CHIMERE-2017: from urban to hemispheric chemistry-transport modeling

Sylvain Mailler1,2, Laurent Menut1, Dmitry Khvorostyanov1, Myrto Valari1, Florian Couvidat3, Guillaume Siour4, Solène Turquety1, Régis Briant1, Paolo Tuccella1, Bertrand Bessagnet3, Augustin Colette3, Laurent Létinois3, Kostantinos Markakis1, and Frédérik Meleux3

1LMD/IPSL, École Polytechnique, Université Paris Saclay, ENS, PSL Research University; Sorbonne Universités, UPMC Univ Paris 06, CNRS, Palaiseau, France
2École des Ponts ParisTech, Université Paris-Est, 77455 Champs-sur-Marne, France
3INERIS, National Institute for Industrial Environment and Risks, Parc Technologique ALATA, 60550 Verneuil-en-Halatte, France
4Laboratoire Interuniversitaire des Systèmes Atmosphériques (LISA), UMR CNRS 7583, Université Paris Est Créteil et Université Paris Diderot, Institut Pierre Simon Laplace, Créteil, France

Correspondence to: Sylvain Mailler (sylvain.mailler@lmd.polytechnique.fr)

Received: 20 July 2016 – Discussion started: 7 September 2016
Revised: 15 May 2017 – Accepted: 18 May 2017 – Published: 28 June 2017

Abstract. CHIMERE is a chemistry-transport model designed for regional atmospheric composition. It can be used at a variety of scales from local to continental domains. However, due to the model design and its historical use as a regional model, major limitations had remained, hampering its use at hemispheric scale, due to the coordinate system used for transport as well as to missing processes that are important in regions outside Europe. Most of these limitations have been removed in the CHIMERE-2017 version, allowing its use in any region of the world and at any scale, from the scale of a single urban area up to hemispheric scale, with or without polar regions included. Other important improvements have been made in the treatment of the physical processes affecting aerosols and the emissions of mineral dust. From a computational point of view, the parallelization strategy of the model has also been updated in order to improve model numerical performance and reduce the code complexity. The present article describes all these changes. Statistical scores for a model simulation over continental Europe are presented, and a simulation of the circumpolar transport of volcanic ash plume from the Puyehue volcanic eruption in June 2011 in Chile provides a test case for the new model version at hemispheric scale.

1 Introduction

Deterministic chemistry-transport modeling is now widely used for the analysis of pollution events, scenarios and forecast (Monks et al., 2009). Numerous models exist and are used from local to global scale, both for gaseous and aerosols modeling (Simpson et al., 2012; Inness et al., 2013, among many others). While models were previously dedicated mainly to specific processes, the latest generation of chemistry-transport models (CTMs) aims at representing the complete set of processes leading to changes in the atmospheric composition in terms of aerosols and trace gases. For regional air quality in the troposphere, several CTMs are currently developed and are able to include all types of emissions: anthropogenic, biogenic, mineral dust, sea salt, vegetation fires and volcanos. Even though all these emission processes are now included in many CTMs, the emitted species have different chemistry and lifetimes, and models often address some specific applications and thus specific spatial areas. This was the case of the CHIMERE model, extensively described in Menut et al. (2013a) for its 2013 version. Originally, CHIMERE was designed for urban areas. It was extended later to western Europe, and then to the northern part of Africa by including mineral dust emissions, but was limited to these areas only, due to limitations in available data...
(such as the anthropogenic emissions). The typical resolution (grid spacing) of the simulation domains range from 4 km for urban-scale domains to about 50 km for regional-scale domains Markakis et al. (2015); Valari and Menut (2008).

The CHIMERE model has been used for a long time for studies at the urban to regional scale. Vautard et al. (2007) has used this model within the CityDelta project over four major urban areas in Europe (Berlin, Milan, Paris and Prague), at a horizontal resolution of 5 km. While this resolution is not sufficient to resolve adequately urban-scale phenomena, Valari and Menut (2008) have shown that due to limitations in the accuracy of the input meteorological fields, increasing the horizontal model resolution to values lower than 10 km might actually degrade model performance. The same authors (Valari and Menut, 2010), show that, actually, rather than increasing the model resolution towards kilometric scale, better results can be obtained by downscaling model results to a kilometric resolution representative of urban scale by mixing model outputs with fine-scale information on emissions. Recent studies using CHIMERE at urban scale include the work of Markakis et al. (2015), using a set of long-term (10 year) CHIMERE simulations at 4 km horizontal resolution for the Paris region, including urban, suburban and rural areas, where the CHIMERE model is used for the present climate but also to test the possible impact of different emission and climate scenarios on air quality in this area. CHIMERE has also been used at continental scale for a long time, including model intercomparison exercises such as AQMEII (Rao et al., 2011; Solazzo et al., 2012b, a), Eurodelta (Schaap et al., 2007) and more recently Eurodelta III Bessagnet et al. (2016). The latter study presents the evaluation of the CHIMERE outputs for the main species of gaseous and particulate atmospheric trace components along with these of six other state-of-the-art models over Europe. The interested reader is therefore referred to Bessagnet et al. (2016) for a detailed comparison of the CHIMERE characteristics and performance compared to other models, and to Terrenoire et al. (2015) for a detailed overview of the CHIMERE performance and scores regarding the concentrations of many gaseous and aerosol species compared to a network of ground measurements over Europe for year 2009. As these studies at continental scale are very recent and dramatic changes in model performance over Europe do not occur from the changes presented here, the present article is not only focused on evaluating the model performance relative to observations but also on describing the generalization of the model scope to hemispheric scales and the inclusion of new processes. For forecasts, the model is applied daily for the French PREV AIR system, (Honoré et al., 2008), the COPERNICUS program, (Copernicus, 2017), as well as in many air quality networks.

In this paper, the CHIMERE-2017 model version is presented. All new developments made since the CHIMERE-2013 version (Menut et al., 2013a) are presented. This mainly consists in an extension of input databases, model grid management, optimization and chemical mechanism. The changes for the grid management are dedicated to build a CTM able to run over a hemispheric domains as well as for smaller regions anywhere in the world. These developments required important changes in the model, as well as the improvement of many processes already included in the previous version: the Fast-JX module for realistic evaluation of the photolysis rates has been added and allows for the calculation of updated photolysis rates at each physical time step, including the optical effects of clouds and aerosols. The mineral dust emissions have been upgraded in order to estimate fluxes in any region. In addition, this new version has also been an opportunity to update the representation of chemical processes by giving the user the choice to use the SAPRC chemical mechanism, which is more widely used than the MELCHIOR chemical scheme developed for the CHIMERE model (Lattuati, 1997; Menut et al., 2013a). Chlorine chemistry has been included, and the representation of physical processes affecting the aerosols, such as nucleation, coagulation and wet deposition, has been improved, while a scheme for traffic-related resuspension of particulate matter in urbanized areas has been included in the model.

CHIMERE-2017 is an offline chemistry-transport model, meaning that it needs to be provided with input meteorological fields, and does not implement any feedback of atmospheric chemistry on atmospheric dynamics. As the CHIMERE model is used for both analysis and forecast, particular attention was given to the optimization of computational performance. Numerous improvements were made in the code and are completely transparent for the user: these changes are described in Sect. 2.

Section 3 presents the changes in the model geometry, including the vertical mesh, as well as changes in the horizontal coordinate system allowing for the application of the model to hemispheric scale domains.

Section 4 presents the improvements in the representation of anthropogenic emissions, including the use of the global HTAP (hemispheric transport of atmospheric pollutants) emission dataset for anthropogenic emissions, and the improvements in modeling mineral dust emissions.

Section 5 describes the changes in the representation of various physical and chemical processes in the model, such as inclusion of the SAPRC scheme for gaseous chemistry and inclusion of chlorine chemistry in the model. This section also presents the evolutions in the modeling of the physical processes affecting aerosols, as well as the implementation of the Fast-JX module for radiative transfers. Another major improvement presented in this section is the ability of CHIMERE-2017 to provide lidar observables as a model output.

Section 6 presents the application of CHIMERE-2017 to simulations of 3 winter months and 3 summer months in a domain covering continental Europe at 50 km resolution, and the scores obtained by the model in comparison with back-
ground observations of gaseous and particulate species in this configuration.

Section 7 presents the application of the new model version to the simulation of the eruption of the Puyehue–Cordon Caulle volcano, in the Chilean Andes, in June 2011. This event provides a good test bed for this new version, since the volcanic plume from this volcanic eruption was dense enough to be observed by satellites all along its circumpolar transport around the South Pole.

Finally, Sect. 8 presents the conclusions of the present study, in terms of applications made possible by this new model version, as well as the outlines for future developments of the CHIMERE model.

2 Optimizations

Several technical changes were made in the CHIMERE code to improve code scalability: these changes regard the parallelization of many preprocessors into the parallelized section of the model, along with improvement of the parallelization strategy for some parts of the model that were already parallelized in order to improve code scalability.

2.1 Parallelization of preprocessors

Compared to the previous model version, several programs that used to be sequential preprocessors executed before the CHIMERE run itself have now been parallelized and included into the main CHIMERE executable. This is the case of the interpolation and treatment of the input meteorological fields. In the new model version, these fields are read and processed at each hourly time step (instead of being processed once and for all at the beginning of the run). This new design has no impact on the model outputs but has two advantages:

1. It allows a reduction of computation time by parallelization of this calculation step.

2. It enables the possibility to develop an online coupled version of the model, in which case the meteorological fields would not be pre-generated.

Note that this “real-time” processing of the meteorological fields is only available for users who use meteorological fields from the WRF (Weather Research and Forecast) model. For users of other sources of meteorological data, such as ECMWF (European Centre for Medium-Range Weather Forecasts) products, offline meteorological preprocessors are still provided with the model. Another important point is that even though the processing of meteorological input has been changed as described here, the version presented here does not take into account any radiative or microphysical feedback of atmospheric chemistry on meteorology. A version including aerosol–radiation interactions through online coupling of CHIMERE with WRF has been developed (Briant et al., 2017), and is available upon request from the lead author of that study. Apart from allowing online coupling between CHIMERE and WRF, the model setup described by Briant et al. (2017) also permits to update the meteorological fields at any time step shorter than 1 h.

Table 1 lists the variables that can be read by CHIMERE from the outputs of the meteorological model, separating the variables that are mandatory from the optional ones.

2.2 Improvement of the parallelization

In 2006, the main CHIMERE loop was parallelized using a master–slave pattern. A Cartesian division of the simulation domain into several sub-domains is done, each sub-domain being attributed to one slave process. Each slave performs the model integration in its own geographical sub-domain as well as boundary condition exchanges with its neighbors in order to permit transport from one slave to the next. In addition, in former CHIMERE versions, a master process was needed in order to gather and scatter data from the various slave processes that performed the actual gridded calculations, and to perform initializations and file input/output.

