refer to the CHIMERE domains, not the WRF domains, if not otherwise stated. Regridding the HTAP dataset to these domains has been accomplished with the emiSURF program without using downscaling with traffic or population proxies (Mailler et al., 2017). Biogenic emissions are from the MEGAN model version 2.04 (Guenther et al., 2006) and mineral dust emissions within the CHIMERE domains are calculated on the basis of the United States Geological Survey landuse dataset (Loveland et al., 2000). The Alfaro and Gomes (2001) saltation and sandblasting scheme, optimized by
- 10 Menut et al. (2005), and the surface wind threshold described in Shao and Lu (2000) are used throughout all experiments. The effect of soil moisture on dust emissions (Fécan et al., 1998) is activated and so are sea-salt emissions. Vertical advection is achieved by the upwind scheme, horizontal advection by the more complex van Leer (1979) scheme. Carbonaceous species and interaction between aerosols and gases are taken into account by the model chemistry. The number of Gauss-Seidel iterations is set to 3, mainly because the model occasionally develops unrealistic waves with lower numbers. Urban correction/reduction
- 15 of the wind speed and the resuspension process are deactivated. A complete list of the internal CHIMERE parameters common to all sensitivity experiments is provided in Table 2. For a full description of these parameters, the interested reader is referred to the CHIMERE user manual available at http://www.lmd.polytechnique.fr/chimere.

Along the lateral boundaries of the coarse domain, the concentrations of the chemical species required by CHIMERE are provided by the forecasts of the ECMWF Composition Integrated Forecasting System (C-IFS) (Flemming et al., 2015), initial-

ized at 00 UTC. This global model comprises 60 vertical levels and has a horizontal resolution of  $\approx 80 km$ . In case a chemical species required by CHIMERE is not provided by C-IFS, the monthly climatological mean values from the MACC reanalysis (Inness et al., 2013) are used instead. Note that the time-varying dust data from C-IFS are scaled by a factor of 0.6 for all experiments since, otherwise, PM concentrations would be overestimated during the two Saharan dust events occurring in the time period considered here (summer 2018).





Table 2. CHIMERE parameters common to all sensitivity test

Parameter	Option
Nr. Gauss-Seidel iterations	3
Chemical time-step	adaptive
Physical time-step	5 minutes
Nr. of aerosol size sections	9
Chemically-active aerosols	yes
Anthropogenic emissions	HTAP v2.2 for the year 2010
Sea-salt emission parameterization	inert, parametrization 0
Biogenic emissions	MEGAN
Mineral dust emission	On
Saltation and sandblasting scheme	Alfaro and Gomes (2001), Menut et al. (2005)
Wind threshold estimation	Shao and Lu (2000)
Effect of soil moisture on mineral dust emissions	Fécan et al. (1998)
Secondary organic aerosol scheme	medium complexity
ISORROPIA coupling	yes
Inclusion of carbonaceous species	yes
Aerosol dry deposition	Zhang et al. (2001)
Horizontal advection scheme	van Leer
Vertical advection scheme	upwind
Urban correction	off
Resuspension process	off
Deep convection	on
Land cover dataset	USGS
Lateral boundary conditions	time-varying C-IFS data





5



**Figure 1.** (a) Horizontal CHIMERE domains used for all sensitivity experiments. The fine domain (orange rectangle) is nested into the coarser one (blue rectangle). At the lateral boundary conditions of the coarse domain, CHIMERE is fed by time varying C-IFS data. (b) The Galician air quality station network, grouped by the main pollution sources. See text for more details.

To eliminate unwanted effects related to the spin-up of the models, the daily WRF forecasts are initialized with the Digital Filtering Initialization (DFI) technique (Skamarock et al., 2008) and the first 3 integration hours are *not* used as meteorological input to CHIMERE. Consequently, CHIMERE is initialized on 03 UTC, using initial conditions from the model execution of the previous day, and is then integrated until 03 UTC of the following day to complete one forecast day. This procedure is repeated for each day from June 20, 2018 to August 31, 2018 and the resulting model output is then concatenated to form time series covering the entire time period.

## 2.2 Specific Configuration of the Sensitivity Tests

To assess the influence of *vertical* resolution on model performance, 10 layer experiments are compared to 20 layer experiments, the lowermost layer being located at 999 hPa and the uppermost at 500 hPa in all cases. Thus, in this study, an increase

- 10 in vertical resolution refers to a refinement of the domain in the lower to middle troposphere. An extension of the model top to, e.g., 200 hPa has been proposed in previous studies since some dust intrusions may extend to pressure levels above 500 hPa (Bessagnet et al., 2017). However, by design of our experiments, most of the dust intrusions' trajectory is simulated by C-IFS rather than internally simulated by CHIMERE and, therefore, elevating the model top is assumed to be of of minor importance here. The effect of an increase in *horizontal* resolution is tested by comparing the model output obtained with the coarse
- 15 resolution domain with that of the fine resolution domain (see Figure 1a). Note that the increase in horizontal resolution is undertaken in both CHIMERE and WRF, meaning that the combined effect is assessed here. Finally, all horizontal and vertical configurations are run twice, one time with the SAPRC mechanism and the other time with the Melchior mechanism, to asses





Table 3. Overview of the applied sensitivity tests, ordered according to their computational cost (from low to high), see text for more details.

Acronym	Horizontal Resolution (lat. $\times$ lon.)	Vertical Layers	Chemical Mechanism
CM10	WRF: $12 \times 12$ km, CHIMERE: $0.15^{\circ} \times 0.15^{\circ}$	10	Full Melchior
CM20		20	Full Melchior
CS10		10	SAPRC
CS20		20	SAPRC
FM10	WRF: $4 \times 4$ km, CHIMERE: $0.05^{\circ} \times 0.04^{\circ}$	10	Full Melchior
FM20		20	Full Melchior
FS10		10	SAPRC
FS20		20	SAPRC

the effects associated with the use of different model chemistries. An overview of these sensitivity tests, ordered from top to bottom according to their computational cost (from low to high), is provided in Table 3.

