Dear Editor and Reviewers,
All remarks and suggestions were taken into account and many corrections were made in the manuscript. Our answers are in blue and after each reviewer’s remark. When a large paragraph is added in the manuscript, it is in a grey box. This letter includes the answers to the two reviewers: since they have some common concerns, only one letter is prepared including all answers.

Best regards,
Laurent Menut
6 septembre 2021

1 Review #1

The document is well written and organized in a clear way.
A positive aspect of the approach in this model development, is the transparency of the model code, good documentation and therefore also the possibility to pursue reproducibility.
In order to improve/increase this reproducibility and to increase the interaction with the users, it would be a good idea to push the source code on a git repository.

Answer:
This is a good suggestion and we are thinking about it. It would be possible for a future version, since many steps have to be achieved before it becomes possible. The model being now distributed with a modified version of WRF for the coupling, we first have to upload our changes in the WRF GIT repository, in collaboration with the WRF developers. Once this is done, we can put our own code on a GIT repository and then distribute the model with two GIT repositories (one for each model).

Some major issues:

• When performing the validation for Europe, you are focusing only on the rural stations, arguing that those stations are indeed most representative for this type of domain. However, the CHIMERE model is at the same time running operational within the framework of Copernicus (CAMS) at a much higher resolution. Why didn’t you consider this resolution in order to also show the performance of the model for local sources (traffic stations, industrial, ..)?

Answer:
It is true that the CAMS resolution is much higher, with \( \Delta x = 10 \text{km} \). But at this resolution, it is not clear if traffic stations may be really simulated with a model. And, changing from 60 to 10km resolution, the cpu time used by the simulation will increase by a factor of 36, becoming non negligible for simulations of one year. This is the reason why we split the model validation in two steps: first, a long term simulation over Europe with a 60 km resolution. Second, a refined simulation using the subgrid scale variability, the only configuration that is truly capable of simulating small scale variability of pollutants in urbanized areas. In section 6 and with Figure 10, we have included a local validation using traffic stations.
Section 4:
The meteorological forcing is ECMWF, please give more details on the model version (ERA5? IFS cycle?) and which resolution is extracted.

Answer:
The ECMWF data used for the simulations is a global 0.25°×0.25° 3-hourly operational forecast of IFS. For each day, we used the analysis, that is the first available day of each daily forecast. This was added in the manuscript as:

In order to quantify the quality of this new version, two simulations are performed, one with v2017r5 and one with v2020r1. The meteorology used corresponds to the IFS/ECMWF fields, with a global 0.25°×0.25° resolution and 3-hourly time-frequency. For each day, we use analysis day of this dataset. The new version is used in offline mode to be homogeneous with the v2017r5 version.

Section 4.5 O3 profiles:
Here I would suggest to be very careful with the choice of the ozonesonde stations, since most of the stations are in the process of reprocessing their data. Please, mention the type of sensors used (ECC, BM) and if a reprocessing took already place. Only consider stations, where you have the full yearly cycle available in order to exclude a possible seasonal effect on your results.

Answer:
Details were added as requested and the new paragraph is:

We have also included validation for vertical profiles of pressure, temperature and ozone using the data downloaded from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC, https://woudc.org/home.php) resulting from ozonesonde measurements. From around 25 stations that are in our domain only 9 have consistent data for the year 2014. All the 9 stations use ECC as the measurement instrument, (https://woudc.org/data/instruments/). Out of the remaining 9 stations, 7 of them are well allocated over the seasons of the year, each representing at least 10 sondes per season. The 2 remaining stations, in Ireland and Italy, have 24 and 28 usable sondes respectively. These two stations present data only for the winter months. Our criteria for choosing the stations was their geographical position: we wanted to present different parts of Europe in the limits of availability of data for the stations, therefore the stations in Ireland, Switzerland and Poland were chosen in the presentation of the results. The number of total sondes over all the stations (561 in total) and their temporal distribution over different seasons is quite enough to be representative of both the geographical and temporal aspects of the simulations.

minor typos:
- line 170: OzoneSonde measurements should be ozonesonde measurements
- line 688: "BCclim"

Answer:
Thanks, it was corrected.
The paper described an important and valuable updates developed for CHIMERE model, which are particularly relevant for the CHIMERE modelers community and also to enlarge this community. However, there are large parts where this paper seems more a manual of the model, than a scientific paper.