The use of a master process limited the efficiency of the parallelized code, since the master process did not perform any computation except gathering and scattering the data to and from the slaves, and that it totally centralized the input and output tasks, a bottleneck effect that limited the gains realized by parallelization, particularly when the simulation domains were very large and split between many slaves.

Therefore, in the CHIMERE-2017 version, this master process has been removed: using the parallel input/output routines of the parallel-netcdf library (Li et al., 2003), each slave process now reads the netcdf input files and writes the output data for its own sub-domain into a single output netcdf file common to all slaves, removing the bottleneck effect due to the centralization of input/output tasks.

This induces some major simplifications of CHIMERE code, including reduction of inter-process communications related to the parallelization of the input/output processes, which were performed in a central way by the master process in previous model version.

3 Model geometry

Major changes have been implemented in CHIMERE-2017 compared to earlier CHIMERE versions, opening the possibility to perform simulations in domains including the pole.

Historically, CHIMERE was first designed as a box model for the region of Paris (Menut et al., 2000). Rapidly, it has been transformed into a Cartesian model on curvilinear Arakawa C-grids (Arakawa and Lamb, 1977; see Fig. 1). However, the formulation of the transport scheme on these curvilinear grids up to CHIMERE-2014b was still based on a longitude–latitude (lat–long) formulation, which implied the
Table 1. Mandatory and optional variables obtained from meteorological input data. If the optional variables are not provided by the raw meteorological model, they are diagnosed during the simulation.

<table>
<thead>
<tr>
<th>CHIMERE name</th>
<th>Variable</th>
<th>Dimension</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>long</td>
<td>Longitude of grid points</td>
<td>2-D</td>
<td>°E</td>
</tr>
<tr>
<td>lat</td>
<td>Latitude of grid points</td>
<td>2-D</td>
<td>°N</td>
</tr>
<tr>
<td>tem2</td>
<td>2 m temperature</td>
<td>2-D</td>
<td>K</td>
</tr>
<tr>
<td>soim</td>
<td>Soil moisture</td>
<td>2-D</td>
<td>m^3 m^-3</td>
</tr>
<tr>
<td>rh2m</td>
<td>2 m relative humidity</td>
<td>2-D</td>
<td>0–1</td>
</tr>
<tr>
<td>lspc</td>
<td>Large-scale precipitation</td>
<td>2-D</td>
<td>kg m^-2 h^-1</td>
</tr>
<tr>
<td>copc</td>
<td>Convective precipitation</td>
<td>2-D</td>
<td>kg m^-2 h^-1</td>
</tr>
<tr>
<td>temp</td>
<td>Temperature</td>
<td>3-D</td>
<td>K</td>
</tr>
<tr>
<td>cliq</td>
<td>Cloud liquid water content (excluding rain water)</td>
<td>3-D</td>
<td>kg kg^-1</td>
</tr>
<tr>
<td>sphu</td>
<td>Specific humidity</td>
<td>3-D</td>
<td>kg kg^-1</td>
</tr>
<tr>
<td>pres</td>
<td>Pressure</td>
<td>3-D</td>
<td>Pa</td>
</tr>
<tr>
<td>alti</td>
<td>Altitude of half layer</td>
<td>3-D</td>
<td>m</td>
</tr>
<tr>
<td>winz</td>
<td>Zonal component of the wind</td>
<td>3-D</td>
<td>m s^-1</td>
</tr>
<tr>
<td>winm</td>
<td>Meridional component of the wind</td>
<td>3-D</td>
<td>m s^-1</td>
</tr>
<tr>
<td>swrd</td>
<td>Shortwave radiation</td>
<td>2-D</td>
<td>W m^-2</td>
</tr>
<tr>
<td>lwrd</td>
<td>Longwave radiation</td>
<td>2-D</td>
<td>W m^-2</td>
</tr>
<tr>
<td>sshf</td>
<td>Surface sensible heat flux</td>
<td>2-D</td>
<td>W m^-2</td>
</tr>
<tr>
<td>shlf</td>
<td>Surface latent heat flux</td>
<td>2-D</td>
<td>W m^-2</td>
</tr>
<tr>
<td>usta</td>
<td>Friction velocity</td>
<td>2-D</td>
<td>m s^-1</td>
</tr>
<tr>
<td>hght</td>
<td>Boundary-layer height</td>
<td>2-D</td>
<td>m</td>
</tr>
<tr>
<td>weas</td>
<td>Water equivalent accumulate snow</td>
<td>2-D</td>
<td>kg m^-2</td>
</tr>
<tr>
<td>snowh</td>
<td>Snow height</td>
<td>2-D</td>
<td>m</td>
</tr>
<tr>
<td>sealice</td>
<td>Sea-ice ratio</td>
<td>2-D</td>
<td>na</td>
</tr>
<tr>
<td>psfc</td>
<td>Surface pressure</td>
<td>2-D</td>
<td>Pa</td>
</tr>
<tr>
<td>rain</td>
<td>Rain water content</td>
<td>3-D</td>
<td>kg kg^-1</td>
</tr>
<tr>
<td>cice</td>
<td>Ice content</td>
<td>3-D</td>
<td>kg kg^-1</td>
</tr>
</tbody>
</table>

Figure 1. Centered (black) and staggered (blue and green) grid points in the Arakawa C-grid.
singular at the pole, to a 3-D Cartesian coordinate system, which has no singularity. In the former CHIMERE, versions the grid centers were represented by their geographical co-
ordinates \((\lambda_{ij}, \phi_{ij})\), and the wind vectors by their projection on the local frame \((u_{\lambda}, u_{\phi})\) (Fig. 2). In the present version, the points are represented by their Cartesian coordinates in the frame centered at the Earth center and with unit vectors \((u_1, u_2, u_3)\), and the wind vectors are represented by their projections on these unit vectors.

This change in the internal representation of spherical geometry has only a small impact on the simulated values, in the sense that it corrects some geometrical errors that appeared due to the assumptions made in the old coordinate system, but these differences have been found to be of very small amplitude, except in the vicinity of the pole where distortions due to the lat–long system become critical. The new coordinate system allows for domains that include the pole, without the need for any particular filtering. This strategy allows for the creation of regional domains from local to hemi-
spheric scale anywhere on the globe, including one pole or even, which opens possible application of CHIMERE-2017 for studies in the polar areas, including circumpolar trans-
pport of polluted air masses, as will be shown in Sect. 7. An example grid on which CHIMERE-2017 can be run is shown on Fig. 3. This grid is a polar stereographic grid centered at the north pole, entirely covering the Northern Hemisphere, and with the four corners of the domains extending slightly into the Southern Hemisphere (as far south as 19.47° S). With this projection and this number of points, the horizon-
tal model resolution varies from 140 × 140 km² at the pole to 70 × 70 km² at the Equator.

In this new coordinate system, the transport is calculated as follows. First, the coordinates of every grid center \(M_{ij}\)

\[
\begin{align*}
\lambda_{ij}^1 &= \cos \phi_{ij} \cos \lambda_{ij}^i \\
\lambda_{ij}^2 &= \cos \phi_{ij} \sin \lambda_{ij}^i \\
\lambda_{ij}^3 &= \sin \phi_{ij}
\end{align*}
\]

are converted from their geographical coordinates \((\lambda_{ij}, \phi_{ij})\) to Cartesian coordinates \((x_1^{ij}, x_2^{ij}, x_3^{ij})\) on a unit sphere as follows:

\[
\begin{align*}
\lambda_{ij}^1 &= \cos \phi_{ij} \cos \lambda_{ij}^i \\
\lambda_{ij}^2 &= \cos \phi_{ij} \sin \lambda_{ij}^i \\
\lambda_{ij}^3 &= \sin \phi_{ij}
\end{align*}
\]
The horizontal wind vector \( U^{ij} \) at the grid center is initially represented by the two classical wind components: \( U^{ij} = u^{ij} \cdot u + v^{ij} \cdot v \), where the zonal and meridional wind components \( u^{ij} \) and \( v^{ij} \) are obtained from the meteorological inputs. Since this representation splitting the horizontal wind into a zonal and a meridional component is singular at the geographical poles, before performing the transport operations, the horizontal wind is split into its three components on the Cartesian frame \((u_1, u_2, u_3)\) using the following formulae for projecting the wind on the Cartesian frame:

\[
\begin{align*}
U^{ij}_1 & = -\sin \lambda \cdot u^{ij} - \sin \phi \cdot \cos \lambda \cdot v^{ij} \\
U^{ij}_2 & = \cos \lambda \cdot u^{ij} + \sin \phi \cdot \sin \lambda \cdot v^{ij} \\
U^{ij}_3 & = \cos \phi \cdot v^{ij}
\end{align*}
\]  

Once the Cartesian coordinates of the grid centers \((x_1^{ij}, x_2^{ij}, x_3^{ij})\) and of the wind-speed vectors \((U^{ij}_1, U^{ij}_2, U^{ij}_3)\) are computed at the grid centers, it is easy to obtain the values of the speed vectors at the staggered cells (Fig. 1) with the following formulae:

\[
\begin{align*}
U^{i+\frac{1}{2},j}_k & = U^{ij}_k + U^{i+1,j}_k (k = 1, 2, 3) \\
U^{i,j+\frac{1}{2}}_k & = \frac{1}{2} (U^{ij}_k + U^{i+1,j+1}_k) (k = 1, 2, 3) .
\end{align*}
\]  

This new formulation with the use of Cartesian coordinates instead of geographical lat–long coordinates for the transport of pollutants removes the constraints that prevented the use of CHIMERE on domains including a geographic pole and/or a date-change line. This new formulation has been tested on the case of the eruption of the Puyehue volcano, in June 2011, a case during which the ash plume from the volcano went around the South Pole through the southern Atlantic, Pacific and Indian oceans back to South America after 15 days (Sect. 7). This case is a perfect test bed for the ability of the model to simulate circumpolar movements, and evaluate its ability to represent the location of an aerosol plume after several days/weeks of travel.

### 3.1 Vertical mesh calculation

The vertical discretization of CHIMERE needs to obey 2-fold requirements. First, as it has been the case since the beginning of the development of the model, the vertical mesh needs to be very refined in the lowest atmospheric layers because these layers are critical for the modeling of boundary-layer contamination, particularly in urban areas, but also in marine areas with sea-salt emissions, and in arid areas with mineral dust emissions. On the other hand, the CHIMERE model is now used not only for studies at urban/regional scale, but also for studies at continental and, from the present version, hemispheric scale. Therefore, a relatively fine vertical resolution is also needed in the free troposphere to be able to simulate the transport of trace gases and aerosols over large distances avoiding excessive numerical diffusion. Therefore, due to these two requirements, the CHIMERE-2017 vertical mesh is defined as described below.

Regarding the vertical discretization, the user has three degrees of freedom:

- The thickness of the first layer. The user can fix the top of the first model layer, by setting the top of the first model layer in sigma coordinates: \( \sigma_1 = 0.997 \) corresponds to a thickness of about 3 hPa for the first model layer, about 30 m.