## 2.3 The Air Quality Monitoring Network in Northwestern Spain

The Galician air quality monitoring network comprises a total of 46 stations which, as a function of the main pollution source or the lack thereof, can be grouped into background, industrial and traffic sites (see Figure 1b). Currently, 14 stations are directly maintained by the Galician regional government (Xunta de Galicia). The remaining 32 stations are maintained by industrial companies which are supervised by the government in order to assure the same measurement standards, specified in

5 industrial companies which are supervised by the national UNE-EN norm.

The quality control of the corresponding data is accomplished *manually* by trained technical staff of the regional government, i.e. is centralised in one institution. First, outlier values are detected by comparing a suspicious value to the typical time series behaviour at the considered site and at the surrounding sites. Once the outlier is detected, its validity is determined

10 taking into account inter-variable relationships, potential power breakdowns, calibration errors, damages and changes in the topographic features surrounding the station. This way, a quality controlled observation dataset has been developed which, at some locations, is now nearly a decade long. This dataset serves as reference for model verification and will hereafter be referred to as the Galician Observational Air Quality Dataset.

# 2.4 Applied Verification Measures

15 Here, CHIMERE's performance with respect to observations is measured in terms of the Pearson correlation coefficient (r), the relative BIAS (see Equation 1), and the standard deviation ratio (see Equation 2):

$$bias = \frac{\overline{m} - \overline{o}}{\overline{o}} \times 100 \tag{1}$$





$$ratio = \frac{\sigma_m}{\sigma_o} \tag{2}$$

20

, where  $\overline{m}$ ,  $\overline{o}$ ,  $\sigma_m$  and  $\sigma_o$  are the modelled and observed values for the temporal mean and standard deviation, respectively. These measures are applied separately for the daily maximum and minimum time series of  $NO_2$ ,  $O_3$ , PM10 and PM2.5. Note that the chosen verification measures are complementary to each other since they cover different time series aspects. Namely, BIAS and RATIO measure the model's capacity to reproduce the observed mean and dispersion whereas r looks at the similarity in day-to-day variability irrespective of errors in the mean and dispersion. The perfect scores for r, BIAS and RATIO are 1, 0 and 1, respectively.

25

In addition, the mean absolute error (MAE) is a good measure of overall performance, and is here applied as a skill score (mean absolute error skill score, MAESS), i.e. as percentage deviation from the error of a reference experiment:

$$MAESS = \left(1 - \frac{MAE_i}{MAE_{ref}}\right) \times 100\tag{3}$$

, where  $MAE_i$  is the error a specific experiment *i* and  $MAE_{ref}$  the error of the experiment CS10, used as reference throughout the present study since it is the computationally least expensive experiment (see Table 3). Positive values indicate performance gains, negative values performances losses with respect to the reference. For more details on these verification measures, the interested reader is referred to Jolliffe and Stephenson (2012).

5 Results are provided by means of the boxplot, each one being computed upon the point-wise verification results at all available stations. A point-verification is excluded if the percentage of missing values present in the observed time series exceeds 20%.

## 3 Results

The verification results are displayed in Figures 2 and 3 for the daily maxima, and in Figure 4 and 5 for the daily minima,
respectively. Due to large differences between the results for specific variables, and also between the maxima and minima, it is not straightforward to use the same axis scaling for all panels. However, the number of different scalings was kept as small as possible to allow for comparison between one variable and another.

According to the default values of Python's Seaborn package, the centre line of each boxplot refers to the median value of the group of point-wise temporal verification results and the box refers the to the interquartile range of this group (IQR).

15 The whiskers extend from the 25th percentile minus  $1.5 \times IQR$  at the lower end to the 75th percentile plus  $1.5 \times IQR$  at the upper end. Outlier verification results lying beyond these limits are indicated by diamonds. Results obtained from the coarse horizontal resolution are depicted in magenta, those obtained from the fine horizontal resolution in green (see Figure 1a). The boxplots are ordered according to their computational demand, the cheapest experiment (CS10) located at the top and the most costly (FM20) at the bottom of each panel. Boxplots located on top of each other are equal in their experimental design,





20 except for the applied chemistry mechanism. This way, the effect of changing the mechanism can be easily catched by eye. All verification results are for the lowermost model layer, whose upper limit is located at 999 hPa, i.e. at roughly 10m above ground.

For the sake of simplicity, when referring to the results for BIAS, r and RATIO in the forthcoming, the range of the 8 spatial median values (1 median value per experiment) will be quoted first, followed by the overall minimum and maximum

25 obtained from all experiments and all station sites. The analyses of the performance gains (or losses) will then be more specific, quoting the spatial median, minimum and maximum MAESS *separately for each experiment*. Note, that the minimum and/or maximum values may be outliers, in which case they are not very representative of the dispersion. In this sense, the IQR is more representative and the reader should always double-check the dispersion having a look at the width of the boxes in Figures 3 and 4.

## 30 3.1 Maximum Values

Among all considered variables, and irrespective of the applied verification measure, CHIMERE's performance is poorest for  $NO_2$  (see see Figure 2a-c), in which case the bias ranges between -38 and -52% for the spatial median and between -91 and +244% at individual station sites (hereafter referred to as "extremes"). The correlation coefficient ranges between +0.11 and 0.24 (extremes: -0.40 and +0.65) and the time series' dispersion tends to be underestimated by the model, as is indicated by a standard deviation ratio between 0.44 and 0.70 (extremes: 0.07 and 5.95). Above all, the results for  $NO_2$  are characterised by a very pronounced spatial spread, i.e. by very large differences between one station and another.

- In comparison to the other chemical species assessed here, model performance is best for  $O_3$  (see Figure 2e-g). The bias ranges between +6% to +13% (extremes: -17% and +61%) and the correlation coefficient between +0.64 and +0.70 (extremes: +0.42 and +0.82). With standard deviation ratios ranging between 0.51 and 0.6 (extremes: 0.39 and 0.96), the time series' dispersion is clearly underestimated by the model. Also, the spatial spread of the verification results is generally smaller for  $O_3$  than for  $NO_2$ .
- 10 With bias values between -49 to -57% (extremes: -13 and -76%), correlation coefficients between +0.41 to +0.45 (extremes: -0.06 and 0.88), standard deviation ratios between 0.27 and 0.34 (extremes: 0.07 and 0.99) and a general increase in the spatial spread, results for PM10 are somewhat poorer than for O3 but still way-better than for  $NO_2$  (compare Figure 3a-c to Figure 2a-c and e-g).

