

Authors' reply to reviewer comments GC-2019-22

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

Reviewer Comment 2

Major Points

1. *There are a number of typographical mistakes, albeit mainly subtle. So please get a native english speaker to proof-read the manuscript. Namely, I have not attempted to pick up all typos.*
We adopt the advice of the reviewer and get a native speaker to edit the manuscript

2. *The methodology (i.e. what was done in the session) needs to be clarified e.g. (i) were specific questions/topics posed for this research exercise [which it was], (ii) elicitation by sticky notes or hands in the air or by the co-authors making notes of what the group said? I think the observational data are (i) L57-558 - a specific session to discuss (by unstated means) the issues (unspecified in detail), and (ii) L20 insights from the work in the rest of the conference (by unstated means). A 'Methods' section needs to be added, which is one place where the questions asked at the session could be stated.*

3. *The 'novelty' (i.e. what is reported here that is not stated elsewhere) is difficult to distinguish, although a Methods section and taking care to phrase the results/discussion in terms of the evidential basis of insights should fix this.*

4. *The Abstract portrays all the thoughts as entirely new, rather than emerging from a context. e.g. L8 'we observed' - we reaffirm? we agree with the informal subject-wide consensus? Please rephrase where appropriate. As an editor of GC, I note that this was submitted as a review article, but it may be better classified as a standard paper.*

The approach followed in the session was similar to a 'focus group' approach where experts in share views and experiences. This paper is not a classical science paper addressing a well posed problem, but synthesizes those experiences from arguably a wide range of specialists. We agree with the reviewer that both the context and type of this research, and the methodology deserve clarification. We adopt the advice of adding a dedicated 'Methods' section (see below for the suggested text). Besides, throughout the paper we will rephrase text to correctly reflect our methodology.

“ Methods : The focus of the conference session was on data and compute intensive approaches that are applied in weather and climate science. The session comprised 10 oral abstract presentations, one keynote talk, and 6 short poster pitches. The 16 participants were either presenters or involved in the organization of the session, and represented domain science, as well as computer and data sciences.

The first part of the session was dedicated to the presentations. The second part was interactive. In three groups of each 5 or 6 persons the participants discussed the "challenges and opportunities regarding open weather and climate science" and noted their findings on a flipchart. The findings of each group were presented and discussed in a following plenary session. Observations and insights from the plenary discussion were documented. The observations in this paper are based on both the insights from the studies presented in the session, and the notes made during the interactive part of the session. The majority of the participants from the session also contributed to this paper. As such this paper represents a shared view of the participants, i.e., a group of experts in weather and climate science, on the digital and open science developments in their field."

Minor Points

Title - The paper's contents are about open access, not digital (see L2&3). Suggest changing title to reflect this.

We agree with the reviewer that this paper is about open science. In fact, we think we really do include both open science and the digital era. We suggest that we include both terms in the title, i.e., Open Weather and Climate Science in the Digital Era. In the introduction we will point out what we mean by "digital era".

L6 - 'the studies in the conference session showed' - How exactly?

We will rephrase the paragraph to clarify its meaning and add concrete examples that illustrate the importance of shared data and software:

"The majority of studies (roughly 80 %) presented in the conference session depended in some way or another on shared data and software. For example, many studies included open datasets from disparate sources to improve accuracy of forecasts on the local scale, or to extend analyses beyond the domain of weather and climate. Furthermore, shared software is a prerequisite for the studies that presented systems like a model coupling framework or a digital collaboration platform. Although these studies showed that sharing code and data is important, the consensus among the participants was that this is not sufficient to achieve open weather and climate science and that there are important issues to address."

L8 - 'we observed' - how (in)formally was this done?

L62 - A brief comment on the limitations/benefits of the approach used to bring together the information for this paper appears necessary in the Methods section.

L99 & 103 - Session/sessions? One 'session' with multiple time blocks?

L103 - A hint of what was done. Good, but please expand in a Methods section. Using the standard Method/Results/Discussion format might help the clarity of the work. Having everything merged into thematic section currently makes determining what this paper adds difficult,

although by clearly stating which evidence comes from where and moving from data to discussion within the existing sections might also work.

L104 - 'Discussed'. Please elaborate. e.g. who is 'we'. The co-authors of this paper? How was it determined what are 'common findings' and 'highlights'?