- The number of layers, typically from 8 to 20 layers for the most common configurations of the model.

- The pressure of the top of the model, \( p_{top} \), can be freely set by the user with typical values from 500 hPa for studies at urban/regional scales to 100 hPa for continental/hemispheric-scale studies.

From these user-defined parameters, a preprocessing tool calculates a vertical grid as follows:

- From the surface to 800 hPa, the layer thickness (in hPa) increases exponentially.

- From 800 hPa to the top of model, the layers are evenly distributed, with equal thickness for each layer.

This procedure outputs the pressure of the level tops, for a reference surface pressure \( p_{ref} \) of 1000 hPa. However, the model levels need to adapt themselves to the variations of the surface pressure, essentially due to orography. This is ensured by scaling linearly the pressure levels between the surface pressure and the pressure at the top of model, \( p_{top} \), producing two sequences of coefficients \( a_i \) and \( b_i \), such that the pressure at the top of layer \( i \) is given by \( p_i = a_i p_{ref} + b_i p_{surf} \). These coefficients are given by the following expressions:

\[
a_i = \frac{p_{top} (p_i - p_1)}{p_{ref} (p_1 - p_{top})} ,\]  

\[
b_i = \frac{p_1 (p_i - p_{top})}{p_{surf} (p_1 - p_{top})} .\]  

The linear scaling of the pressure levels by these two sequences of coefficients ensures that the pressure levels never cross each other, and that their relative thickness stays the same even above high topography, as shown in Fig. 4. Vertical transport on this mesh can be calculated using either a slope-limited Van Leer scheme (Van Leer, 1979) or a upwind scheme, depending on user’s choice, also taking into account turbulent mixing and, optionally, deep-convection fluxes, following the Tiedtke (1989) formulation.
Figure 4. Model pressure levels with 20 vertical levels: thickness of the first model layer is 3 hPa, top of model set at 200 hPa. Pressure levels are represented across an idealized mountain with a top at 500 hPa.

4 Emissions

4.1 The anthropogenic emissions

4.1.1 Overall description

CHIMERE needs to be forced at least by input meteorological fields, and by anthropogenic emissions. A preprocessor for anthropogenic emissions, named emisurf, is provided to the users. This preprocessor was historically developed for the downscaling and reformatting of the raw emissions from the EMEP (European Monitoring and Evaluation Program) emission inventory at 50 km resolution, but can be adapted by users to any other raw dataset they need to use. The main steps for this are described in Menut et al. (2012):

- A first step projects the annual masses from the “raw” EMEP grid to the CHIMERE grid. The spatial emission distribution from the EMEP grid to the CHIMERE grid is performed using proxies like population density, as described by Fig. 5a–d. Proxies used by emisurf for this process include land-use data (either GLCF, USGS or GlobCover), large point source database (such as the EPER database for Europe), etc.

- Second, monthly, weekly and hourly profiles are prescribed to convert annual totals to hourly fluxes used as input for CHIMERE. These factors are derived largely from data provided by the University of Stuttgart (IER) as part of the GENEMIS project (Friedrich and Reis, 2004), and are available as data files from the EMEP model website, www.emep.int.

- A last step consists in converting the species available in the raw data into the model species. Generally, a minimum of seven species are available: CO, SO\(_x\), NO\(_x\), NH\(_3\), NMVOC (non-methane volatile organic compounds), PM\(_{2.5}\) and PM\(_{coarse}\) (difference between PM\(_{10}\) and PM\(_{2.5}\)). In CHIMERE, depending on the chemical scheme, about 30 species are emitted. NO\(_x\) is split into NO, NO\(_2\) and HONO. Usually, 5 to 10 % is assigned for NO\(_2\) emissions for all sectors, except for traffic emissions where 20 % should be assigned to NO\(_2\) for modern fleets (post-2010). For NMVOC, the data used are derived from the detailed United Kingdom specification given in Passant (2002). For SO\(_x\), 99 % is assigned to SO\(_2\) and 1 % for primary sulfate to account for very fast and local sulfate production. The lumping procedure accounts for the reactivity of VOC species following Middleton et al. (1990).

The vertical distributions were originally based upon plume-rise calculations performed for different types of emission sources, which are thought typical for different emission categories, under a range of stability conditions (Vidic, 2002), but have since been simplified and adjusted to reflect the more recent findings of (Bieser et al., 2011). The main changes have been for the residential sector where now 100 % of the emissions are placed in the lowest 20 m of the atmosphere, reflecting the large dominance of domestic combustion for this emission category. Also, emissions from large combustion facilities in SNAP (Selected Nomenclature for Air Pollutants) sectors 1 and 4 corresponding to large industrial facilities burning fossil fuels are attributed to lower layers than in Vidic (2002), resulting in enhanced concentrations of primary species such as NO\(_x\) and SO\(_x\) in the boundary layer, in better agreement with routine surface observations, as discussed in Mailler et al. (2013). The vertical distribution profiles that are used for each SNAP sector are constant profiles depending only on the SNAP sector, and are presented in Terrenoire et al. (2015).

4.1.2 Recent changes

The main recent changes have been focused on the use of proxies to better reallocate in space the raw emissions. This specialization can be performed from the raw gridded data or directly from the annual country totals (Terrenoire et al., 2015).
The European Pollutant Release and Transfer Register (E-PRTR) data are used to precisely place the emissions from the main industrial sources. E-PRTR is the Europe-wide register that provides easily accessible key environmental data from industrial facilities in European Union member states and in Iceland, Liechtenstein, Norway, Serbia and Switzerland.

To treat road traffic emissions at the European scale, a spatial proxy to distribute the annual country emissions has been developed. This proxy provides a unitless value for a given cell at 1 km resolution over Europe. It is built by crossing several databases (population, land cover data, roads, etc.); it consists of a linear regression of several parameters such as population density, length of road, and surface of urban areas in a given fine grid cell. The regression coefficients are calculated over France thanks to the use of the French high-resolution bottom-up inventory and applied everywhere over Europe (Fig. 6).

For the extrapolation at the European level, it uses the best source of information among the following proxies: CORINE land cover (from the European Environment Agency), road data of the ETISplus European project (European Transport policy Information System) for 2010 over Europe. ETISplus combines data, analytical modeling with maps (GIS) and a single online interface for accessing the data. Default European GIS road data from EuroglobalMap, default worldwide GIS road data from natural Earth data\(^1\), and population database by Gallego (2010) over Europe and data from Center for International Earth Science Information Network (CIESIN) for the rest of the world. All of these data were not available on the whole domain. Therefore, three tiers of information were defined to cover all countries with different levels of confidence:

- Countries covered by all the data: Iceland, Norway, Turkey, Bosnia Herzegovinia, Serbia, Montenegro, Kosovo, Macedonia, Albania and all the EU28 except Greece.

- Countries without CLC coverage but with ETIS or EuroglobalMap data: Belarus, Ukraine, Moldavia and Greece.

- Other countries are only covered by the world road map and population data.

\(^1\)http://www.naturalearthdata.com/
For shipping emissions (SNAP 8), a proxy was developed using an inventory of shipping routes obtained from the US National Center for Ecological Analysis and Synthesis. A database of pressure on marine ecosystems has been developed for the year 2008 by Halpern et al. (2015) but the dataset remains non-exhaustive, the data being collected only on voluntary vessels.

4.2 Mineral dust emissions

Mineral dust modeling is an important process for understanding climate evolution but also for air quality regional modeling. For many regions over the world, it becomes necessary to manage air pollution knowing the relative part of anthropogenic and natural contributions. For this, even over small regions, it is important to have the same level of knowledge for mineral dust emissions as for anthropogenic or biogenic emissions. In this new model version, many improvements were done for mineral dust emissions. They are related to input databases, the emission schemes themselves and additional options to better take into account the impact of meteorological conditions on emissions.

4.2.1 Soil, land use and roughness length

For the calculation of mineral dust emissions, several variables have to be known: land use, soil characteristics, aeolian roughness length and erodibility. Originally, CHIMERE used a database limited to North Africa and the Arabian Peninsula. For simulations over Africa or Europe, this spatially limited database was considered adequate, Sahara being the major source in this region. But for this new CHIMERE-2017 version, the goal is to enable calculations of mineral dust emissions anywhere in the world. A large part of this change was already done in Menut et al. (2013b) for land use, soil and roughness length. The soil and land use used are now those from NCAR USGS land-use dataset (Homer et al., 2004) and STATSGO-FAO soil dataset (Wolock, 1994). The roughness length is estimated using the global 6 km horizontal resolution “Global Aeolian Roughness Lengths from ASCAT and PARASOL” dataset (Prigent et al., 2012).

In addition to these changes, the option to evaluate the soil erodibility based on satellite data was added. Therefore, three options are now available in CHIMERE 2017:

1. Calculate the erodibility from the land-use database: cropland, grassland, shrubland and barren or sparsely vegetated areas, are then considered as partly erodible. This was the only option offered in earlier CHIMERE versions. In this case, constant percentages are applied for each land-use category.

2. Use the global erodibility dataset derived from MODIS (Grini et al., 2005), included and used in CHIMERE as described by Beegum et al. (2016).

3. Use a mix between these two strategies, using MODIS only over desert areas and the USGS land uses categories elsewhere.

4.2.2 The Kok’s scheme for mineral dust emissions

In this model version, the Kok mineral dust emissions parameterization is proposed, in addition to the Marticorena and Bergametti (1995) and Alfaro and Gomes (2001) schemes. The Kok scheme is fully described in the articles Kok et al. (2014b), Kok et al. (2014a) and Mahowald et al. (2014). The vertical dust flux is calculated as

\[
F_d = C_d f_{\text{bare}} f_{\text{clay}} \frac{\rho_a (u_*^2 - u_{st*}^2)}{u_{st*}^2} \left( \frac{u_a}{u_{st*}} \right)^C \frac{\nu + \nu_{\text{mod0}}}{\nu_{\text{mod0}}} .
\]  

(6)
where $f_{\text{bare}}$ and $f_{\text{clay}}$ represent the relative fraction of bare soil and clay soil content, respectively. The flux is calculated only if $u > u_{st}$. The threshold friction velocity, $u_{st}$, is calculated using the Iversen and White (1982) or the Shao and Lu (2000) scheme (a user’s choice). The corresponding $u_{st}$ is this friction velocity but for a standard atmospheric density $\rho_0 = 1.225 \text{ kg m}^{-3}$.

$$u_{st} = u_{st} \left( \frac{\rho_s}{\rho_0} \right),$$

where $u_{st(0)}$ represents $u_{st}$ for an optimally erodible soil and was chosen as $u_{st(0)} = 0.16 \text{ m s}^{-1}$ in Kok et al. (2014b). The dimensionless coefficient $C_k$ is chosen as $C_k = 2.7$.

The dust emission coefficient $C_d$ represents the soil erodibility as

$$C_d = C_{d0} \exp \left( -C_e \frac{u_{st} - u_{st(0)}}{u_{st(0)}} \right),$$

with the constant dimensionless coefficients $C_e = 2.0$ and $C_{d0} = 4.4 \times 10^{-5}$.