Results for PM2.5 (see Figure 3e-g) are slightly better than for PM10. The bias improves to values between -27 and -39% (extremes: -69 and +14%), the correlation coefficient to values between +0.36 and +0.45 (extremes: 0 and +0.84) and

the standard deviation ratio to values between 0.43 and 0.56 (extremes: 0.09 and 1.43). The spatial spread of the results is comparable to that obtained for PM10.

The experiments' performance with reference to the base experiment (CS10) is displayed in Figures 2 and 3, panels d+h. Improvement is indicated by positive MAESS values and worsening by negative ones, respectively.

20 The MAESS results for  $NO_2$  are not satisfactory since none of the considered experiments yield a performance increase with respect to the base experiment and, even worse, an increase in horizontal resolution greatly inflates the dispersion of the results







**Figure 2.** Temporal verification results for daily near-surface *maximum*  $NO_2$  (left) and  $O_3$  (right). Row 1: bias, row 2: Pearson correlation coefficient, row 3: ratio of standard deviations, row 4: mean absolute error skill score (MAESS) with reference to the base experiment CS10. Each column pertains to a specific chemical species. Boxplots are calculated upon the point-wise verification results at all available stations. Abbreviations are explained in Table 3. See text for more details.





(see Figure 2d). For the remaining chemical species, model performance can be improved by increasing both, the horizontal and vertical resolution, particularly when using the SAPRC mechanism (i.e. experiment FS20). For this particular configuration, the spatial median performance increase is +5%, ranging between -10 and +19% at individual stations, when considering O<sub>3</sub>
(see FS20 in Figure 2h), +5% (extremes: -7 and +14%) when considering PM10 (see FS20 in Figure 3d), and +6% (extremes: -14 and +16%) for PM2.5 (see FS20 in Figure 3h). When using the high horizontal and vertical resolution in combination with full Melchior instead of SAPRC, on the one hand, even larger performance gains are yielded for O<sub>3</sub> (see FM20 in Figure 2h) but, on the other, no improvements are obtained for PM (see FM20 in Figure 3d+h) meaning that, *overall*, the best results for daily maximum values are achieved with the SAPRC mechanism. Importantly, an increase in horizontal resolution generally
inflates the spatial spread of the results (compare green with magenta boxplots in the last column of Figure 2 and 3). However, for the spatial spread of the results (compare green with magenta boxplots in the last column of Figure 2 and 3). However, for the spatial spread of the results (compare green with magenta boxplots in the last column of Figure 2 and 3).

for the particular case of  $O_3$ , this unwanted effect is alleviated by additionally augmenting the vertical resolution (see FS20 and FM20 in Figure 2h).

## 3.2 Minimum Values

In comparison to the daily maxima, CHIMERE's performance for the daily *minima* is poorer for  $O_3$ , similar for PM10 and PM2.5, and equally poor for  $NO_2$  (compare Figures 4 and 5 with Figures 2 and 3).

With bias values between -77 and -74% (extremes: -99 and +237%) and standard deviation ratios between 0.26 and 0.47
(extremes: 0.02 and 3.12), the model's performance for *NO*2 is again very poor, as was the case for the daily maxima (see Figure 4a+c). However, correlation coefficients lay in between +0.34 and 0.43 (extremes: -0.31 and +0.69), i.e. are better for the minima than for the maxima (compare Figure 4b with 2b).

For  $O_3$ , the bias ranges between +87 and +127% (extremes: +10 and 577%) and the correlation coefficient between +0.21 and 0.37 (extremes: -0.14 and 0.63). Thus, both the mean values and the day-to-day sequence of the time series are repro-

- 10 duced worse for the minima than for the maxima (compare Figure 4e+f with Figure 2e+f). However, with standard deviation ratios laying between 0.80 and 0.95 (extremes: 0.42 and 2.48), the simulated time series are underdispersed only in case the coarse horizontal mesh is used (see magenta boxplots in Figure 4g). The ratios for the fine horizontal mesh are almost equally distributed around unity, indicating underdispersion for roughly one half of the station sites and overdispersion for the other half (see green boxplots in Figure 4g). However, albeit better standard deviation ratios are obtained for the minima than for
- 5 the maxima if one looks at the spatial median, results are clearly worse for the minima when considering the spatial spread (compare Figure 4g with 2g).

Unlike the results for the maxima, the median bias for the minima is near to zero for PM10, ranging between -9 and +9% (extremes: -57 and +108%), and positive for PM2.5, ranging between +39 and +63% (extremes: -29 to 306%), indicating a larger bias magnitude for the fine fraction (compare panels a and e in Figure 5). With values ranging between 0.31 to 0.34 (0.07 to 0.64) for PM10 and 0.33 to 0.40 (extremes: 0.14 to 0.58) for PM2.5, correlation coefficients are similar for the two size groups, with slight advantages for the fine fraction (compare panels b and f in Figure 5). The standard deviation ratio lays between 0.76 and 0.96 (extremes: 0.44 and 2.05) for PM10 and between 0.95 and 1.23 (extremes: 0.68 and 2.80) for PM2.5.







Figure 3. As Figure 2, but for PM10 and PM2.5.







Figure 4. Temporal verification results for daily near-surface *minimum*  $NO_2$  (left) and  $O_3$  (right). Row 1: bias, row 2: Pearson correlation coefficient, row 3: ratio of standard deviations, row 4: mean absolute error skill score (MAESS) with reference to the base experiment CS10. Each column pertains to a specific chemical species. Boxplots are calculated upon the point-wise verification results at all available stations. Abbreviations are explained in Table 3. See text for more details.







Figure 5. As Figure 4, but for PM10 and PM2.5.

5 In comparison to the maxima, this indicates a closer match with unity on the one hand, but a larger spatial spread on the other (compare Figure 5c+g with Figure 3c+g).