Answer:
Yes, we agree with this point of view, and, for us, it is not a problem for this article. If we wanted to make a "scientific" paper (meaning a paper revolving around an environmental scientific question), we would publish it in ACP or JGR, among many others. Here, it is a paper in GMD, a journal dedicated to the model development. This is the reason why we describe the more technical aspects of the model novelties, giving details on the databases or the available flags, something that is not possible in the scope of another journals.

Besides that, there is too much updates included which do not allow to enter in detail in each one of them, in particular to perform an adequate model evaluation for the different developments. I would suggest that this paper could be divided into 2 parts (A and B), where the different type of model developments could be analysed in more detailed, and evaluated with test cases not so general (the domain is too coarse and the model validation is presented only in terms of average over the all domain! This is not sufficient to understand the benefits and advantages of the mode updates.

Answer:
First, it is not really necessary to split the paper in two parts, since GMD is an online journal only. Two PDFs of 20 pages or one of 40 pages gives the same outcome, while making it harder or the prospective user to find the information they need. Second, the goal of this paper is to provide an overview of all the changes made in this new version. Of course, it is not feasible to make all possible experiments, at all scales in one paper; however, that was never the goal of this paper : this paper must be easy to read and light. It must quickly show to the reader what the novelties are and provide a discussion about the "small" changes and the "important changes". The changes may be related to one database, or one dynamic process, or one chemical choice : many different things, not possible to directly inter-compare. Therefore, we are doing one simulation for each change and we have to construct very synthetic scores : it is the only way to quantify, with one simple figure, the impact of one change compared to another one (see Figure 13 for example).

In conclusion : this paper is dedicated to provide an overview of all changes of the new model version. Synthetic scores were defined to see what novelty induces the largest change and for what species. For more details, we are also working on "scientific papers", giving more informations about each novelty. It is described in the Introduction with the text, line 24 : 

1. The online coupling with WRF via OASIS3-MCT, to calculate the possible retroactions of aerosols on radiation (direct effect) and cloud formation (indirect effect), (Briant et al., 2017; Tuccella et al., 2019).

2. Natural emissions: the addition of dimethylsulfide, new schemes for sea-salt, improvements of the mineral dust and biomass burning, NO\textsubscript{x} by lightning, (Cholakian et al., 2018; Menut et al., 2018; Menut, 2018; Turquety et al., 2020).

3. Aerosols chemistry: the addition of the Volatility Basis Set (VBS) scheme (Cholakian et al., 2018), and the new Hydrophylic/Hydrophobic Organics (H\textsubscript{2}O) aerosol scheme, (Couvidat et al., 2012, 2018), coupled with the Secondary Organic Aerosol Processor (SOAP, (Couvidat and Sartelet, 2015)) thermodynamic model.

4. The subgrid scale variability of emissions: for cells with anthropogenic emissions from several activity sectors, it is possible to have pollutants concentrations for each. This option corresponds to Valari and Menut (2010).


6. The core of the model: with the possibility to use the splitting operators or the integrated production/loss terms approaches.

Taking into account these comments (and others following) I would recommend major changes before accept it to publications. Please consider the following additional major and minor changes:

1. Line 7: "The NOx emissions from lightning are added"
   
   \textit{Answer:}
   Yes, thank you, it was corrected.

2. Figure 1: Map of CHIMERE users in 2021
   
   \textit{Answer:}
   Corrected.

3. Line 27: "biomass burning" should be specified that is related to forest fires (there is also biomass burning in residential combustion sector, which was already included in previous version
   
   \textit{Answer:}
   Ok, the new sentence is:

   Natural emissions: the addition of dimethylsulfide, new schemes for sea-salt, improvements of the mineral dust and biomass burning due to forest fires, NO\textsubscript{x} by lightning...

4. Tables caption: should the legend of the table be placed below the table? (usually is on the top)
   
   \textit{Answer:}
   Why not. The placement of the caption depends on the journal. It is the choice of
the editor not ours. For this article, we are using the Latex package proposed by Copernicus. After the review, they can edit the article to fit their own rules as they see fit, since this is not the author’s choice.

5. Line 144 : please correct this sentence

Answer:
OK. The comment is not very explicit but we consider that the problem was about the general meaning of the sentence. The new sentence is now:

Several model configurations are defined in order to answer the following questions:

6. Line 150 : besides the quantification of the coupling changes, it would be important to show how the model perform better with this update

Answer:
It’s difficult to understand the ’besides’ because the two are evaluated. On one side, we evaluate the performances of the online coupling in section 5. On the other side, we made improvements in this new version and their impact is quantified in Section 9. And we show how the model performs with these updates (BCclim, nofire, wesely, despres and HTAP set-up) since the results are compared to a ’ref’ simulation.