The vertical dust flux is integrated over the whole size distribution. This flux is thus redistributed into the model dust size distribution as

$$\frac{dV_d}{d\ln D_d} = \frac{D_d}{C_v} \left[ 1 + \text{erf} \left( \frac{\ln(D_d/D_s)}{\sqrt{2} \ln \sigma_s} \right) \right] \exp \left[ -\left( \frac{D_d}{\lambda} \right)^3 \right].$$

where $V_d$ is the volume of mineral dust aerosols for each mean mass median diameter $D_d$, $C_v = 12.62 \mu\text{m}$, $\sigma_s = 3.0$, $D_s = 3.4 \mu\text{m}$ and $\lambda = 12.0 \mu\text{m}$.

### 4.2.3 Impact of vegetation on dust emissions

The vegetation evolves during the year and this variability will impact the mineral dust emissions. Contrarily to the previous model version, more focused on Saharan areas, this version is able to model mineral dust all around the world. For example in areas such as the Sahelian region or Europe, mineral dust are observed but are very dependent on the vegetation variability. To take into account this variability, the vegetation fraction is diagnosed from the USGS 30 s resolution database and acts as a limiter to the erodibility factor.

### 4.2.4 Impact of rain on dust emissions

The possibility to inhibit or moderate dust erosion in case of rainfall was improved in this model version. In the previous model versions, the complete inhibition of mineral dust emissions during a rainfall event was already considered. In this version, a “rain memory function” was added in order to take into account the possible crusting of the soil (Ishizuka et al., 2008) and thus the fact that emissions are also reduced after a rainfall event. For this calculation, a simple factor $f_p$ is applied to moderate the dust emissions fluxes when a precipitation is diagnosed and during the next hours as

$$f_p = E_{\text{dust}} \left( 1 - \exp \left( \frac{-2\pi \Delta t_p}{\tau} \right) \right),$$

where $\Delta t_p$ is the time since the last precipitation event and $\tau$ the period after which the surface mineral dust fluxes $E_{\text{dust}}$ is fully taken into account, considering that the inhibiting effect of precipitation is finished. For this study, $\Delta t_p$ is in hours and $\tau = 12$. This function is displayed in Fig. 7.

### 4.2.5 Impact of soil moisture on dust emissions

In the absence of precipitation, the soil moisture may also inhibit mineral dust erosion. This effect is taken into account using the Fécan et al. (1999) parameterization. This scheme considers that soil moisture will increase the threshold friction velocity, $u^T$, used to determine if erosion occurs or not. To distinguish between soil conditions, the dry and wet threshold friction velocities are defined, and noted $u^T_d$ and $u^T_w$, respectively. $u^T_w$ is estimated as a possible increase of $u^T_d$ depending on the modeled gravimetric soil moisture $w$ (in kg kg$^{-1}$):

$$u^T_w = f(w) u^T_d.$$ (11)

In the model, the dry threshold friction velocity, $u^T_d$ is calculated following the scheme of Shao and Lu (2000). The $f(w)$ factor is estimated as

$$f(w) = \begin{cases} 1 & \text{for } w < w' \\ \left[ 1 + A (w - w')^{b'} \right]^{0.5} & \text{for } w > w' \end{cases},$$ (12)

where $A$ and $b'$ are constants to estimate, and $w'$ corresponds to the minimum soil moisture from which the threshold velocity increases. The values of $A$, $b'$ and $w'$ are dependent on the soil texture. For $A$ and $b'$, the values are fixed to $A = 1.21$ and $b' = 0.68$. Using measurements data, Fécan et al. (1999)
showed that the value of \( w' \) is mainly dependent on the clay content of the soil and proposed the following fit:

\[
 w' = 0.0014 (\% \text{ clay})^2 + 0.17 (\% \text{ clay}). \tag{13}
\]

Note that in Eq. (12), the gravimetric soil moisture \( w \) has to be expressed in %, \( w' \) being in % in Eq. (13) (a conversion is done from kg kg\(^{-1}\) to %).

### 4.3 Traffic-related resuspension

The resuspension process is important for particulate matter and may induce a large increase of the emission flux in case of dry soils, for locations where traffic and industries produce particles that may be deposited on the ground and therefore become available for resuspension. In this model version, the resuspension flux is active only for cells containing an urbanized surface. This flux is applied as primary particulate matter resuspension flux is active only for cells containing an urbanized surface. This flux is applied as primary particulate matter resuspension flux is activated, and a constant tuned in order to approximately close the PM\(_{10}\) emissions only and thus considered in the model as an anthropogenic process.

The formulation is derived from the bulk formulation originally proposed by Loosmore (2003). The resuspension rate \( \lambda \), in \( \text{s}^{-1} \), is expressed as

\[
 \lambda = 0.01 \frac{u^1.43}{\tau^{1.03}}, \tag{14}
\]

where \( \tau \) is the time after the start of resuspension. This time is taken into account considering that particles are first deposited then resuspended. The detail of the processes leading to resuspension are essentially unknown, and we assume here that the available concentration of particulate matter depends only on the wetness of the surface. In this empirical view, the resuspension flux is assumed to be

\[
 F = P f(w) u_s^{1.43}, \tag{15}
\]

where \( f(w) \) is a function of the soil water content and \( P \) is a constant tuned in order to approximately close the PM\(_{10}\) mass budget over Europe estimated in Vautard et al. (2005). It was found to give a correct amount of additional PM\(_{10}\).

In this model version, \( P \) is approximated as \( P = 4.72 \times 10^{-2} \mu \text{g m}^{-2} \text{s}^{-1} \) if we consider European mean conditions with a soil water content of 25% and a friction velocity of \( u_s = 0.5 \text{ m s}^{-1} \).

The soil water function \( f(w) \) is estimated as

\[
 f(w) = \frac{w_s - w}{w_s - w_t}, \tag{16}
\]

where \( w_t = 0.1 \) is a soil moisture threshold below which resuspension is activated, and \( w_s \) is the maximum of soil moisture ponderated by the ratio of water and soil densities as

\[
 w_s = w_{\text{max}} \frac{D_{\text{water}}}{D_{\text{soil}}}, \tag{17}
\]

where \( w_{\text{max}} = 0.3 \) is a constant value representing the maximum soil moisture value, \( D_{\text{water}} \) is the water density (assumed to be unity) and \( D_{\text{soil}} \) is the dry porous soil density.

\( D_{\text{soil}} \) is itself estimated as

\[
 D_{\text{soil}} = (1 - \text{satsm}) D_{\text{mine}}, \tag{18}
\]

where \( \text{satsm} = 0.4 \) is the saturation volumetric moisture content and \( D_{\text{mine}} = 2.5 \), the non-porous soil density.

This resuspension flux is calculated only for model cells having a non-zero urban land use. This flux is thus ponderated in the whole cell by considering the relative surface of the urban area. Finally, the flux is projected onto the model size distribution considering that two-thirds of the flux is in the fine mode, one-third in the coarse mode. The fine and coarse modes are those defined for the anthropogenic emissions fluxes for particulate matter.

### 5 Processes and chemistry

#### 5.1 Integration of the SAPRC chemical scheme

##### 5.1.1 The general gas-phase mechanism

Two gas-phase chemical schemes were implemented in the CHIMERE model. The most detailed chemical scheme, called MELCHIOR1, represents the oxidation of around 80 gaseous species according to 300 reactions. The other mechanism, called MELCHIOR2, is a reduced version of MELCHIOR1 developed using chemical operators (Derognat et al., 2003; Carter, 1990). MELCHIOR2 represents the oxidation of around 40 gaseous species according to 120 reactions. These chemical mechanisms are described in detail in Menut et al. (2013a). Comparisons between MELCHIOR2 and three detailed mechanisms (MCM, Jenkin et al., 2003; SAPRC99, Carter, 2000; GECKO-A, Aumont et al., 2005) show a good agreement between the chemical schemes, with differences in HCHO yields under low- and high-NO conditions lower than 20% between the simulated results (Dufour et al., 2009). SAPRC99 chemical mechanism had already been used in CHIMERE for particular studies (Lasry et al., 2007; Coll et al., 2009) but had never been distributed in a previous CHIMERE release.

Since the development of the MELCHIOR mechanisms in 2003, progress has been made in atmospheric chemistry, particularly concerning the VOC ozonolysis. One of the most up to date chemical schemes currently available in the literature is the SAPRC-07 (Carter, 2010a). This mechanism is widely used and evaluated against chamber data (≈ 2400 experiments). The detailed SAPRC-07 chemical mechanism contains 207 species and 466 reactions. This detailed mechanism has been used to develop several reduced mechanisms designed for CTM applications (Carter, 2010b). The less reduced mechanism, SAPRC-07A, has been implemented in the 2016 CHIMERE model. This chemical scheme contains 72 species and 218 reactions. Two CHIMERE simulations using SAPRC-07A and MELCHIOR2 chemical schemes respectively were compared with AirBase measurements of
NO$_x$ and ozone over Europe during summer 2005. The two chemical schemes were found to provide good correlation with ozone measurements (Pearson’s correlation rate 0.71 for both mechanisms), with a slightly smaller bias for ozone concentrations obtained using SAPRC-07A (8.19 ppb vs. 9.29 ppb, Menut et al., 2013a).

5.1.2 The chlorine mechanism

Over the past decade, several studies have shown that halogens (chlorine, bromine, iodine) chemistry could influence ozone concentrations in the troposphere. A recent review by Simpson et al. (2015) presents the state of art on this topic.

The role of halogen chemistry was traditionally considered limited to the marine boundary layer, recent observations have shown significant ClNO$_2$ concentrations from few parts per trillion in mid-continental urban environment (Mielke et al., 2011) to 2000 ppt in the coastal marine boundary layer (Riedel et al., 2011). This compound can act as a nitrogen reservoir with a long lifetime capable of long-range transport. In previous versions of CHIMERE, it was possible to have the chemical composition (Na, Cl, H$_2$SO$_4$) of sea-salt emissions based on mean composition described in Seinfeld and Pandis (1997). The chlorine chemistry is not described in MELCHIOR chemical schemes but Carter (2010b) proposed in SAPRC-07A a chlorine mechanism with nine inorganic species and three products formed by the reactions with VOCs. In SAPRC-07A, the chlorine chemistry is represented by 68 reactions, which have been implemented in CHIMERE-2017 only if the SAPRC-07A mechanism is chosen by the user.

5.2 Evolution of the aerosol scheme

5.2.1 Discretization of the aerosols size distribution

The CHIMERE model accounts for the size distribution of the aerosols using a size-bin approach: the aerosol particles for each of the model species are distributed in $N$ size bins, covering a diameter range from $D_{\text{min}}$ to $D_{\text{max}}$. Given these three user-defined parameters, a preprocessor computes a sequence $(d_i)_{i=1,N+1}$ of cut-off diameters that meets the following requirements:

- 2.5 and 10 µm are retained as cut-off diameters: two indices $i_1$ and $i_2$ such that $d_{i_1} = 2.5$ µm and $d_{i_2} = 10$ µm must exist.