10

20

As for the daily maxima, the poor performance for NO2 cannot be improved by changing the experimental design and the use of the fine horizontal mesh greatly inflates the spread of the results (see Figure 4d). For both  $O_3$  and PM, full Melchior works better than SAPRC (see panels Figure 4h and 5d+h), i.e. the model's response to changes in the chemistry mechanism is more homogeneous for the minima than for the maxima (in which case full Melchior performed better for  $O_3$  but worse for PM). As for the maxima, optimal results for  $O_3$  can be obtained with a fine vertical resolution, without the need to additionally increase the horizontal resolution, yielding a performance gain of +11% for the spatial median, ranging between +7 and +18% at individual station sites (see CM20 in Figure 4h). If the horizontal resolution is increased in addition, even better results are obtained for the spatial median (+18% for FM20) but the spread of the results increases and performance losses are obtained

5 at some sites (extremes for FM20: -10% to +51%). Hence, when solely looking at the  $O_3$  minima, the most homogeneous performance gains are obtained with the fine vertical and coarse horizontal mesh.

The most balanced improvements for PM10 are yielded by using full Melchior and the coarse horizontal resolution. In this configuration, a +2% gain is obtained for both the 10 and 20 layer set-up (see CM10 and CM20 in Figure 5d), obtaining a lower spatial spread for the latter (-12 and +14% versus -3 and +7%). An increase in horizontal resolution is associated with

10 a slight spread growth (compare FM10 and FM20 to CM10 and CM20). The magnitude of the performance gains is generally smaller for PM10 than for  $O_3$  (compare Figure 5d to 4h).

For PM2.5, using full Melchior is even more advantageous than for PM10 and the unfavourable spread increase associated with the fine *horizontal* mesh is much reduced if 20 vertical layers are used in addition (see Figure 5h). Thus, experiment FM20 is a reasonable choice for this variable, yielding a performance gain of +3% (extremes: -10 to +14%). The most homogeneous

gains, however, are achieved with the computationally cheap CM10 experiment (spatial median: +8%, extremes: -8% to +11%),
 i.e. using the coarse horizontal and vertical mesh.

## 3.3 Verification Results per Pollution Source

spatial median is displayed) in order to provide a visual summary.

Figure 6 shows the spatial median MAESS with reference to the base experiment CS10 for all locations (column 1), as well as for background, industry and traffic locations (columns 2 to 4). The first row refers to the daily maxima, the second to the minima, respectively. Improvement over the base experiment is indicated by green, worsening by red colour shadings. The results in column 1 have been already stated in Section 3.1 and 3.2 and are here displayed again in a simplified form (only the

Irrespective of the main pollution source, performance improvements are larger for the daily minima than for the maxima (compare row 1 wit row 2 in Figure 6). For the minima, particularly large gains are found at those sites influenced by anthro-

- pogenic emissions, i.e. at industrial and traffic sites (see Figure 6g and h). For the daily minima at background sides, however, virtually no gains are obtained for PM and performance losses are obtained for  $NO_2$ , meaning that sophisticated model configurations are inefficient or even counterproductive if the observed pollutant concentrations are very low (see first, third and fourth column in Figure 6f). On the contrary, if the observed concentrations are appreciable, as is the case for minimum  $O_3$ , then the performance *can* be increased with more sophisticated configurations (see CS20, CM20, FS20 and FM20 in Figure
- 30 6f, second column).







**Figure 6.** Spatial median Mean Absolute Error Skill Score (MAESS) with respect to the base experiment CS10 for (row 1) the daily maxima and (row 2) daily minima at (column 1) all available stations, (column 2) background stations, (column 3) industrial stations and (column 4) traffic stations. See text for more details.

## 4 Conclusions

In this study, a series of sensitivity tests was carried out with the chemical weather forecasting model CHIMERE over the northwestern Iberian Peninsula for the 2018 summer season. The model was driven by meteorological data from WRF, the HTAP v2.2 emission inventory and lateral boundary data from the ECMWF C-IFS.

- For daily *maximum* surface ozone, optimal results are obtained with 20 vertical layers between 999 and 500 hPa, a relatively coarse horizontal resolution  $(0.15^{\circ} \text{ longitude} \times 0.15^{\circ} \text{ latitude})$  and the full Melchior mechanism. To achieve the best possible results for *daily maximum particulate matter* it is also necessary to use 20 vertical layers but, in this case, a finer horizontal mesh  $(0.05^{\circ} \times 0.04^{\circ})$  should be used in combination with the SAPRC mechanism. For the daily *minimum* values, full Melchior performs better than SAPRC irrespective of the considered variable/chemical species. To yield optimal results for  $O_3$ , it is
- 10 important to additionally use 20 vertical layers and the fine horizontal mesh. A pollution source oriented analysis revealed that the performance increases of the more sophisticated experiments are largest when the observed concentrations are low but yet appreciable.

We conclude that the best overall results are obtained by using the fine horizontal and vertical mesh. If daily maxima are considered more important than daily minima, the SAPRC mechanism provides the best overall performance for ozone and

15 PM. If daily minima are considered equally important, the full Melchior mechanism should be chosen. Irrespective of the





5

experimental configuration, CHIMERE's performance for  $NO_2$  is very poor, which might point to deficiencies in the HTAP v2.2 emission inventory for this chemical species in our region of interest.

Particularly the promising results for the 20 layer setup confined to the lower and middle troposphere (999 to 500 hPa) might as well yield performance improvements for CHIMERE experiments run on larger scales, e.g. for the CHIMERE implementation participating in the Europe-wide ensemble hosted by Copernicus (Marécal et al., 2015), or for the hindcast runs

20 plementation participating in the Europe-wide ensemble hosted by Copernicus (Marécal et al., 2015), or for the hindcast runs covering several decades that have been accomplished in the EURODELTA-Trends initiative (Colette et al., 2017). It would be interesting to test these hypotheses in future studies.

*Code availability.* The CHIMERE and WRF source codes are publicly available from http://www.lmd.polytechnique.fr/chimere and http://www2.mmm.ucar.edu/wrf/users/downloads.html, respectively.

