

# Authors' reply to reviewer comments GC-2019-22

Weather and Climate Science in the Digital Era, Martine G. de Vos et al.

## Short Comment 1

*P2, L 45. There is a line that talks about the “third development”. The construction of this paragraph could be slightly modified to explicitly present the three developments, for a better flow.*

We adopt the suggestion of the reviewer and will modify the construction of the paragraph

*P3, L 63. Section 2, Consider eliminating too many “and” conjunctions.*

We adopt the suggestion of the reviewer and will check the text for unnecessary “and” conjunctions

*P4, L 94-96. Examples or relevant references cited will improve the effectiveness of this statement.*

In fact, open sharing of data, software and vocabularies is only true common practice in a few fields such as astronomy and genomics. Most scientific fields, including weather and climate science, can be considered lagging behind. We will add a few references to support this.

*P4, L 106 onwards. Some parts in 3.1 Open Data seem to fall under 3.2 Open software. But, this could also mean they are very coupled. No changes necessarily needed here.*

*P5, L 118. While interpreting , “Making data and software findable..”, software may include tools that lead to the data. I think some level of paraphrasing may be required in this paragraph to make the message from the paper more evident, about making all the components adhere to FAIR goal as a whole.*

The reviewer is right, data and software in are connected and both should adhere to the FAIR principles. We will modify the text of this paragraph (and if necessary other parts of the paper) to clarify this message.

*P5, L 126. This paragraph does provide good insights. But, the final message is not translated well enough as to how this affects open data/science.*

*P5, L 131. Just a note- Removing the need for post-processing by incorporating as many steps as possible within the model itself can make the model computationally even more expensive. Thus, when there is a use-case to share model source code, one may still find it challenging, though open. Though there is one helpful cloud computation reference cited, I would have expected to see more bits about cloud computing in this paper, in this particular section.*

We agree with the reviewer that the impact on open data/science can be stated more clearly. We included a more elaborate description that producing FAIR model data is necessary, but can not be achieved through traditional post-processing pipelines.

Furthermore, we agree with the reviewer that cloud computing technologies, like xarray, Dask, and Apache SPARK, could be useful, since data processing and analysis pipelines usually do not require communication between parallel jobs. One of the key aspects, however, is the capability of the developer, usually a meteorologist or climate scientist, to adopt a new programming paradigm that allows the parallel execution of the workflow on cloud infrastructure. Here research software engineers may play a key role by, e.g., building useful tooling on top of existing low-level platforms like Apache Spark or Dask.

We will rephrase the paragraph accordingly.

*P6, L 161. Punctuation. Add comma after conference.*

We will rephrase the sentence

*P6, L 178 The message/action item here seems to have not translated well here. It does sound contradictory, but the essence of the message might be lost, regarding the technical challenges and reduced scope for multi-discipline collaboration. Please paraphrase this to improve the paragraph.*

We will rephrase the paragraph to clarify the message:

“The use of software as presented above, motivated by open science principles, requires a suitable digital infrastructure. The cloud appears to be a potential avenue, as it enables individual researchers to gain access to high computing resources, vast amounts of storage and a suite of software tools. In our session, several digital platforms were presented, that use cloud technologies to create a virtual research environment where scientific end-users can store, analyze and share their data. The participants also observed, however, that current platforms, like the Open Geospatial Consortium and JRC Earth Observation Data and Processing Platform, do not seem to increase the extent of scientific collaboration, especially across disciplines. This may be partly due to the fact that these platforms each have implemented their own set of standards for both data formats and interfaces to access these data. Since scientists are required to invest time and effort in working with a specific platform, the heterogeneity poses hurdles to their collaboration with researchers on another platform.”

*P7, L 194 Punctuation. Replace “here” with “there.*

*We will rephrase the sentence*

*P7, L 216 This statement is well put in terms of sharing knowledge. I hope this can be reflected more in the paper.*

We thank the reviewer for this comment. Throughout the paper we will rephrase text to be more specific on our observations and how these support our story. At the end of the paper, we will compile a list of action points or conclusions, i.e., to improve the current situation, that are described in the different sections of the paper.