7. Section 3.2 : a large coarse domain was selected but since there is only available air quality observations for Europe, why is this domain considered? Maybe the best would be to focus on Europe (and North Africa eventually to model dust) and perform a more detail and interesting validation exercise

Answer:
The domain was defined to completely have the African mineral dust sources explicitly modelled. Several previous studies have shown that long-range transport of mineral from Africa to Europe may affect aerosol surface concentration budget in a large extent. With this domain, we ensure to have an explicit calculation of the transport of these mineral dust. With a reduced domain, we perhaps could increase the resolution (from 60 to 40km perhaps), but we take the risk of having erroneous results, hence simulations would be more difficult to explain and validate against available measurements.

8. Line 178 : why do not used the Delta-Tool developed for model assessment pruposed within the FAIRMODE community?

Answer:
Yes, why not. For a next version, we could use the Delta-Tool code. We know that some research teams using CHIMERE are also using Delta-Tool. For this study, we made the choice of using the same tool as the INERIS forecast team in the context of the PREVAIR system : it ensures to have consistent scores between the model development for research and the forecast version used in PREVAIR and CAMS. All in all, the calculation of a bias between measurements and model output is the same with any tools.

9. Section 4.1 : More than know the average value of the statistical indicators over all the domain (!!) would be more valuable to have the idea of the dispersion of the error (box plots). For the BIAS, for example, considering only the average of the domain,
it would leave to misunderstanding of the results (if there is regions with positive and others with negative systematic errors values)

**Answer:**
For the sake of the clarity of the presentation of the results, it is necessary to sum up some results. Averaging over the domain is one way, providing time averages is another. It enables us to present results for several statistical scores and for several pollutants while still staying concise. Averaging the results does not add a risk of misunderstanding the results: the important point here is to have one synthetic value to be able to compare several very different processes. For example, in Figure 13, we compare the impact of boundary conditions to the dry deposition schemes. These processes are completely different. But here the goal is just to quantify globally the impact of each change, and to inter-compare their effect. Of course, it is probable to have different impacts spatially. For example, for the boundary conditions, the impact will be more important where the inside flux is more important. For the dry deposition scheme, the impact will be the most important when there is more pollutants above surfaces sensitive to deposition. For these model changes, we added more realistic databases and parametrizations schemes: we expect that the change will be at any place where they act.

10. Line 215: please explain. I did not understand: "The v2017r5 was also more accurate...For all situations, the v2020r1 is more able to represent..."?? ?? (please also correct v2002r1)

**Answer:**
Ok, probably the text was not clear. It was rewritten as:

An example is presented in Figure 4 with the count of exceedances of 180 µg.m⁻³ for the daily maximum for surface ozone concentrations. With the observations, two major peaks are recorded: the 10th June and 22th July 2014. Four to five other less numerous peaks were also recorded. For the peak on the 10th June, 23 stations recorded an exceedance in Western Europe. The model v2020r1 was able to reproduce 15 exceedances, whereas the v2017r5 was only able to reproduce 2. For the peaks on the 22th July, 15 exceedances are observed, 15 are modelled by v2020r1, 12 by v2017r5. For these two cases, the v2002r1 version improves the capability of the model to simulate ozone daily maximum surface concentrations.

11. Line 219-220: can the authors explain the low values given by v2017r5?

**Answer:**
The sentence is: "The new version gives lower values than v2017r5, but with a better day to day variability". It means that the v2020r1 version simulates lower values than v2017r5. The two versions simulate concentrations lower than the observations. As for all models, there are many possible reasons for this bias and this is impossible to explain with one or two processes. It is clearly a non-linear impact of all processes taken into account in the model.

12. Line 224: why only O3 is mentioned? there are other important pollutants

**Answer:**
For the specific case of the text line 224, it is the section dedicated to the vertical profiles. Only ozone concentration is available with the sondes; therefore only ozone
has been mentioned. But in the others sections, when other pollutants are available, the comparison is done. For example, Table 5 presents results for O$_3$, NO$_2$, PM$_{2.5}$ and PM$_{10}$.

