

## ***Interactive comment on “A N-dimensional Fortran Interpolation Program (NterGeo.v2020a) for Geophysics Sciences – Application to a back-trajectory program (BACKPLUMES.v2020r1) using CHIMERE or WRF outputs” by Bertrand Bessagnet et al.***

### **Anonymous Referee #2**

Received and published: 20 July 2020

Review for <https://www.geosci-model-dev-discuss.net/gmd-2020-88/gmd-2020-88.pdf>

The manuscript "A N-dimensional Fortran Interpolation Program (NterGeo.v2020a) for Geophysics Sciences – Application to a back-trajectory program (BACKPLUMES.v2020r1) using CHIMERE or WRF outputs" by Bertrand Bessagnet et al., describes an interpolation subroutine written in Fortran and its application to analysis of CHIMERE data. The topic lies within the scope of GMD, but the manuscript in

C1

the current state needs additional work in order to be considered for publication. The writing is not sufficiently clear, which makes the reading difficult. The current level of English is inadequate for publication in GMD. It needs to be revised and edited by a proficient speaker. It would be ideal if a native English speaker could language edit the manuscript.

Specific comments are listed below:

Title "A N-dimensional Fortran Interpolation Program (NterGeo.v2020a) for Geophysics Sciences - Application to a back-trajectory program (BACKPLUMES.v2020r1) using CHIMERE or WRF outputs" triggers a question. CHIMERE being an Eulerian off-line chemistry-transport model, it doesn't generate wind fields. Offline trajectory models require wind fields in order to run. In what sense BACKPLUMES uses CHIMERE output? CHIMERE outputs are mainly composition fields. Does it mean that BACKPLUMES uses CHIMERE interpolated meteorological fields from WRF?

P1

Abstract

I4 "Fortran is a powerful and fast language, highly portable and easy to interface with other existing Fortran models." This phrase is problematic: Fortran is not a model. Most languages should allow interfacing code. Is the point that there are many existing geophysical models written in Fortran? It could be replaced by something like: "Fortran is a fast and powerful language highly portable. It is easy to interface models written in Fortran with each other."

I5 "any Fortran compiler": did the authors try all possible options? Why don't just say that is written in standard Fortran and tested with two compilers? In fact there are KIND statements and those can be dialect dependant.

L5

I6 "novel optional parameter" This is unclear: novel with respect to what? Optional with

C2

respect to what?

I6 "For the general program, the inverse distance is used for the weight calculation with a distance defined as a p-distance". This phrase is undefined at this point. Which weights? What is p? The abstract should be self contained.

I8 "Moreover, a real case of geophysics application embedding this interpolation program is provided and discussed" Is it a "geophysical application"?

I8 "it consists in determining back-trajectories using atmospheric dispersion or mesoscale meteorological model outputs, respectively from the widely used models CHIMERE and WRF." Here "atmospheric dispersion" is unclear when applied to CHIMERE, which as far as I know is an Eulerian chemistry model itself. "Dispersion" usually refers to Lagrangian models. The phrase reads as : "NterGeo.v2020a determines back-trajectories using atmospheric dispersion (?) from CHIMERE and meteorology from WRF." Is that the intended meaning?

I13 "used in its recent version V4.9.2 to propose horizontal and vertical interpolations" "used to propose interpolations" please revise the English expression.

I17 "full coupling strategy": Full coupling of what?

I18 Is ISORROPIA the chemistry and thermodynamics module? then please place the reference at the end of the phrase. Or it is an interpolation routine?

I19 "In this case" do you mean in such cases? Is this related to what is said above? If so please do not open a new paragraph which implies a new topic.

P2

I25 "In some studies," is too informal, please revise.

I31 (Scipy, 2014), unfortunately -> (Scipy, 2014),]. Unfortunately...

I32 "not enough optimized for our objective": Please make explicit which objective.

C3

I33 "irregular (varying intervals) but structured" is unclear. What is the structure? Do you mean a square array?

I35 It is still unclear what the program precisely does and why the existing interpolation algorithms are not sufficient.

I36 "Atmospheric models (physics and/or chemistry) are commonly used in the Geophysics community." This phrase as it is contains little information. Mainly Lagrangian models are named later in the paragraph, maybe the authors mean "Lagrangian models (physics and/or chemistry) are commonly used in the Atmospheric Science community?"

I39 It is "atmospheric motion" not "motions".

I40 "A new back-trajectory plume has been developed" do you mean "A new back-trajectory plume model has been developed"?

I41 spatial interpolation of what? Please specify.

I41 "the other codes": please specify which codes.

41 Unspecified: "some additional characteristics are implemented," please list the most important.

Message of section 1: The introduction suggests that the aim is to interpolate 3d aerosol data on trajectories calculated with BACKPLUMES.

Section 2)

I51 "The program supports": which program, NterGeo?

here it is still not completely clear what "irregular structured" mean. This may suggest an arbitrary triangular mesh. But it appears that what is meant is that the nodes are in multidimensional grid array with varying edge lengths, i.e., there is an  $N_1 \times \dots \times N_n$  array of cells that are hexahedrons with arbitrary trapezoids as faces.

C4

L57 "As it includes not specific options or function, version of a compiler, there is no reason to have limitations or errors with other compilers." Which compilers did the authors try? Better simply to say that "for portability, the code is written in standard Fortran without elements from any particular dialect."

P3

L58 "The top shell calling script in the package provide two sets of options for ": provide\*s\* two sets of options...