25 *Data availability.* The CHIMERE and observational data underlying the verification results presented here have been made available by the authors and can be retrieved from the public Dropbox link https://www.dropbox.com/s/eb6y5pjsb6phbf2/Brands\_et\_al\_2019\_acpd\_ underlying\_data.zip?dl=0. The GFS data used to drive WRF and the C-IFS data used as lateral boundary input to CHIMERE can be down-loaded from https://rda.ucar.edu/datasets/ds084.1 and ftp://dissemination.ecmwf.int, respectively.

*Author contributions*. Swen Brands designed and executed the CHIMERE experiments, built the figures, analysed the results and wrote the manuscript. Guillermo Fernández-García and Marcos Tesouro Montecelo provided the meteorological input from WRF and contributed to the manuscript. Nuria Fernández García, Anthony Estévez Saunders, Paula Costa Tomé and Cristina Otero were responsible for the quality control of the observations and contributed to the manuscript. Pablo Enrique Carracedo García, Anabela Neto Venancio, Pedro Daniel Melo Costa, María Luz Macho and Juan Taboada contributed to the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. The authors would like to thank the CHIMERE development team working at LMD Paris and also the WRF development team, for providing the models' source codes and technical support. They gratefully acknowledge the computational resources and technical support provided by the Centro de Supercomputación de Galicia (CESGA), as well as the free availability of the global predictions from the
 10 GFS and C-IFS forecasting systems maintained by NCEP and ECMWF/Copernicus, respectively.





#### References

- Alfaro, S. C. and Gomes, L.: Modeling mineral aerosol production by wind erosion: Emission intensities and aerosol size distributions in source areas, Journal of Geophysical Research: Atmospheres, 106, 18075–18084, https://doi.org/10.1029/2000JD900339, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900339, 2001.
- 5 Balzarini, A., Pirovano, G., Honzak, L., Žabkar, R., Curci, G., Forkel, R., Hirtl, M., José, R. S., Tuccella, P., and Grell, G.: WRF-Chem model sensitivity to chemical mechanisms choice in reconstructing aerosol optical properties, Atmospheric Environment, 115, 604 – 619, https://doi.org/https://doi.org/10.1016/j.atmosenv.2014.12.033, http://www.sciencedirect.com/science/article/pii/ S1352231014009819, 2015.

Banzhaf, S., Schaap, M., Kerschbaumer, A., Reimer, E., Stern, R., van der Swaluw, E., and Builtjes, P.: Implementation and evalu-

- 10 ation of pH-dependent cloud chemistry and wet deposition in the chemical transport model REM-Calgrid, Atmospheric Environment, 49, 378 – 390, https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.10.069, http://www.sciencedirect.com/science/article/pii/ S135223101101199X, 2012.
  - Banzhaf, S., Schaap, M., Kranenburg, R., Manders, A. M. M., Segers, A. J., Visschedijk, A. J. H., Denier van der Gon, H. A. C., Kuenen, J. J. P., van Meijgaard, E., van Ulft, L. H., Cofala, J., and Builtjes, P. J. H.: Dynamic model evaluation for secondary inorganic aerosol
- 15 and its precursors over Europe between 1990 and 2009, Geoscientific Model Development, 8, 1047–1070, https://doi.org/10.5194/gmd-8-1047-2015, https://www.geosci-model-dev.net/8/1047/2015/, 2015.
  - Beegum, S. N., Gherboudj, I., Chaouch, N., Couvidat, F., Menut, L., and Ghedira, H.: Simulating aerosols over Arabian Peninsula with CHIMERE: Sensitivity to soil, surface parameters and anthropogenic emission inventories, Atmospheric Environment, 128, 185 – 197, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.01.010, http://www.sciencedirect.com/science/article/pii/ \$1352231016300188\_2016
- **20** \$1352231016300188, 2016.
  - Bessagnet, B., Menut, L., Colette, A., Couvidat, F., Dan, M., Mailler, S., Létinois, L., Pont, V., and Rouïl, L.: An evaluation of the CHIMERE Chemistry Transport Model to simulate dust outbreaks across the Northern Hemisphere in March 2014, Atmosphere, 8, https://doi.org/10.3390/atmos8120251, http://www.mdpi.com/2073-4433/8/12/251, 2017.

Brasseur, G. P., Xie, Y., Petersen, A. K., Bouarar, I., Flemming, J., Gauss, M., Jiang, F., Kouznetsov, R., Kranenburg, R., Mijling, B.,

- Peuch, V.-H., Pommier, M., Segers, A., Sofiev, M., Timmermans, R., van der A, R., Walters, S., Xu, J., and Zhou, G.: Ensemble forecasts of air quality in eastern China – Part 1: Model description and implementation of the MarcoPolo–Panda prediction system, version 1, Geoscientific Model Development, 12, 33–67, https://doi.org/10.5194/gmd-12-33-2019, https://www.geosci-model-dev.net/12/33/2019/, 2019.
- Caplan, P., Derber, J., Gemmill, W., Hong, S.-Y., Pan, H.-L., and Parrish, D.: Changes to the 1995 NCEP Operational
   Medium-Range Forecast Model Analysis–Forecast System, Weather and Forecasting, 12, 581–594, https://doi.org/10.1175/1520-0434(1997)012<0581:CTTNOM>2.0.CO;2, https://doi.org/10.1175/1520-0434(1997)012<0581:CTTNOM>2.0.CO;2, 1997.
  - Chang, J. C. and Hanna, S. R.: Air quality model performance evaluation, Meteorology and Atmospheric Physics, 87, 167–196, https://doi.org/10.1007/s00703-003-0070-7, https://doi.org/10.1007/s00703-003-0070-7, 2004.
- Colette, A., Granier, C., Hodnebrog, Ø., Jakobs, H., Maurizi, A., Nyiri, A., Bessagnet, B., D'Angiola, A., D'Isidoro, M., Gauss, M., Meleux,
  F., Memmesheimer, M., Mieville, A., Rouïl, L., Russo, F., Solberg, S., Stordal, F., and Tampieri, F.: Air quality trends in Europe over the past decade: a first multi-model assessment, Atmospheric Chemistry and Physics, 11, 11657–11678, https://doi.org/10.5194/acp-11-11657-2011, https://www.atmos-chem-phys.net/11/11657/2011, 2011.