- The sequence of the cut-off diameters covers exactly the size interval requested by the user: $d_1 = D_{\text{min}}$ and $d_{N+1} = D_{\text{max}}$.

The first requirement is set to allow for a meaningful evaluation of PM$_{2.5}$ and PM$_{10}$ in the model, since these quantities are typically available from routine measurements.

The default (and recommended) values of the extreme diameters are $D_{\text{min}} = 0.01$ µm and $D_{\text{max}} = 40$ µm. Using these values, the produced size distributions for various values of the number of intervals $N$ are shown in Table 2 according to the requested number of bins, $N$. If $N \geq 12$, then the ratio of two successive cut-off diameters is always such as $d_{i+1}/d_i \leq 2$; all particles within a single size bin have comparable diameters at least within a factor 2, which is a good way to ensure that all the size-depending processes affecting the aerosols (sedimentation, coalescence, etc.) are treated in a realistic way. However, when calculation speed is a critical requirement, for example for operational pre-vision, the number of size bins could be lowered to $N = 6$, still ensuring that $d_{i+1}/d_i \leq 4$.

5.2.2 Wet diameter and density of aerosols

In many processes, the diameter and the density of aerosols are used (deposition, absorption, coagulation, etc.). These processes have to take into account that the diameter and the density of aerosols change with humidity due to the amount of water absorbed into the particles. Therefore, the notion of wet diameter and wet density was introduced in CHIMERE-2017. Particles are distributed between bins according to their dry diameter. The wet diameter of the particles is calculated as a function of humidity and the composition of the particle.

To compute the wet density and wet diameter for each aerosol size bin, the amount of water in each bins is computed with the “reverse mode” of ISORROPIA (Nenes et al., 1998) by using the composition of particles, assuming that only sulfate, nitrate, ammonium and sea salts have a high enough hygroscopicity to absorb a significant amount of water. The density of the aqueous phase of particles is computed according to composition following the method of Semmler et al. (2006). The density and mass of the inorganic aqueous phase (sulfate, nitrate, ammonium and sea salts and water) and the density and mass of other compounds (dust, organic, black carbon, etc.) are used to compute the total density of the particle and then its wet diameter, assuming internal mixing for each size bin.

5.2.3 Absorption

Absorption is described by the “bulk equilibrium” approach of Pandis et al. (1993). In this approach, all the bins for which condensation is very fast are merged into a “bulk particulate phase”. Following Debry et al. (2007), a cutting diameter of 1.25 µm is used to separate bins, which are inside the “bulk particle” (with a diameter lower than the cutting diameter) from other bins.

Thermodynamic models are used to compute the partitioning between the gas phase and the bulk particle phase and estimate the gas-phase concentrations at equilibrium. For semi-volatile inorganic species (sulfate, nitrate, ammonium), concentrations $G_{\text{eq}}$ at equilibrium are calculated using ISORROPIA. This model also determines the water content of
Table 2. Values of the diameter intervals, $I_v$ (µm), obtained for $D_{\text{min}} = 0.01$ µm, $D_{\text{max}} = 40$ µm, and 14 different values of bins ($N = 3$ to $N = 16$).

<table>
<thead>
<tr>
<th>$I_v$</th>
<th>$N = 3$</th>
<th>$N = 4$</th>
<th>$N = 5$</th>
<th>$N = 6$</th>
<th>$N = 7$</th>
<th>$N = 8$</th>
<th>$N = 9$</th>
<th>$N = 10$</th>
<th>$N = 11$</th>
<th>$N = 12$</th>
<th>$N = 13$</th>
<th>$N = 14$</th>
<th>$N = 15$</th>
<th>$N = 16$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>2.50</td>
<td>0.16</td>
<td>0.06</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>10.00</td>
<td>2.50</td>
<td>0.40</td>
<td>0.16</td>
<td>0.09</td>
<td>0.09</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>4</td>
<td>40.00</td>
<td>10.00</td>
<td>2.50</td>
<td>0.63</td>
<td>0.27</td>
<td>0.27</td>
<td>0.16</td>
<td>0.11</td>
<td>0.11</td>
<td>0.08</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>2.50</td>
<td>0.83</td>
<td>0.83</td>
<td>0.40</td>
<td>0.23</td>
<td>0.23</td>
<td>0.16</td>
<td>0.12</td>
<td>0.12</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>6</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
<td>0.63</td>
<td>0.40</td>
<td>0.40</td>
<td>0.27</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
<td>0.63</td>
<td>0.40</td>
<td>0.40</td>
<td>0.27</td>
<td>0.20</td>
</tr>
<tr>
<td>8</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
<td>0.63</td>
<td>0.40</td>
<td>0.40</td>
<td>0.27</td>
</tr>
<tr>
<td>9</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
<td>0.63</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>10</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
<td>0.63</td>
<td>0.40</td>
</tr>
<tr>
<td>11</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
<td>0.63</td>
</tr>
<tr>
<td>12</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
<td>1.14</td>
</tr>
<tr>
<td>13</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
<td>1.14</td>
</tr>
<tr>
<td>14</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
<td>2.50</td>
</tr>
<tr>
<td>15</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
<td>5.00</td>
</tr>
<tr>
<td>16</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>40.00</td>
<td>10.00</td>
</tr>
<tr>
<td>17</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>4.00</td>
</tr>
</tbody>
</table>

If a particle shrinks or grows due to condensation/evaporation, the mass of this particle has to be redistributed over diameter bins. The mass redistribution algorithm of Gelbard and Seinfeld (1980); Seigneur (1982) is used.

5.2.4 Coagulation

The flux of coagulation $J_{\text{coag},b,i}$ of a compound $i$ inside a bin $b$ is computed with the size binning method of Jacobson et al. (1994):

$$ J_{\text{coag},b,i} = \sum_{j=1}^{b} \sum_{k=1}^{b} f_{j,k} b K_{j,i} A_{p,j}^b N^k - A_{p,j}^b \sum_{k=1}^{b} K_{b,j} N^k, \quad (22) $$

where $N^k$ is the volumic number of particles in bin $k$, $K_{j,i}$ the coagulation kernel coefficient between bins $i$ and $j$ and $f_{j,k}$ the partition coefficient (the fraction of the particle created from the coagulation of bins $j$ and $k$, which is redistributed into bin $b$). The coagulation kernel and the partition coefficients are calculated as described in Debry et al. (2007).

5.2.5 Wet deposition

For the in-cloud scavenging of particles, the deposition of particles is assumed to be proportional to amount of water lost by precipitations. The deposition flux is written as

$$ \left[ \frac{dQ_i^k}{dt} \right] = -\frac{\epsilon_l P_t}{w_l h} Q_i^k, \quad (23) $$

where $P_t$ is the precipitation rate released in the grid cell (kg m$^{-2}$ s$^{-1}$), $w_l$ the liquid water content (kg m$^{-3}$), $h$ the cell thickness (m) and $\epsilon_l$ an empirical uptake coefficient (in the
For the below-cloud scavenging of particles, particles are scavenged by raining drops following Henzig et al. (2006). A polydisperse distribution of raining drops is applied:

\[ N(R) = 1.98 \times 10^{-5} A R^{-0.384} R^{2.93} \exp \left( -5.38 P^{-0.186} R \right), \]  

where

\[ A = 1.047 - 0.0436 \ln P + 0.00734 (\ln P)^2, \]

where \( P \) is the precipitation rate in mm h\(^{-1} \) and \( R \) the radius of the droplet. The below-cloud scavenging rate is written as

\[
\frac{\mathrm{d}Q(t)}{\mathrm{d}t} = -Q(t) \int_{R}^{R_{\text{d}}} \pi R^{2} u_{g}(R) E(R, r) N(R) \, \mathrm{d}R,
\]

where \( R \) is the radius of the raindrop (in m), \( r \) the radius of the particle (in m), \( u_{g} \) the terminal drop velocity (in m s\(^{-1} \)), \( E(R, r) \) the collision efficiency of a particle with a raindrop and \( N(R) \) (in m\(^{-1} \)) the raindrop size distribution.

### 5.3 Online calculation of photolysis rates using the Fast-JX module

#### 5.3.1 Modeling strategy

CHIMERE-2017 includes the module Fast-JX version 7.0b (Wild et al., 2000; Bian and Prather, 2002) for the online calculation of the photolysis rates. Fast-JX is a module that solves the equations of radiative transfer in an atmospheric column taking into account the solar zenith angle, the vertical profile of ozone and water-vapor concentrations, the ice and water clouds, the radiative effect of scattering and absorption by aerosols and the surface albedo.

Following the recommendations of the Fast-JX developers, the effective size of ice particles is estimated following Heymsfield (2003) as \( R_{\text{eff}_{i}} = 164 \times \text{IWC}^{3/5} \), where \( R_{\text{eff}_{i}} \) (\( \mu m \)) is the effective radius of ice particles, and IWC is the ice content of the atmospheric particles (g m\(^{-3} \)). Regarding water droplets, their radius is estimated also following the recommendations of Fast-JX developers, as 9.60 \( \mu m \) for clouds at low altitudes (below 810 hPa), 12.68 \( \mu m \) for high clouds (above 610 hPa), and linearly interpolated between these two values for intermediate altitudes.

Taking these factors (and their real-time simulated variations) into account, Fast-JX computes the photolysis rates for all the relevant photochemical reactions that have been designed in order to be easily introduced in chemistry-transport models, which has already been done in various CTMs such as PHOTOMCAT (Voulgarakis et al., 2009), Polair3D (Real and Sartelet, 2011), UKCA (Telford et al., 2013) and GEOS-Chem (Eastham et al., 2014).

CHIMERE-2013 did not take into account all of these processes (Menut et al., 2013a), relying instead on a very simplified calculation of the photolysis rates, as shown in Table 3. The photolysis rates were evaluated from tabulated values using TUV (Madronich, 1987), depending only on the solar zenith angle and the altitude. These tabulated values were calculated assuming a vertical profile for ozone that was typical of the Northern Hemisphere midlatitudes, neglecting the effect of the aerosols, and assuming a constant and uniform surface albedo. The effect of clouds was parameterized as an exponential reduction of the photolysis rates as a function of the cloud optical depth. While this set of approximations was acceptable when the CHIMERE model was used as boundary-layer regional CTM for locations in Europe, this had strong limitations for its use for longer-term simulations including long-range transport in the free troposphere over geographical domains including polar and/or tropical zones. Photolysis rates for the photodissociation of ozone and nitrogen dioxide as computed by the Fast-JX model inside CHIMERE have been compared favorably to in situ measurements at the island of Lampedusa (Italy), even in presence of aerosols (Mailler et al., 2016).