13. Figure 7: this scheme could be improved in order to give more detailed...

*Answer:*
We understand the text could contain "more details". But here, it is a synthetic scheme just to have the processes involved. The details are provided in the text (Sections 5.1, 5.2) and in the reference papers of (Briant et al., 2017; Tuccella et al., 2019). In these papers, several more detailed schemes are presented.


*Answer:*
More detail is provided in (Tuccella et al., 2019). For this article, we added the following explanations:

The activation scheme of Thompson and Eidhammer (2014) is replaced by the one of Abdul-Razzak and Ghan (2002), this latter being able to use a sectional approach for aerosols as designed in CHIMERE. For each aerosol section, the scheme is able to calculate the fraction of activated aerosol.

15. Line 289-290: please review this sentence (english grammar)

*Answer:*
The sentence was replaced by:

The diagnostic of these grid-resolved clouds depends on the horizontal resolution. For a coarse resolution, they are less well calculated than for a fine resolution.

16. Line 291: remove "version"

*Answer:*
OK thank you, done.

17. Line 300: please identify the 5 simulations

*Answer:*
This paragraph was added:

These simulations correspond to:

(a) offline-wrf: the CHIMERE model is forced by hourly meteorological fields stored in the WRF model output files (as in the previous CHIMERE version). It corresponds to option \texttt{online=0}.

(b) cpl1: The is no coupling but CHIMERE received WRF meteorological fields using the OASIS-MCT coupler. The coupling frequency is now the physical time-step (for example 15mn).

(c) cpl2: Coupling with the aerosol direct effect

(d) cpl3: Coupling with the aerosols indirect effect

(e) cpl4: Coupling with the direct and indirect effects.
18. Line 311: please review this sentence (english grammar)

*Answer:* OK the sentence is now:

All configurations show an overestimation of the ozone daily mean concentration, with a bias between 3 and 4 $\mu$g.m$^{-3}$, compared to observations. The variability between the several model configurations is only 1 $\mu$g.m$^{-3}$, showing that the choice of the coupling is not the main reason of the bias.

19. Section 6: this development should have a paper dedicated to that. There is too much information/detail that should be given, which is not possible to analyse in a couple of lines (max 20).

*Answer:* Yes, we agree and this paper already exists, Valari and Menut (2010). The development proposed in CHIMERE v2020r1 is exactly the same as the one described in this article. This is why we present here only a short summary of the development and one new Figure as an example. For more details, the reader is invited to read this article (or the CHIMERE documentation).

20. Line 419: a brief explanation of what is being presented below should be given here

*Answer:* Yes, OK, the following lines were added line 419:

The properties of the primary semi-volatile compounds are shown in Table 8 and are: the name of the surrogate species, the molecular weight of this species, noted MW (g.mol$^{-1}$), the partitioning constant $K_p$ (m$^3$.µg$^{-1}$) and the enthalpy of vaporization $\Delta H_{vap}$ ([kJ.mol$^{-1}$]).

21. Section 7.3.1: there is no section 7.3.2: it make sense to exist only one?

*Answer:* No, of course, it does not make sense. It is a latex typing error, this subsubsection being a subsection. It was corrected.

22. Line 575-580: references to support this are missing

*Answer:* It is a very general paragraph about mineral dust modelling. We have thus added here some references to well known articles. The new paragraph is now:

The mineral dust emissions calculation remains highly uncertain in transport models, Cuevas et al. (2015); Middleton (2017). This uncertainty is linked to the fact that these emissions are estimated based on a threshold process computed with very uncertain forcings, Alfaro and Gomes (2001); Darmenova et al. (2009). The emission measurements are rare and, when they exist, representative mainly of a very local environment (a few tens of meters around the measurement site), Kok et al. (2012). In CHIMERE v2020r1, several improvements were done on the mineral dust emissions flux calculation: (i) implementation of several subgrid scale approaches to take into account soil and surfaces variabilities, and (ii) the mineralogy to explicitly model several species instead of only one.
23. Line 581-586 : references to support this are missing

*Answer:*
It is a paragraph about mineral dust variability. It is more on generalities about the modeling of a process when there is a threshold. We added these references as :

The problem of using a threshold is that the transition between a zero and a non-zero value is very sensitive to the accuracy of the forcing parameter and on its spatial and temporal representativity. For the mineral dust emissions calculation, the key parameters are the 10m wind speed and the surface and soil type, including the aeolian roughness length, Shao (2001); Engelstaedter et al. (2006). Their spatial representativity is the one of the model cell and may be different than the local phenomenon to model.