30

- Colette, A., Andersson, C., Manders, A., Mar, K., Mircea, M., Pay, M.-T., Raffort, V., Tsyro, S., Cuvelier, C., Adani, M., Bessagnet, B., Bergström, R., Briganti, G., Butler, T., Cappelletti, A., Couvidat, F., D'Isidoro, M., Doumbia, T., Fagerli, H., Granier, C., Heyes, C., Klimont, Z., Ojha, N., Otero, N., Schaap, M., Sindelarova, K., Stegehuis, A. I., Roustan, Y., Vautard, R., van Meijgaard, E., Vivanco, M. G.,
- 5 and Wind, P.: EURODELTA-Trends, a multi-model experiment of air quality hindcast in Europe over 1990–2010, Geoscientific Model Development, 10, 3255–3276, https://doi.org/10.5194/gmd-10-3255-2017, https://www.geosci-model-dev.net/10/3255/2017/, 2017.
  - Crippa, P., Sullivan, R. C., Thota, A., and Pryor, S. C.: The impact of resolution on meteorological, chemical and aerosol properties in regional simulations with WRF-Chem, Atmospheric Chemistry and Physics, 17, 1511–1528, https://doi.org/10.5194/acp-17-1511-2017, https://www.atmos-chem-phys.net/17/1511/2017/, 2017.
- 10 Deroubaix, A., Menut, L., Flamant, C., Brito, J., Denjean, C., Dreiling, V., Fink, A., Jambert, C., Kalthoff, N., Knippertz, P., Ladkin, R., Mailler, S., Maranan, M., Pacifico, F., Piguet, B., Siour, G., and Turquety, S.: Diurnal cycle of coastal anthropogenic pollutant transport over southern West Africa during the DACCIWA campaign, Atmospheric Chemistry and Physics, 19, 473–497, https://doi.org/10.5194/acp-19-473-2019, https://www.atmos-chem-phys.net/19/473/2019/, 2019.
- EU: Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe, Official Journal of the European Union, 2008.
- Fécan, F., Marticorena, B., and Bergametti, G.: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semiarid areas, Annales Geophysicae, 17, 149–157, https://doi.org/10.1007/s00585-999-0149-7, https://doi.org/ 10.1007/s00585-999-0149-7, 1998.

Flamant, C., Deroubaix, A., Chazette, P., Brito, J., Gaetani, M., Knippertz, P., Fink, A. H., de Coetlogon, G., Menut, L., Colomb, A.,

20 Denjean, C., Meynadier, R., Rosenberg, P., Dupuy, R., Dominutti, P., Duplissy, J., Bourrianne, T., Schwarzenboeck, A., Ramonet, M., and Totems, J.: Aerosol distribution in the northern Gulf of Guinea: local anthropogenic sources, long-range transport, and the role of coastal shallow circulations, Atmospheric Chemistry and Physics, 18, 12 363–12 389, https://doi.org/10.5194/acp-18-12363-2018, https: //www.atmos-chem-phys.net/18/12363/2018/, 2018.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness,

- 25 A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geoscientific Model Development, 8, 975–1003, https://doi.org/10.5194/gmd-8-975-2015, https://www.geosci-model-dev.net/8/975/2015/, 2015.
  - Giorgi, F. and Francisco, R.: Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM, Climate Dynamics, 16, 169–182, https://doi.org/10.1007/PL00013733, https://doi.org/10.1007/PL00013733, 2000.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmospheric Chemistry and Physics, 6, 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, https://www.atmos-chem-phys.net/6/3181/2006/, 2006.
- Im, U., Brandt, J., Geels, C., Hansen, K. M., Christensen, J. H., Andersen, M. S., Solazzo, E., Kioutsioukis, I., Alyuz, U., Balzarini, A., Baro,
- R., Bellasio, R., Bianconi, R., Bieser, J., Colette, A., Curci, G., Farrow, A., Flemming, J., Fraser, A., Jimenez-Guerrero, P., Kitwiroon, N., Liang, C.-K., Nopmongcol, U., Pirovano, G., Pozzoli, L., Prank, M., Rose, R., Sokhi, R., Tuccella, P., Unal, A., Vivanco, M. G., West, J., Yarwood, G., Hogrefe, C., and Galmarini, S.: Assessment and economic valuation of air pollution impacts on human health over Europe and the United States as calculated by a multi-model ensemble in the framework of AQMEII3, Atmospheric Chemistry and Physics, 18, 5967–5989, https://doi.org/10.5194/acp-18-5967-2018, https://www.atmos-chem-phys.net/18/5967/2018/, 2018a.