#### 5.3.2 Surface albedo

The surface albedo in the near-UV spectral region, which is determinant for the calculation of photolysis rates (Dickerson et al., 1982), is highly variable according to the land use and to the presence or absence of snow. It is worth noting that the albedo of all the continental and oceanic surfaces is smaller than 0.1, while the albedo of snow ranges from 0.3 to over 0.8 according to the type of land use. Therefore, the absence/presence of snow will modulate very substantially the values of the modeled photolysis rates, and therefore the concentration of trace gases such as ozone. Even though strong ozone peaks generally occur in summertime in a context of strong anthropogenic NO\(_X\) production and in the absence of snow, it has been shown recently that strong ozone peaks can occur in wintertime over the continental United States in zones of oil and gas extraction due to the combination of the strong anthropogenic concentrations of VOCs in a very shallow boundary layer with relatively strong photolysis rates due to the high surface albedo (Edwards et al., 2014; Schnell et al., 2009). It is therefore important that CTMs take into account the impact of snow on surface albedo, in order to be able to reproduce correctly such cases.

The surface albedo in the UV band in CHIMERE-2017 is evaluated according to Laepple et al. (2005) in the absence of snow (tested as snow depth less than 1 cm), and from Tanskannen and Manninen (2007) in the presence of snow, tested as snow depth greater than 10 cm. Values are displayed in Table 4.

The snow depth is read from the WRF or ECMWF meteorological inputs, if available. If any other model is used, the snow cover will be assumed inexistent. If the snow cover
Table 3. Taking into account the various factors affecting the photolysis rates in CHIMERE-2013 and CHIMERE-2017.

<table>
<thead>
<tr>
<th>Factor</th>
<th>CHIMERE-2013</th>
<th>CHIMERE-2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZA</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Altitude</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Clouds</td>
<td>Parameterized</td>
<td>✓</td>
</tr>
<tr>
<td>Tropospheric ozone column</td>
<td>Constant profile</td>
<td>✓</td>
</tr>
<tr>
<td>Stratospheric ozone column</td>
<td>Constant profile</td>
<td>Month- and latitude-dependant climatology</td>
</tr>
<tr>
<td>Water-vapor concentration</td>
<td>Constant profile</td>
<td>✓</td>
</tr>
<tr>
<td>Aerosol effect</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Variable albedo</td>
<td>X</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 4. Tabulated values from Laepple et al. (2005) and Tanskannen and Manninen (2007) used for the calculation of the albedo in the UV band. In the presence of sea ice over ocean, the albedo of the ice surface is assumed equal to the Tanskannen and Manninen (2007) value for > 10 cm of snow on barren land.

<table>
<thead>
<tr>
<th>No.</th>
<th>Land use</th>
<th>Albedo for snow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&lt; 1 cm</td>
</tr>
<tr>
<td>1</td>
<td>Agricultural land/crops</td>
<td>0.035</td>
</tr>
<tr>
<td>2</td>
<td>Grassland land-use type</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>Barren land/bare ground</td>
<td>0.10</td>
</tr>
<tr>
<td>4</td>
<td>Inland water</td>
<td>0.07</td>
</tr>
<tr>
<td>5</td>
<td>Urban</td>
<td>0.035</td>
</tr>
<tr>
<td>6</td>
<td>Shrubs</td>
<td>0.05</td>
</tr>
<tr>
<td>7</td>
<td>Needleleaf forest</td>
<td>0.025</td>
</tr>
<tr>
<td>8</td>
<td>Broadleaf forest</td>
<td>0.025</td>
</tr>
<tr>
<td>9</td>
<td>Ocean</td>
<td>0.07</td>
</tr>
</tbody>
</table>

is thinner than 1 cm in the model, the albedo is assumed to be that of dry land. If the snow cover is thicker than 10 cm, the albedo is assumed to be that of snow-covered land. In-between, a linear interpolation is performed. Even though the case of sea ice is not explicitly treated in Tanskannen and Manninen (2007), the assumption is made in CHIMERE-2017 that the albedo of sea ice is the same as that of a thick layer of snow covering barren land.

5.3.3 Implementation

The physical calculations performed by Fast-JX are split in two steps.

First, the Legendre coefficients for the scattering phase function for all aerosol species and diameter bin are calculated using Michael Mischenko’s sph.err code (Mischenko et al., 2002), assuming sphericity of the aerosol particles. This calculation is performed for each of the $n_{\text{spec}} \times n_{\text{bins}}$ species, and for the five wavelengths that are used for the Mie scattering processes in Fast-JX. This step is performed once and for all before the first simulation step, and lasts from a couple of seconds to a couple of minutes according to the number of aerosol species and diameter bins. The refractive indices reproduced in Table 5 are the ones provided along with the model, essentially based on the values compiled in the framework of the ADIENT project\(^2\), as described by the corresponding technical report by E. J. Highwood\(^3\). However, the specification of these parameters is in a parameter file, and can be changed by the user to other values. In the same way, the user can easily introduce more species in the optical treatment for specific studies, e.g., volcanic ashes.

After the preprocessing phase, at each time step and in each model column, the Fast-JX module resolves the radiative transfer in the model atmospheric column, computing the actinic fluxes at each model level and integrating them over $N$ wavelength bins in order to produce accurate photolysis rates. In the configuration adopted for CHIMERE-2017, $N$ is set to 12, which is the value recommended by Fast-JX developers for tropospheric studies. These 12 wavelength bins include the seven standard Fast-J wavelength bins from 291 to 850 nm, as described in Wild et al. (2000). The seven standard Fast-J wavelength bins are essentially concentrated from 291 to 412.5 nm, which is the spectral band relevant for tropospheric photochemistry. Following the recommendations of Fast-JX model developers, these seven standard wavelength bins are complemented by five additional wavelength bins, from 202.5 to 291 nm, which are only relevant in the upper tropical troposphere. In a typical simulation framework, it has been found that the increase in computational time relative to the simulation with tabulated photolysis rates is below 10 % (Mailler et al., 2016).

5.4 Online calculation of lidar profiles

During the model integration, some additional diagnostic variables are estimated: (i) the clouds optical depth and the aerosol optical depth (AOD) using the Fast-JX module, and (ii) the lidar profiles.

The lidar profiles are calculated using the aerosol contributions only, as detailed in Stromatas et al. (2012). They are proposed as output after a simulation and are designed to be

\(^2\)http://www.reading.ac.uk/adient/refractiveindices.html, last access: 17 January 2017

\(^3\)www.reading.ac.uk/adient/REFINDS/Techreportjul09.doc, last access: 17 January 2017
directly comparable to ground-based or spatial lidars. Three different profiles are calculated both in nadir and zenith lidar configurations: (i) the attenuated scattering ratio, \( R'(z) \), (ii) \( \beta'(z, \lambda) \) and \( \beta'_m(z, \lambda) \), respectively, the total and molecular attenuated backscatter signal.

By definition, \( R'(z) \) is equal to 1 in absence of aerosols/clouds and when the signal is not attenuated. In the presence of aerosols, \( R'(z) \) would be greater than one. Following Winker et al. (2009), this ratio is expressed as

\[
R'(z) = \frac{\beta'(z)}{\beta'_m(z)}.
\]

The total attenuated backscatter signal \( \beta'(z, \lambda) \) is calculated as

\[
\beta'(z, \lambda) = \left[ \frac{\sigma_{m}^{\text{sca}}(z, \lambda)}{S_m(z, \lambda)} + \frac{\sigma_{p}^{\text{sca}}(z, \lambda)}{S_p(z, \lambda)} \right] \exp\left( -2 \int_{z}^{\text{TOA}} \sigma_{m}^{\text{ext}}(z', \lambda) dz' \right) + \eta' \int_{z}^{\text{TOA}} \sigma_{p}^{\text{ext}}(z', \lambda) dz'
\]

and the molecular attenuated backscatter signal \( \beta'_m(z, \lambda) \) as

\[
\beta'_m(z, \lambda) = \frac{\sigma_{m}^{\text{sca}}(z, \lambda)}{S_m(z, \lambda)} \cdot \exp\left( -2 \int_{z}^{\text{TOA}} \sigma_{m}^{\text{ext}}(z', \lambda) dz' \right),
\]

where \( \sigma_{p}^{\text{scat/ext}}(z, \lambda) \) and \( \sigma_{m}^{\text{scat/ext}}(z, \lambda) \) are the extinction/scattering coefficients for particles and molecules (in km\(^{-1}\)). \( S_m \) and \( S_p \) are the molecular and particular extinction-to-backscatter ratios (in sr). \( \eta'(z) \) represents the particles multiple scattering and \( z \) represents the distance between the emitter and the studied point. Note that for the case of a space lidar the integration begins from the top of the atmosphere (TOA) while for a ground lidar the integration begins from 0 (ground level) to \( z \). Further details about these calculations are provided in Stromatas et al. (2012).

### Table 5. Refractive indices for the main aerosol species in CHIMERE at 200, 300, 400, 600 and 1000 nm.

<table>
<thead>
<tr>
<th>Species</th>
<th>λ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>200</td>
</tr>
<tr>
<td>PPM</td>
<td>1.53</td>
</tr>
<tr>
<td>OCAR</td>
<td>1.60</td>
</tr>
<tr>
<td>BCAR</td>
<td>1.85</td>
</tr>
<tr>
<td>SALT</td>
<td>1.38</td>
</tr>
<tr>
<td>SOA</td>
<td>1.56</td>
</tr>
<tr>
<td>DUST</td>
<td>1.53</td>
</tr>
<tr>
<td>H2SO4</td>
<td>1.50</td>
</tr>
<tr>
<td>HNO3</td>
<td>1.53</td>
</tr>
<tr>
<td>NH3</td>
<td>1.53</td>
</tr>
<tr>
<td>WATER</td>
<td>1.35</td>
</tr>
</tbody>
</table>

### Table 6. Number of EMEP stations per species and per season used for performance statistics. Stations CH01, CH04, CH05, DE03, DE08, AT05, AT48, IT01, IT04, ES78 and DE44 were excluded from the analysis due to their topography difficult to simulate with a 0.5° resolution.

<table>
<thead>
<tr>
<th>Species</th>
<th>Winter</th>
<th>Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td>O3</td>
<td>96</td>
<td>93</td>
</tr>
<tr>
<td>NO2</td>
<td>40</td>
<td>34</td>
</tr>
<tr>
<td>SO2</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>PM10</td>
<td>26</td>
<td>20</td>
</tr>
<tr>
<td>PM25</td>
<td>22</td>
<td>16</td>
</tr>
</tbody>
</table>

### 6 Model scores for two test cases over Europe

The performance of CTMs is often evaluated by comparing simulation results to data of measurements, either from routine networks (Solazzo et al., 2012a, b) or from dedicated field campaigns (e.g., Menut et al., 2015; Petetin et al., 2015). Simon et al. (2012) presented an overview of performance evaluation studies for a large set of models and studied cases. A statistical evaluation with measurement data is performed for two 3-month-long simulations with CHIMERE-2017: summer (June–August 2008) and winter (January–March 2009). Each of the simulation periods analyzed were 2017: summer (June–August 2008) and winter (January–March 2009).
Figure 8. Performance statistics for the main model species and for daily averaged values. The numbers on the right axes give the overall scores (Pearson’s correlation, MFE, and MFB), while the box plots show the variability among the EMEP stations. The boxes extend from the lower- to upper-quartile values of the data. The center lines show the medians, and the red squares show the means over stations. The whiskers indicate the 5 and 95 percentile values, and the values on the right axis of each panel are the overall value of the considered indicator, i.e., merging all the stations into a single statistical dataset as described in Jacobson (2005).