24. Line 605-608 : references to support this are missing

*Answer:*
It is a paragraph about mineral dust mineralogy. There aren’t many studies about this topic, but we added some references as :

Among many others, one interest of the mineral dust modeling is to calculate accurate aerosol deposition fluxes, Sokolik and Toon (1999). For biogeochemical studies, the deposition of mineral dust over the ocean is of interest, in particular for phytoplankton growth, Zhang et al. (2015). In this new model version, the detailed mineralogy of dust was implemented, following the development of Balkanski et al. (2007). Instead of one mineral dust species called ‘dust’, twelve mineral species are emitted, transported, and deposited. They are : calcite, chlorite, feldspar, goethite, gypsum, hematite, illite, kaolinite, mica, quartz, smectite, and vermiculite, Claquin et al. (1999); Balkanski et al. (2007). Each one has a specific density, silt/clay ratio, and refractive index. As with the mean mineral dust, these species are described with a binned size distribution (in general 10 bins as recommended in the parameter file).

25. Section 8.4 : see my comment above about saying only "biomass burning"

*Answer:*
OK, the section title was changed to : "Forest biomass burning emissions".

26. Figure 11 : I think Figure 11 should be placed only after being mentioned and not before

*Answer:*
OK, it was changed (but it is also Latex package dependent). It will probably be optimized by the journal editing team.

27. Section 9 : This section is too limited for so much work that should have be done and presented. That’s why my suggestion would be to split it into 2 papers, in order to be possible to perform a more completed model evaluation, in order to understand the viability and add-value of the different model developments

*Answer:*
It is true that the section is not very long. It was done this way to remain synthetic, the presentation of this model version being already in a long manuscript. The reviewer
points out the "viability" and the "added value" of the new model developments. We think that all these new developments correspond to a better accuracy of databases or parameterizations. Then, even if statistical scores are not always better (many possible compensation errors may exist), the new functions are going toward a more precise and realistic model. And we don’t think we have to give more details in this section since the goal we had in mind is reached: provide an overview of the changes and their variability, their impact on the modelled concentrations.

28. Line 733: please review this sentence (English grammar)

Answer:
OK the sentence was corrected and is now:

The physical time-step $dt_{phys}$ is defined by the user in the chimere parameter file. In offline mode, it indicates the value that the user wants for the transport calculation.

29. Section 10, 11: I have doubts if these sections should be included since there is no comparison and model evaluation. It is just information for the model users, which could be just mentioned, referred and reported in the (online) manual

Answer:
We think these sections are really relevant in this article. An article in GMD is not for comparison to data only. The way the time management is performed is an important model development. It is the same for the nesting/nudging in case of a regional model: the way they are performed impacts the results a lot.

30. Section 11: an introductory text should exist before Figure 15

Answer:
Yes, correct. Text was added and some sentences were changed as:

For the nesting, the way how it is managed with this new version is presented in Figure 15. As with the previous version, CHIMERE authorizes nesting up to three levels. However, the namelist files can be easily changed to increase the depth of the nesting.

In offline mode, CHIMERE can be used as usual with meteorological forcing through WRF, ECMWF, or from other sources of data. In this case, WRF can be run in 1-way or 2-way nesting, then CHIMERE is run after, as usual, for each domain with a 1-way nesting.

31. Conclusions: there is lack of explanation of how model validation was performed

Answer:
The way the model evaluation is performed is presented all along the article and we thought it was not necessary to repeat it here. But if the reviewer considers it as being important, we can add some sentences. The following was added to the Conclusion:

The validation of this version was performed using surface measurements of pollutants concentrations, aerosol optical depth and vertical ozone soundings. Statistical scores were calculated including bias, correlation, RMSE as well as scores on ozone daily maximum concentrations, being of importance for users doing forecast.
Références


Balkanski, Y., Schulz, M., Claquin, T., and Guibert, S. : Reevaluation of Mineral aerosol radiative forcings suggests a better agreement with satellite and AERONET data, Atmospheric Chemistry and Physics, 7, 81–95, doi: 10.5194/acp-7-81-2007, 2007.


Darmenova, K., Sokolik, I. N., Shao, Y., Marticorena, B., and Bergametti, G. : Development of a physically based dust emission module within the Weather Research and Forecasting