- Im, U., Christensen, J. H., Geels, C., Hansen, K. M., Brandt, J., Solazzo, E., Alyuz, U., Balzarini, A., Baro, R., Bellasio, R., Bianconi, R., Bieser, J., Colette, A., Curci, G., Farrow, A., Flemming, J., Fraser, A., Jimenez-Guerrero, P., Kitwiroon, N., Liu, P., Nopmongcol, U., Palacios-Peña, L., Pirovano, G., Pozzoli, L., Prank, M., Rose, R., Sokhi, R., Tuccella, P., Unal, A., Vivanco, M. G., Yarwood, G.,
- 5 Hogrefe, C., and Galmarini, S.: Influence of anthropogenic emissions and boundary conditions on multi-model simulations of major air pollutants over Europe and North America in the framework of AQMEII3, Atmospheric Chemistry and Physics, 18, 8929–8952, https://doi.org/10.5194/acp-18-8929-2018, https://www.atmos-chem-phys.net/18/8929/2018/, 2018b.
  - Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J.,
- 10 Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, Atmospheric Chemistry and Physics, 13, 4073–4109, https://doi.org/10.5194/acp-13-4073-2013, https://www.atmos-chem-phys.net/13/4073/2013/, 2013.
  - Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP v2.2: a mosaic of regional and
- 15global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmospheric Chemistry and Physics, 15,<br/>11411–11432, https://doi.org/10.5194/acp-15-11411-2015, https://www.atmos-chem-phys.net/15/11411/2015/, 2015.
  - Jolliffe, I. T. and Stephenson, D. B., eds.: Forecast verification: a practitioner's guide in atmospheric science, Wiley, 2012.
  - Jonson, J. E., Simpson, D., Fagerli, H., and Solberg, S.: Can we explain the trends in European ozone levels?, Atmospheric Chemistry and Physics, 6, 51–66, https://doi.org/10.5194/acp-6-51-2006, https://www.atmos-chem-phys.net/6/51/2006/, 2006.
- 20 Karlický, J., Huszár, P., and Halenka, T.: Validation of gas phase chemistry in the WRF-Chem model over Europe, Advances in Science and Research, 14, 181–186, https://doi.org/10.5194/asr-14-181-2017, https://www.adv-sci-res.net/14/181/2017/, 2017.
  - Kukkonen, J., Olsson, T., Schultz, D. M., Baklanov, A., Klein, T., Miranda, A. I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S., Sofiev, M., Sokhi, R., Lehtinen, K. E. J., Karatzas, K., San José, R., Astitha, M., Kallos, G., Schaap, M., Reimer, E., Jakobs, H., and Eben, K.: A review of operational, regional-scale, chemical weather forecasting models in
- Europe, Atmospheric Chemistry and Physics, 12, 1–87, https://doi.org/10.5194/acp-12-1-2012, https://www.atmos-chem-phys.net/12/1/ 2012/, 2012.
  - Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, International Journal of Remote Sensing, 21, 1303–1330, https://doi.org/10.1080/014311600210191, https://doi.org/10.1080/014311600210191, 2000.
- 30 Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P., Bessagnet, B., Colette, A., Létinois, L., Markakis, K., and Meleux, F.: CHIMERE-2017: from urban to hemispheric chemistry-transport modeling, Geoscientific Model Development, 10, 2397–2423, https://doi.org/10.5194/gmd-10-2397-2017, https://www.geosci-model-dev.net/10/2397/2017/, 2017.
  - Marécal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado,
- 35 A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Denier van der Gon, H. A. C., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouïl, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, A., van Velthoven, P., van Versendaal, R., Vira, J., and Ung, A.: A regional





5

10

15

air quality forecasting system over Europe: the MACC-II daily ensemble production, Geoscientific Model Development, 8, 2777-2813, https://doi.org/10.5194/gmd-8-2777-2015, https://www.geosci-model-dev.net/8/2777/2015/, 2015.

Markakis, K., Valari, M., Perrussel, O., Sanchez, O., and Honore, C.: Climate-forced air-quality modeling at the urban scale: sensitivity to model resolution, emissions and meteorology, Atmospheric Chemistry and Physics, 15, 7703-7723, https://doi.org/10.5194/acp-15-7703-

2015, https://www.atmos-chem-phys.net/15/7703/2015/, 2015.

- Mazzeo, A., Huneeus, N., Ordoñez, C., Orfanoz-Cheuquelaf, A., Menut, L., Mailler, S., Valari, M., van der Gon, H. D., Gallardo, L., Muñoz, R., Donoso, R., Galleguillos, M., Osses, M., and Tolvett, S.: Impact of residential combustion and transport emissions on air pollution in Santiago during winter, Atmospheric Environment, 190, 195 - 208, https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.06.043, http://doi.org/10.1016/j.atmosenv.2018.06.043, http://doi.org/10.1016/j.atmosenv.2018.06.043, https://doi.org/10.1016/j.atmosenv.2018.06.043, https://doi.04144, https://doi.041444, https://doi.04144, https://doi.04144, https://doi.04144, https://doi.04144, https://doi.04144, https://doi.04144, https://doi.04144, https://doi.0414 //www.sciencedirect.com/science/article/pii/S1352231018304345, 2018.
- Menut, L.: Sensitivity of hourly Saharan dust emissions to NCEP and ECMWF modeled wind speed, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009522, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009522, 2008.

Menut, L., Schmechtig, C., and Marticorena, B.: Sensitivity of the sandblasting flux calculations to the soil size distribution accuracy, Journal of Atmospheric and Oceanic Technology, 22, 1875–1884, https://doi.org/10.1175/JTECH1825.1, https://doi.org/10.1175/JTECH1825.1, 2005.

- Menut, L., Bessagnet, B., Colette, A., and Khvorostiyanov, D.: On the impact of the vertical resolution on chemistry-transport modelling, Atmospheric Environment, 67, 370 - 384, https://doi.org/https://doi.org/10.1016/j.atmosenv.2012.11.026, 2013.
- Menut, L., Flamant, C., Turquety, S., Deroubaix, A., Chazette, P., and Meynadier, R.: Impact of biomass burning on pollutant surface concentrations in megacities of the Gulf of Guinea, Atmospheric Chemistry and Physics, 18, 2687-2707, https://doi.org/10.5194/acp-18-20 2687-2018, https://www.atmos-chem-phys.net/18/2687/2018/, 2018.
- Misenis, C. and Zhang, Y.: An examination of sensitivity of WRF/Chem predictions to physical parameterizations, horizontal grid spacing, and nesting options, Atmospheric Research, 97, 315 - 334, https://doi.org/https://doi.org/10.1016/j.atmosres.2010.04.005, 2010.

Monteiro, A., Russo, M., Gama, C., and Borrego, C.: How important are maritime emissions for the air quality: At European and national scale, Environmental Pollution, 242, 565 - 575, https://doi.org/10.1016/j.envpol.2018.07.011, http://www.sciencedirect. com/science/article/pii/S026974911830962X, 2018.