The emission data were those from EMEP at 0.5°, and the boundary conditions for the concentrations from the LMDz-INCA model for gases and chemically active aerosols and from the GOCART model for dust. The simulation was performed with the MELCHIOR2 chemical mechanism for gaseous species, 10 bins for aerosol size distribution and the SOA (secondary organic aerosols) scheme of Bessagnet et al. (2008), 5 min chemistry time step and the Van Leer numerical scheme for both horizontal and vertical transport. The Wesely (1989) aerosol dry deposition and Loosmore (2003) resuspension schemes were used. The on-line coupling with ISORROPIA model was used.

The statistical scores are computed between modeled and observed daily averaged values, using surface concentration measurements from the EMEP monitoring sites, after filtering out the stations with complex topography (CH01, CH04, CH05, DE03, DE08, AT05, AT48, IT01, IT04, ES78 and DE44) that cannot be simulated appropriately at 0.5° resolution. Stations from the EMEP monitoring sites have been chosen for this study because their location has been selected in order to minimize local influences and be representative of large areas (Tørseth et al., 2012). For each simulation period only stations containing at least 70% of time series data were retained.

Figure 8 shows the performance statistics for the main model species. The number of EMEP stations used for each species for winter and summer is shown in Table 6.

The standard metrics used for air quality modeling (Simon et al., 2012) were employed, namely the Pearson’s correlation (PCOR), the mean fractional error (MFE) and the mean fractional bias (MFB).

Ozone shows the best scores among all the species, both for summer and winter, with PCOR = 0.70, MFE = 17%, MFB = 5% in summer and PCOR = 0.72, MFE = 25%, MFB = 2% in winter. It also shows the smallest variability of scores among the stations (93 available stations in summer and 96 in winter). As noted by Simon et al. (2012), the ozone overestimation often reported for CTMs is related to the averaging over the hours with high and low concentrations, so the scores are dominated by performance at low concentrations, which occur much more often than high concentrations. Indeed, the MFB computed from daily maximum ozone concentrations (not shown) is quite lower: 1% for summer and 7% for winter.

The NO2 shows quite larger MFE: 62% in summer and 53% in winter, with a large variability of both MFE and MFB between stations. The bias is negative in winter, slightly positive in summer but with a high negative values (NO2 underestimation) at some stations. For this particular species, with strong emissions horizontal gradients, the model resolution of 0.5° is not enough even when surface concentrations are measured at the background rural sites. Also, as discussed by Terrenoire et al. (2015), the negative bias could be partly related to the general underestimation of the emissions in the

The SO$_2$ shows the largest MFE for both summer (76%) and winter (81%) and the lowest correlation in summer (0.20). It shows positive bias: MFB = 32% in winter and 14% in summer. The difficulty in SO$_2$ simulation could be related to the uncertainties in the emission vertical profiles, which is a particularly sensitive factor in SO$_2$ modeling, because industrial stack emissions represent a substantial part of SO$_2$ emissions (Pirovano et al., 2012; Mailler et al., 2013). While some CTMs have included a plume-in-grid model for subgrid treatment of point emissions depending on the actual meteorological conditions and flux characteristics, this is not the case of the CHIMERE model, which can also limit the performance of the model regarding SO$_2$ concentrations. The conversion of SO$_2$ to sulfate can also be a source of error in SO$_2$ concentrations, as mentioned by Ciarelli et al. (2016) and Bessagnet et al. (2016), who observed very different behavior of models far from emission sources, probably due to the chemical mechanisms. The lower correlation coefficient in summertime was found in all the CTMs examined in Bessagnet et al. (2016).

The performance for PM is affected by compensating effects of several chemical components, such as dust, primary organics and secondary species like sulfates, nitrates and SOA.

The PM$_{10}$ concentrations are generally overestimated in winter (MFB = 12%), with correlation values lower in winter (0.50) and summer (0.23) than for the whole year, as reported by Terrenoire et al. (2015). In summer the PM$_{10}$ bias is quite low MFB = 0%, and the MFE (42%) shows small variability between the stations.

The PM$_{25}$ concentrations show a larger overestimation than PM$_{10}$ in winter (MFB = 35% vs. 12% for PM$_{10}$) and have also a positive bias in summer (MFB = 25%). The winter correlation is higher though (0.65 vs. 0.50), and its variability between the stations is smaller. The PM$_{25}$ overestimation can be associated to the overestimation of ammonium (MFB = 77% in summer and 65% in winter) and sulfate (MFB = 32% in summer and 33% in winter, not shown).

Boylan and Russell (2006) defined performance goals and criteria to be attained by air quality models. Their performance goal is attained for particulate matter when the MFE is less or equal to 50%, and |MFB| is less than 30%. The performance criteria are attained when the MFE is less or equal to 75%, and |MFB| is less than 60%. The performance goal is thus a more demanding condition than the performance criteria.

The PM$_{10}$ simulation satisfies the performance goal for both summer and winter. As for PM$_{25}$, it satisfies the performance goal in summer and the performance criteria in winter.

7 Application to the Puyehue–Cordon Caulle eruption (June 2011)

A simulation with the present version of CHIMERE has been performed for the Southern Hemisphere, from 15 May to 30 June 2011, a period covering the eruption of Puyehue–Cordon Caulle (Chile). This eruption emitted an important plume containing volcanic ashes and sulfur dioxide into the troposphere and the lower stratosphere. This plume had severe consequences on air traffic over Argentina as well as other countries in the Southern Hemisphere. While the eruption began on 4 June, the plume went around the entire Southern Hemisphere and was back in the vicinity of the emission source by 14 June (Global Volcanism Program, 2013; Klüser et al., 2013). This volcanic eruption case provides a perfect test bed to evaluate the new abilities of the CHIMERE model to simulate as accurately as possible transport at hemispheric scale, including cases where the transported plume undergoes a complete circumpolar trajectory around the South Pole.

7.1 Model configuration

The meteorological simulation has been performed using the WRF meteorological model, version 3.5.1, on a simulation domain covering most of the Southern Hemisphere at a resolution of about 55 × 55 km at 45° S, with 20 vertical levels from the surface to 100 hPa. For the gaseous chemistry, the MELCHIOR-2 chemical mechanism has been used. The horizontal domain is composed of 250 × 250 cells and is centered at the South Pole and covering the entire extratropical Southern Hemisphere. The horizontal resolution varies with latitude: 65 × 65 km (at the pole), 55 × 55 km (at 45° S) and 36 × 36 km (at 25° S).

The anthropogenic and biogenic emissions are taken into account and produced from the HTAP dataset and MEGAN model, respectively. Mineral dust emissions have not been included in this simulation, since the focus of this test bed study was in the circumpolar transport of ash emissions from the Puyehue volcano. The novelty of this simulation is the addition of the volcanic emissions of SO$_2$ and volcanic ashes.

7.2 Volcanic emissions

The total mass flux emitted in the form of particles has been represented according to Mastin et al. (2009), using the following equation:

\[ \dot{V} = \left( \frac{H}{2.00} \right)^{\frac{1}{3}} \]

\[ M = \rho \dot{V}, \]  

(30)

where \( H \) is the column height expressed in kilometers, \( \dot{V} \) is the volume flux expressed in m$^3$ s$^{-1}$, \( M \) is the mass flux in kg s$^{-1}$ and \( \rho = 2500$ kg m$^{-3}$ is the ash density. The altitude
of the ash column has been taken from Collini et al. (2013), and is reproduced here in Table 7. Only the fine fraction of the emissions, with particle diameter smaller than 63 µm has been included. The conversion from the total emitted mass flux has been performed using a conversion factor \( m_{63} \) taken from Mastin et al. (2009) for S2 type volcanoes, i.e., \( m_{63} = 0.4 \). It is worth noting at this point that the uncertainty on the value of this parameter, \( m_{63} \), is very strong, with values ranging from 0.02 to 0.6 depending on the characteristics of the considered eruption, and that therefore the uncertainties on the resulting mass of fine ash is very strong. The particles emitted with a diameter greater than 63 µm have not been considered because they are not supposed to be relevant for long-range transport due to their rapid sedimentation.

The emitted ashes have been distributed evenly from the altitude of the crater (2200 m a.s.l) to the altitude of the top of the column, obtained by summing the column height to the altitude of the crater.

The refractive indices of the volcanic ashes from Deriman et al. (2012) have been used. However, as these authors provided the refractive indices of volcanic ash only in the visible, the values at 200 and 300 nm have been taken as equal to the value given at 440 nm.

The granulometry of the ashes are taken as 80 % in a coarse mode, with a lognormal distribution centered at 30 µm and 20 % in a finer mode with a lognormal distribution centered at 4 µm, consistent with the results of Durant et al. (2009).

The SO\(_2\) mass flux has been taken from Theys et al. (2013), who prescribe mass flux estimates based on IASI measurements for the first 48 h of the eruption. Since these authors do not provide an estimation for the subsequent part of the eruption, we assumed that the SO\(_2\) fluxes are null after the first 48 h of the eruption. This hypothesis is of course questionable, but nevertheless the study of Theys et al. (2013) shows in a convincing way that most of the SO\(_2\) emission occurs during the first 48 h of the eruption.

### 7.3 Analysis of the circumpolar transport

The simulation is initialized by climatological concentrations for aerosols and trace gases from the LMDZ-INCA chemistry-transport model. These two datasets are also used to provide the top and lateral boundary conditions during the simulation. The simulation itself, covering the 15 May through 30 June, can be divided into two successive phases: first, from 14 May to 4 June, the model undergoes a spin-up period, with the concentrations of gaseous and particulate species building up due to the emissions of sea-salt and anthropogenic contaminants (Fig. 9a). At the end of this spin-up period, significant AOD values, from 0.05 to 0.20 appear over the Southern Ocean from 30 to 70° S, mostly due to sea-salt emissions, consistent with the findings of Jaeglé et al. (2011), and consistent with the satellite-based climatology of these authors, which represent a mean value about 0.15

---

### Table 7. Main characteristics of the volcanic emissions used for the hemispheric simulation. H: column height (km); \( \dot{V} \): volume flux (m\(^3\) s\(^{-1}\)); \( M \): Mass flux (kg s\(^{-1}\)); \( M_{63} \): emitted mass (kg); \( M_{63} \): emitted mass for the fraction with diameter < 63 µm.

