



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-441-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-441

Anonymous Referee #1

Referee comment on "Improving trajectory calculations by FLEXPART 10.4+ using single image superresolution" by Rüdiger Brecht et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-441-RC1>, 2022

Review of "Improving trajectory calculations by FLEXPART 10.4+ using deep learning inspired single image superresolution by Brecht et al., submitted to EGU sphere.

The study by Brecht et al. presents an application of a neural network to provide higher-resolution wind information from a coarser resolution. Specifically, the authors analyse the impact of their approach on the accuracy of trajectory calculations, which rely heavily on interpolation operations. The paper describes interesting and relevant new approaches to interpolation in an atmospheric framework. However, there are a number of issues with the methodology, the presentation of the results, the discussion of limitations, and the overall writing style, that need to be addressed before the manuscript can be considered publishable.

Major items

1. Currently, the authors sub-sample a higher resolution model field to obtain the coarse resolution wind field. This approach is in my impression inconsistent with what a coarser-resolution model would provide. A coarser-resolution model would provide an average of what is represented by several grid points of a finer-scale model. I think the sub-sampling makes it harder for the linear interpolation to provide good results. That sub-sampling approach is also the origin of the checkerboard patterns apparent in e.g. Fig. 4. I recommend to redo the analysis by averaging the model fields rather than sub-sampling.

2. Is it correct that the neural network works on a 3x3 grid point stencil on a lat-lon grid? It seems that the footprint of the interpolation operator would thus see very different area sizes near the pole than near the equator. How is this affecting the training and application of the neural network?

3. Some sections are poorly written and therefore hardly comprehensible (for example section 3.2). Since the authors introduce new methods into the field of atmospheric science, I recommend to make an extra effort to clearly define all terminology (e.g., "channel attention" is never defined). However, I do not think this is mainly about the content material, but rather about how the sentences and paragraphs are compiled. For example, section 3.2 could start with a short paragraph, providing a break-down of the steps involved in the architecture, before describing each part in following paragraphs. I recommend the authors to take a look at Gopen and Swan (1990) regarding how to write more clearly and effectively.

4. The issue of non-conservation of interpolation algorithms is a major concern in a physical application as presented here. The basic model equations are derived from principles of mass and energy conservation. Therefore, if there are conservation violations induced by this method, this aspect needs to be clearly brought forward throughout the manuscript, to make sure it is not overlooked by readers. This aspect can be mentioned in Sec. 3.4, brought up as part of the results, for example by comparing the kinetic energy and the velocity spectra from the fine-resolution and interpolated fields. The short discussion in L. 234 onwards might be more suitable in a discussion paragraph.

5. The results need to be structured more clearly. Right now, the sequence of results and examples in Sec. 4.1 appear somewhat arbitrary. While there certainly is some reasoning behind the structure and examples, it is not spelled out clearly, and thus the reader is left guessing about how to "connect the dots". Coming from an atmospheric science perspective, I suggest a structure that starts one specific case as an example, such as one of the frontal bands shown in Fig. 3., where the linear interpolation has clear deficiencies. Thereby, it would be helpful to also show additional atmospheric variables to illustrate the case (for example surface pressure or air temperature. A tropical cyclone or a Rossby wave breaking could be other interesting situations to present. After stepping through the example case, more statistically robust information could be provided, from considering a larger number of days or cases. Finally, you proceed to the application with the trajectory calculations, before considering energy conservation.

6. On many occasions, the results are presented in a qualitative way (closer/larger/etc.). In order to connect the results to the figures, and to make it possible to follow the interpretation and evaluation of the authors, it would be very useful to include concrete numbers alongside the qualitative interpretation, while referring to the respective figure panels. Examples are L. 182 and onward, L. 198 and onward.

7. The writing style of in particular the results section should be more distanced or objective. Now, the authors frequently use expressions like "we demonstrate", "we show" in the start of a paragraph, i.e. before actually having presented the evidence. As a critical reader, one might get the impression you are overselling the results. I strongly recommend changing this unnecessary forceful writing style to a more distanced, objective style. Let the reader see the evidence for themselves, while guiding them through the material, before drawing conclusions. Many paragraphs in the results are currently "upside down" in that way.

8. As another, related aspect, the figures are not properly described. At present, the length of the text describing the results is very much out of balance with the number of figures. For example in Sec. 4.1.1, L. 183, an entire 3 figures are referred to within just 3 sentences, but none of the sentences describes what actually is seen within the figures. Rather than leaving it up to the author to interpret the figures, use some sentences for each figure to describe what is displayed, and highlight what is important to take away. This applies to all figures in the manuscript.

9. On several occasions (including Fig. 2, 4, 5), the figure captions contain information about the method or results that are not mentioned in the text. Such information must be placed in the main text.

10. What are the limitations of the method in terms of computational effort? In L. 192, it is briefly mentioned that the computation time is a factor 10 larger than linear interpolation. Is there still an advantage of neural network approach compared to for example quadratic interpolation? This could be worth a short section in the discussion. The improved conservation of other properties is also interesting, but unfortunately not shown in more detail.

Detailed comments:

Title: "deep learning inspired": unclear what this expression means, consider to remove/replace. State what aspect of trajectory calculations is improved (accuracy).