I58 "Assuming the X array, the result of the function f transforming X to Y array in R can be expressed as: " is confusing, please rewrite. It is stated that Y is a scalar, not an array, otherwise express the suitable space  $R^m$ .

I62: Typography, please use "`\verb|`" or the verbatim environment and not italics for file names.

I63 "a classical bilinear interpolation" before you mentioned arbitrary dimensions, do you mean multilinear?

I 70 Style: please don't start a phrase with a symbol.

I 79 1-dimensionAL array

eq 5 please check and define all variables involved. Is the summation from  $k=1$  or from  $j=1$ ?

I 85 "closest neighbours" or vertices of containing grid cell? It may be that the closest vertices correspond to a neighbouring cell

I 94 an hard coded option -> A hard coded option

I 99 in a 1D array. -> into a 1D array.

I 125 "If the distance is too high compared to the characteristic" -> "If the distance is larger than the characteristic"...

C5

I 141 "The time consuming is somehow proportional to the number of points in which to interpolate." This phrase is not grammatically correct and as stated does not add any important information. It is trivial that performing more operations takes longer time, the authors could replace "somehow" for something more precise.

I 143 "There is a discontinuity of the NMSE from  $p = 1$  to  $p = 1+$ ": why?

I 144 "The NMSE decreases with the number of points but a slight increase is observed 145 from  $200 \times 200$  from  $300 \times 300$ ." why?

153 This could certainly depends -> This could certainly depend

I 175 "can use output files": what kind of data? CHIMERE and WRF output contains a very large number of variables, which are necessary?

"This kind of model has mainly one calculation to do several times" please revise this phrase. It is difficult to make sense out of it. Do you mean: This kind of models (Lagrangian) contain a time loop that performs many times the same kind on calculation?

I178 "The first advantage of BACKPLUMES is to use the results of a simulation already performed. " do you mean that BACKPLUMES is an offline model?

I 179 "homogeneous with the "direct" simulation" do you mean consistent with the Eulerian forward simulation?

I 180 "BACKPLUMES is different than other back-trajectories models, such as Hysplit or Flexpart." In what is different?

Please correct the English: different than other -> different to other

I 181 "Since it is difficult to calculate correct probabilities back in time, the choice was made to randomly launch numerous trajectories and try to cover all possible origins." This phrase is very unclear, please clarify.

I 181 Flexpart's goal is not to "estimate the most likely trajectory as an envelope of

C6

numerous possible trajectories."

enveloppe -> envelope

l 185: "The BACKPLUMES model is an open-source code and is available on the CHIMERE model web site." You should add a link with this statement. I looked in the website <https://www.lmd.polytechnique.fr/chimere/> and I couldn't find any reference to BACKPLUMES.

186-189 This paragraph is useful but incomplete. Since it is not described before (although it has been used in a couple of publications) you should specify better the numerics of the model.

192 " The number of trajectories is a up to the user and may be from one to several hundred of tracers.": "tracers" usually denote chemical species. Do you mean particles?

193: This paragraph starts with a discussion of the time step and the CFL condition and then jumps to the WRF specific calculation of the z coordinate. It is difficult to follow the logic and has many English language errors. Please rewrite paying attention to the logical order in the definition of the variables for the equations.

202 "the horizontal grid cell" do you mean dependent on the horizontal coordinate?

Eq 15: does 1 correspond to the lowest level?

207 " is eta values on full (w) levels " please revise the English.

"The layer thicknesses varying in space but also in time, this calculation is done for all trajectories and all time-step. " please revise the English.

200-208: This section is very unclear. You have to better explain the rationale of the log pressure vertical interpolation and possibly move the technical details (maybe expanded) to an appendix.

## C7

Figure 5: "Methodology for the vertical distribution as a function of the diurnal cycle of the boundary layer height." Please define the circles, squares and the black thick line in the caption.

210-215: How do you know that this Euler scheme is accurate enough?

219-220 There is also mixing in a neutrally stable boundary layer, even in the absence of thermal convection. How do you treat such cases?

225 "Therefore a random function is applied..." Please define Rand.

228 "It is considered that 15 mn is representative of a well-mixed convective layer." Why? Can you add e.g. a reference?

233 "its possible vertical motion with values between 0 and  $w/2$  m s<sup>-1</sup>." Where do these values come from?

237 Define Rand (as all symbols) the first time is mentioned.

The treatment of mixing is very simplistic. The well mixed condition is not even mentioned. Although simplified approaches may find application in certain situations, its limitations and shortcomings should be better presented and discussed. There is abundant literature on Lagrangian methods in the atmosphere to refer to.

239: "The difference between the two models is 240 the number of vertical levels." How many in each case?

241: "The wind field is also the same, WRF sending this information to CHIMERE." Please rewrite this phrase revising the English.

242: " WRF using the (Hong et al., 2006) schemes and CHIMERE using the (Troen and Mahrt, 1986) scheme." Are there many Hong et al. schemes? Please use `\citet` and not `\citep` in order to place the parenthesis in the right place.

245 Please use a proper LaTeX degree symbol.

## C8

259: " But, in both cases, the answer in our case is clearly that the main contribution of the air masses located at the starting point are mainly coming from the North." There is very little quantitative information in this statement.

267: "This interpolation program can be used for any application in Geophysics and Engineering Sciences but also to explore large structured matrices for Machine Learning applications." This conclusion is not supported by the main body of the manuscript. Especially the mention to Machine Learning. There are no results nor discussion of this topic. Please rework it in the discussion or remove from the conclusions.

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-88>, 2020.