- 25
- Otero, N., Sillmann, J., Mar, K. A., Rust, H. W., Solberg, S., Andersson, C., Engardt, M., Bergström, R., Bessagnet, B., Colette, A., Couvidat, F., Cuvelier, C., Tsyro, S., Fagerli, H., Schaap, M., Manders, A., Mircea, M., Briganti, G., Cappelletti, A., Adani, M., D'Isidoro, M., Pay, M.-T., Theobald, M., Vivanco, M. G., Wind, P., Ojha, N., Raffort, V., and Butler, T.: A multi-model comparison of meteorological drivers of surface ozone over Europe, Atmospheric Chemistry and Physics, 18, 12 269-12 288, https://doi.org/10.5194/acp-18-12269-
- 30 2018, https://www.atmos-chem-phys.net/18/12269/2018/, 2018.
- Palmer, T. N., Doblas-Reyes, F. J., Weisheimer, A., and Rodwell, M. J.: Toward seamless prediction: calibration of climate change projections using seasonal forecasts, Bulletin of the American Meteorological Society, 89, 459-470, https://doi.org/10.1175/BAMS-89-4-459, 2008.
- Schaap, M., Cuvelier, C., Hendriks, C., Bessagnet, B., Baldasano, J., Colette, A., Thunis, P., Karam, D., Fagerli, H., Graff, A., Kranenburg, R., Nyiri, A., Pay, M., Rouïl, L., Schulz, M., Simpson, D., Stern, R., Terrenoire, E., and
- 35 Wind, P.: Performance of European chemistry transport models as function of horizontal resolution, Atmospheric Environment, 112, 90 - 105, https://doi.org/10.1016/j.atmosenv.2015.04.003, http://www.sciencedirect.com/science/article/pii/ \$1352231015300066, 2015.
  - Shao, Y. and Lu, H.: A simple expression for wind erosion threshold friction velocity, Journal of Geophysical Research: Atmospheres, 105, 22 437-22 443, https://doi.org/10.1029/2000JD900304, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900304, 2000.





Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., and Powers, J.: A description of the Advanced Research WRF version 3, NCAR Technical Note NCAR/TN-475+STR, National Center for Atmospheric Research, Boulder, Colorado, USA, 126pp., 2008.

Teixeira, J., Carvalho, A., Tuccella, P., Curci, G., and Rocha, A.: WRF-chem sensitivity to vertical resolution during a saharan dust event,

- 5 Physics and Chemistry of the Earth, Parts A/B/C, 94, 188 195, https://doi.org/https://doi.org/10.1016/j.pce.2015.04.002, http://www. sciencedirect.com/science/article/pii/S1474706515000364, 2016.
  - Theobald, M. R., Vivanco, M. G., Aas, W., Andersson, C., Ciarelli, G., Couvidat, F., Cuvelier, K., Manders, A., Mircea, M., Pay, M.-T., Tsyro, S., Adani, M., Bergström, R., Bessagnet, B., Briganti, G., Cappelletti, A., D'Isidoro, M., Fagerli, H., Mar, K., Otero, N., Raffort, V., Roustan, Y., Schaap, M., Wind, P., and Colette, A.: An evaluation of European nitrogen and sulfur wet deposition and their trends estimated by six chemistry transport models for
- 10 of European nitrogen and sulfur wet deposition and their trends estimated by six chemistry transport models for the period 1990–2010, Atmospheric Chemistry and Physics, 19, 379–405, https://doi.org/10.5194/acp-19-379-2019, https://www.atmos-chem-phys.net/19/379/2019/, 2019.

Thunis, P., Degraeuwe, B., Pisoni, E., Meleux, F., and Clappier, A.: Analyzing the efficiency of short-term air quality plans in European cities, using the CHIMERE air quality model, Air. Qual. Atmos. Health., 10, 235–248, https://doi.org/10.1007/s11869-016-0427-y, 2016.

- 15 Valari, M. and Menut, L.: Does an increase in air quality models' resolution bring surface ozone concentrations closer to reality?, Journal of Atmospheric and Oceanic Technology, 25, 1955–1968, https://doi.org/10.1175/2008JTECHA1123.1, 2008.
  - van Leer, B.: Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method, Journal of Computational Physics, 32, 101–136, https://doi.org/https://doi.org/10.1016/0021-9991(79)90145-1, http://www.sciencedirect.com/science/article/ pii/0021999179901451, 1979.
- 20 Vautard, R., Szopa, S., Beekmann, M., Menut, L., Hauglustaine, D. A., Rouil, L., and Roemer, M.: Are decadal anthropogenic emission reductions in Europe consistent with surface ozone observations?, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026080, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL026080, 2006.
  - Vivanco, M. G., Theobald, M. R., García-Gómez, H., Garrido, J. L., Prank, M., Aas, W., Adani, M., Alyuz, U., Andersson, C., Bellasio, R., Bessagnet, B., Bianconi, R., Bieser, J., Brandt, J., Briganti, G., Cappelletti, A., Curci, G., Christensen, J. H., Colette, A., Couvidat, F.,
- 25 Cuvelier, C., D'Isidoro, M., Flemming, J., Fraser, A., Geels, C., Hansen, K. M., Hogrefe, C., Im, U., Jorba, O., Kitwiroon, N., Manders, A., Mircea, M., Otero, N., Pay, M.-T., Pozzoli, L., Solazzo, E., Tsyro, S., Unal, A., Wind, P., and Galmarini, S.: Modeled deposition of nitrogen and sulfur in Europe estimated by 14 air quality model systems: evaluation, effects of changes in emissions and implications for habitat protection, Atmospheric Chemistry and Physics, 18, 10199–10218, https://doi.org/10.5194/acp-18-10199-2018, https://www.atmos-chem-phys.net/18/10199/2018/, 2018.
- 30 Wilson, R. C., Fleming, Z. L., Monks, P. S., Clain, G., Henne, S., Konovalov, I. B., Szopa, S., and Menut, L.: Have primary emission reduction measures reduced ozone across Europe? An analysis of European rural background ozone trends 1996–2005, Atmospheric Chemistry and Physics, 12, 437–454, https://doi.org/10.5194/acp-12-437-2012, https://www.atmos-chem-phys.net/12/437/2012/, 2012.
  - Zhang, L., Gong, S., Padro, J., and Barrie, L.: A size-segregated particle dry deposition scheme for an atmospheric aerosol module, Atmospheric Environment, 35, 549 – 560, https://doi.org/https://doi.org/10.1016/S1352-2310(00)00326-5, http://www.sciencedirect.com/
- **35** science/article/pii/S1352231000003265, 2001.