<table>
<thead>
<tr>
<th>Day</th>
<th>H</th>
<th>( \dot{V} )</th>
<th>( M )</th>
<th>( M_{63} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/06</td>
<td>10</td>
<td>794.9</td>
<td>1.99 × 10(^{06})</td>
<td>2.86 × 10(^{10})</td>
</tr>
<tr>
<td>05/06</td>
<td>10</td>
<td>794.9</td>
<td>1.99 × 10(^{06})</td>
<td>1.72 × 10(^{11})</td>
</tr>
<tr>
<td>06/06</td>
<td>10</td>
<td>794.9</td>
<td>1.99 × 10(^{06})</td>
<td>1.72 × 10(^{11})</td>
</tr>
<tr>
<td>07/06</td>
<td>6.5</td>
<td>133.0</td>
<td>3.33 × 10(^{05})</td>
<td>2.87 × 10(^{10})</td>
</tr>
<tr>
<td>08/06</td>
<td>8</td>
<td>180.9</td>
<td>4.52 × 10(^{05})</td>
<td>3.91 × 10(^{10})</td>
</tr>
<tr>
<td>09/06</td>
<td>8.5</td>
<td>405.0</td>
<td>1.01 × 10(^{06})</td>
<td>8.75 × 10(^{10})</td>
</tr>
<tr>
<td>10/06</td>
<td>8</td>
<td>314.9</td>
<td>7.87 × 10(^{05})</td>
<td>6.80 × 10(^{10})</td>
</tr>
<tr>
<td>11/06</td>
<td>6.5</td>
<td>133.0</td>
<td>3.33 × 10(^{05})</td>
<td>2.87 × 10(^{10})</td>
</tr>
<tr>
<td>12/06</td>
<td>7</td>
<td>180.9</td>
<td>4.52 × 10(^{05})</td>
<td>3.91 × 10(^{10})</td>
</tr>
<tr>
<td>13/06</td>
<td>8</td>
<td>314.9</td>
<td>7.87 × 10(^{05})</td>
<td>6.80 × 10(^{10})</td>
</tr>
<tr>
<td>14/06</td>
<td>7</td>
<td>240.9</td>
<td>6.02 × 10(^{05})</td>
<td>5.20 × 10(^{10})</td>
</tr>
<tr>
<td>15/06</td>
<td>8</td>
<td>314.9</td>
<td>7.87 × 10(^{05})</td>
<td>6.80 × 10(^{10})</td>
</tr>
<tr>
<td>16/06</td>
<td>7</td>
<td>180.9</td>
<td>4.52 × 10(^{05})</td>
<td>3.91 × 10(^{10})</td>
</tr>
<tr>
<td>17/06</td>
<td>5.5</td>
<td>66.5</td>
<td>1.66 × 10(^{05})</td>
<td>1.44 × 10(^{10})</td>
</tr>
<tr>
<td>18/06</td>
<td>5</td>
<td>44.8</td>
<td>1.12 × 10(^{05})</td>
<td>9.68 × 10(^{09})</td>
</tr>
<tr>
<td>19/06</td>
<td>4</td>
<td>17.7</td>
<td>4.44 × 10(^{04})</td>
<td>3.83 × 10(^{09})</td>
</tr>
<tr>
<td>20/06</td>
<td>4</td>
<td>17.7</td>
<td>4.44 × 10(^{04})</td>
<td>3.83 × 10(^{09})</td>
</tr>
</tbody>
</table>

### Table 8. H: column height (km); \( \dot{M} \): Mass flux (kt d\(^{-1}\)); \( M \): emitted mass of SO\(_2\) (kg).

<table>
<thead>
<tr>
<th>Day</th>
<th>H</th>
<th>( \dot{M} )</th>
<th>( M )</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/06, 19:00–24:00 UTC</td>
<td>10</td>
<td>250</td>
<td>5.21 × 10(^{7})</td>
</tr>
<tr>
<td>05/06, 00:00–08:00 UTC</td>
<td>10</td>
<td>250</td>
<td>8.33 × 10(^{7})</td>
</tr>
<tr>
<td>05/06, 08:00–20:00 UTC</td>
<td>10</td>
<td>110</td>
<td>5.50 × 10(^{7})</td>
</tr>
<tr>
<td>06/06, 20:00–24:00 UTC</td>
<td>10</td>
<td>60</td>
<td>1.00 × 10(^{7})</td>
</tr>
<tr>
<td>06/06, 00:00–19:00 UTC</td>
<td>10</td>
<td>60</td>
<td>6.00 × 10(^{7})</td>
</tr>
</tbody>
</table>

in these areas. In the subsequent time steps, the volcanic ash plume from the Puyehue volcano becomes the dominant feature of the AOD structure in the Southern Hemisphere. While it is difficult to compare the simulated values to measured ones because of the large uncertainties on the mass flux and size distribution of the volcanic ashes, it is possible to compare the modeled trajectory of the ash plume with spaceborne observations. For this purpose, we will rely on the space images and analyses provided by Klüser et al. (2013) and Global Volcanism Program (2013). Figure 9b for 6 June at 12:00 UTC (08:00 a.m. local time) can be compared to Fig. 2 of Klüser et al. (2013), reproduced here for the reader’s convenience as Fig. 10, which shows that at this time, about 36 h after the onset of the eruption, the initial direction of the volcanic plume is eastward, with a slight southward tilt, consistent with the CHIMERE simulations. On 8 June (Fig. 9d), the simulated pattern for ash transport also fits very well the pattern that is visible on Fig. 11 (also taken from Klüser et al., 2013), with the initial portion of the ash plume traveling
Figure 9. Simulated AOD at 600 nm every 48 h from 4 June, 12:00 UTC to 14 June, 12:00 UTC.

southward over the southern Atlantic and reaching towards the southern Pacific ocean over Cape Horn, a pattern that is observed in both CHIMERE observations and the satellite observations. The older parts of the plume are located off the Atlantic coasts of Argentina, also covering a large part of southern Brazil in the model but not so in the infrared AOD data (Fig. 11). Finally, the plume from the initial explosions are located at that time in the southern ocean, in-between the southern tip of the African continent and the Antarctic. It can also be observed that while the ash plume is continuous in the CHIMERE simulation, it is not so in the observations. This reflects the succession of explosive phases and quiet phases of the volcanic eruption, while the flux imposed to the CHIMERE model is continuous, as discussed in Boichu et al. (2013), who also present a possible workaround for this problem by assimilation of satellite data.
On 12 June, 4 days later, the leading edge of the volcanic ash plume is located at about 135° W and 55° S above the southern Pacific Ocean, while other portions of the plume are located above New Zealand, Tasmania and areas of continental Australia and southern Africa (Figs. 9e and 11). Later on, on 14 June, the leading edge of the ash plume reaches back to the southern coasts of Chile, as visible in both the simulation outputs (Fig. 9f) and the report of Global Volcanism Program (2013), which indicates that part of the plume was reaching South America from the Pacific ocean at that time between 35 and 50° S while other parts of the ash plume were located further to the south, close to the Antarctic Peninsula, consistent with Fig. 9f. On 14 June and during the following days, the plume from the initial explosion of 4 June and the following days is passing over the Puyehue volcano again, a fact that is correctly captured by the CHIMERE model.

8 Conclusions

CHIMERE-2017 is a model version, which presents several major improvements compared to the earlier version described in Menut et al. (2013a). Compared to the previous model version, anthropogenic emissions can be generated anywhere in the world from the HTAP emission inventory, as well as mineral dust emissions, which were available only for North Africa and the Arabian Peninsula in previous model versions. With the same objective of permitting the use of the model in any part of the world and at any scale from...
urban to hemispheric scale, an important limitation of the model has been removed by improving the internal treatment of the transport on the sphere, allowing for domains up to the hemispheric scale, and possibly including a geographic pole. Much attention has also been paid to the physical processes, including a major update in the representation of the physical processes affecting the aerosols, as well as the effect of the modeled aerosol on the photolytic reaction rates. Other efforts have been made to improve the user’s experience with the model: this includes improvements in the parallelization of the model in order to reduce computation time, as well as providing key observable variables such as the aerosol optical depth and lidar backscatter coefficients, which permits the user to compare the outputs of the model directly with the results of remote sensing observations.

These improvements pave the way to many applications that were out of reach for the CHIMERE model up to now: CHIMERE 2017 has the necessary abilities to give new insights on questions, such as the radiative impact of aerosols on photochemistry, at all scales, from urban to hemispheric, including mineral dust emissions and deposition anywhere in the world. The possibility to run hemispheric simulations also allows for the use of this CTM for the study of transport of aerosol and gaseous contamination plumes between the different continent within a hemisphere. It contributes to bridge the gap between global chemistry-transport models such as LMDz-INCA, MOZART or Geos-CHEM and regional models. While CHIMERE has already been used successfully for the evaluation of the decadal trends in air quality over Europe (Colette et al., 2011), as shown by the study Xing et al. (2015) with the hemispheric version of CMAQ.

Figure 11. Figure by L. Klüser, T. Eberstedter and J. Meyer-Arnok, published in Klüser et al. (2013) as Fig. 3, with the following description: “The PCCE ash plume on its way around the Southern Hemisphere for descending MetOp orbits from 7, 8, 12 and 13 June.”
hemispheric versions of regional CTMs are tools that can be used successfully to study long-term trends in regional air quality with added value from models simulated in regional domains, only because they can perform a consistent simulation over the entire hemisphere without relying on boundary conditions provided by global CTMs relying on different assumptions and parameterizations.

**Code availability.** The present article refers to the CHIMERE-2017 release, which is freely available and provided under the GNU general public license\(^4\). The source code along with the corresponding technical documentation can be obtained from the CHIMERE web site at http://www.lmd.polytechnique.fr/chimere/.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** For anthropogenic emissions, EuroglobalMap products include Intellectual Property from European National Mapping and Cadastral Authorities and is licensed on behalf of these by EuroGeographics. Original product is freely available at www.eurogeographics.org. Terms of the license available at http://www.eurogeographics.org/form/topographic-data-eurogeographics. The MACC boundary conditions dataset was provided by the MACC-II project, which is funded through the European Union Framework 7 program. It is based on the MACC-II reanalysis for atmospheric composition; full access to and more information about this data can be obtained through the MACC-II web site http://www.copernicus-atmosphere.eu.

We acknowledge C. Prigent for providing the global high-resolution aeolian aerodynamic roughness length. The authors would also like to acknowledge Lars Klüser, Thilo Eberstede and Julian Meyer-Arnek for publishing the figures here reused as Figs. 10 and 11 with an Creative-Commons license, permitting reuse of these figures. We are also indebted to the two anonymous reviewers, who helped improve a lot this study from a scientific and editorial point of view.

Edited by: Alex B. Guenther
Reviewed by: two anonymous referees

**References**


Honeröl, C., Rouil, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., Elichegaray, C., Fland, J., Malherbe, L., Meleux, F., Menut, L., Martin, D., Peuch, A., Peuch, V., and Pois-