L. 20: Can you back up this statement by a reference/example?

L. 21: "where a dense network" rephrase. If the point is that the numerical weather prediction process produces large amounts of gridded data, then it would be sufficient to state just that, without mentioning observations (which are not at all part of this manuscript). Remove "reanalysis model", a reanalysis is generated from regular NWP models.

L. 25 onward: check citation of references, missing brackets.

L. 27: remove "just to name a few"

L. 30: logical gap, what is the connection to the previous paragraph?

L. 34: remove "surprisingly", this entirely depends on the perspective of the writer.

L. 34: briefly define "convolutional neural network".

L. 44: what do you mean by "variable-scale"?

L. 44: what to you mean by "deep" - how deep?

L. 45: Rephrase: "showcase" sounds like snapshots or illustrations, but as a reader I look for reliable evidence.

Section 2: "Related work"

This section does currently not serve a clear purpose, and is somewhat duplicate with the introduction. I recommend to delete this section here, and partly incorporate bits in the introduction, partly into a clearer method description.

Section 3: "Methods"

This section would benefit from a first paragraph that explains your overall approach, followed by a section that discusses the choice of the neural network, based on the range of choices that exist, in an accessible writing style.

Section 3.1: "Training data"

The training data would be more natural to place after sections that describe the actual neural network and approach.

L. 82: Why could this seem little data? How much training is commonly needed?

L. 107: rephrase using more distanced and objective terms. It could provide depth to the study to present a less well-performing approach in an appendix.

L. 114: remove "for testing purposes"

L. 117: 50 or 88 -> 50 and 88

Figure 1: several abbreviations and terms of the operations in the figure are not defined, include in caption or describing text. What do the bracketing lines indicate? The hierarchy between (a), (b) and (c) and between (a) and (d) could be made clearer in the figure, e.g. by lines that indicate "zooming in".

L. 123: Add a statement about the purpose of the error metric, i.e. what is to be assessed.

L. 127: here and elsewhere: ground truth -> truth. (ground truth would only make sense in a remote sensing context)

The notations for RMSE and SSIM could be simplified and clarified, for example using \hat{z} for the interpolated quantity, and using a, b instead of x, y (which is commonly used for spatial coordinates) for the two figures in the SSIM metric. How important is the "perceived similarity of two images" for the given application? This would be a suitable place to mention conservation issues due to interpolation.

L. 142, 144: unclear what "this" refers to.

L. 145-149: unclear, please rephrase

L. 152 onward: the emphasised names do not appear to be re-used in the remainder of the manuscript. Maybe rather introduce 3-letter abbreviations, such as REF, LIN, NNI that then can re-appear in the results and figures.

L. 160: place references at the end of sentence

L. 163: clarify whether X_n, Y_n are vectors with m elements, or for a specific time along the trajectory

L. 172: This section seems to describe your approach, and would be better placed in the methods.

Figure 2: the lines for lin u and lin v are exactly equal, is this coincidence? x-axis is lacking a unit. RMSE is defined with an index, but given without index here. Please explain in the result text how to interpret this figure.

Table 1: What do the arrows indicate? The caption contains a key result, that should be moved to the main text.

L. 183 to 188: Need to guide the reader through the results. The comparison needs more structure, and quantitative examples from the figure where available to support the qualitative conclusions.

L. 191: Maybe express in relative terms, hardware-dependent?

L. 195: distinguish "evaluate" and "apply" - an objective way to present the results would be to apply the method, display the results, and thereafter evaluate based on the error metrics.

Figure 3: Lacking panel labels. The top row does not seem to give additional information to the bottom row. I recommend using a continuous color scale; the two-color scale gives unjustified importance to errors larger than 5 m/s. Maps are missing coordinates. The RMSE in the title should be part of the text rather than a caption title. It would be useful to present a specific situation with meteorological fields for context.

L. 198: "before": rephrase

L. 199: "relative error reduction": where shown?

L. 200: "This holds...": can this information be presented as part of a more aggregated and thus robust result?

L. 202: unclear, rephrase

Figure 4: See comments about Fig. 3, the color scale gives unjustified emphasis to wind errors above 2.5 m/s. Indicate in Fig. 3 where this zoom is taken. Arrows are difficult to see, take to separate panels, and use meteorological fields (e.g. sea level pressure or potential temperature) as reference in both sets of panels.

Figure 5: This figure needs more explanation. What bins have been used? With only 10 bins, it may be more appropriate to show the lines as step function. What error metric has been used? Can this figure be constructed on more than just one day to make it more robust?

Figure 6: see Fig. 3.

Table 2: see Table 1

Figure 7: consider to remove this Figure. At this point, quantitative information may be sufficient/more useful than another illustration

L. 209: likely -> conceivable, provide reference

L. 215: it would be useful to briefly re-cap how these results are obtained. One case, several cases, specific region? How are trajectory errors distributed on a global map, do they mirror the interpolation errors?

L. 221: "smaller" - should this be "larger"?

L. 223: "directly corresponds" - is this a result, your interpretation, or an assumption?

L. 228: These paragraphs would better fit into a discussion section, together with other limitations. If possible, it would be useful to give more details, such that other studies can refer to your work.

L. 242: remove "just to name a few"

L. 245: would be useful to connect to weather phenomena here

L. 250: remove "see Fig. 2"

L. 263: this is an important limitation and should be taken up at different locations in the manuscript, including a discussion section. If non-conservation is an issue here, it would be useful to quantify. This would also give some balance to the study, which now mainly focuses on the advantages.