L118 - Example of where evidential basis could be clarified. 'we recognized': we as co-authors discussing and concluding, we in the session, and how was this recognized (e.g. large majority in room, or someone mentioned, or did all participants agree to a circulated notes/minutes?).

We agree with the reviewer that both the context and type of this research, and the methodology deserve clarification. We adopt the advice of adding a dedicated 'Methods' section (see our reply in 'major points' for the suggested text). Besides, throughout the paper we will rephrase text to correctly reflect our methodology.

L9 - Typo - 'there' not here

We will rephrase the sentence

L11 - 'primarily due to'? i.e. either these were refined from a list for some reason, or is this the complete list of possibilities?

This statement refers to the section where these barriers are described in more detail. For instance, many data sources come from private industry who may see a competitive advantage to maintaining privacy. But those data may prove useful to the weather community for improving initial conditions of forecast models. Such conundrums may be solved by signing nondisclosure agreements and allow weather service to act as trusted agents who use the data for the public good without disclosing their details.

We will include this explanation in the abstract and in the corresponding section.

L19-20 - It is claimed that 'much faster progress' is being made as 'observed from the studies presented in the conference'. This is quite a leap of logic, and is one illustration of how the manuscript could be more tightly argued and/or presented. If this is simply the authors impression, this is fine, but should be clarified by adding 'we believe' or similar. If written as a statement, and evidential basis should be provided in the new data collected. If this is simply a confirmation of what is in the existing literature (i.e. L52-53) then this should be also clarified.

The reviewer rightly points out that this is the authors' view. We will rephrase the sentence accordingly.

L21 - Typo - .. computationally intensive ...

We will rephrase the sentence

L22 - Introduction. A wide range of topics and issues are introduced here. They are placed in historical context, which is good. But, the treatment of these becomes quite vague when the actual session is mentioned (L58- 59)

The introduction discusses the role and use of technology in weather and climate science in history as well as the 'digital era'. We will clarify this as mentioned in the reply on the first comment.

We will move the description of the session to the new Methods section.

L39-48 - This paragraph is currently un-referenced. Please add these.

We adopt the reviewers suggestion and will add the following references to the paragraph:

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47–55. <https://doi.org/10.1038/nature14956>

Huntingford, C., Jeffers, E. S., Bonsall, M. B., Christensen, H. M., Lees, T., & Yang, H. (2019). Machine learning and artificial intelligence to aid climate change research and preparedness. *Environmental Research Letters*, 14(12), 124007. <https://doi.org/10.1088/1748-9326/ab4e55>

Schneider, T., Lan, S., Stuart, A., & Teixeira, J. (2017). Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted High-Resolution Simulations. *Geophysical Research Letters*, 44(24), 12,396–12,417. <https://doi.org/10.1002/2017GL076101>

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>

Ruti, P., Tarasova, O., Keller, J., Carmichael, G., Hov, Ø., Jones, S., ... Yamaji, M. (2019). Advancing Research for Seamless Earth System Prediction. *Bulletin of the American Meteorological Society*, (August 2019), 23–35. <https://doi.org/10.1175/bams-d-17-0302.1>

L40 - 'exascale' - I don't know this word. Please add a reference or two so that non-specialists can inform themselves.

The term 'exascale' computing refers to 10^{18} operations per second, a factor of 1000 beyond current machines.

We will explained the term in the text and add a reference to the sentence:

Reed, D. A., & Dongarra, J. (2015). Exascale computing and big data. *Communications of the ACM*, 58(7), 56–68. <https://doi.org/10.1145/2699414>

L62 - Open science. This appears to be a literature review, unrelated to the session mentioned. Was the session simply used as a brainstorming exercise to get the information together for such literature reviews? If so, again this is fine, but include a Methods section to state this, even if it's only a paragraph long. When the paper is revised, I would expect to distinguish whether the information is (i) in the literature, and being brought together here (ii) views of people in the room etc And, this will allow the contribution of this paper to be clarified/determined. If this

section is a review, say 'review' not 'explore', but my Methods points still stand w.r.t later sections.

The reviewer rightly points out that this section contains a literature review on open science. We will clarify this both in this section, i.e., rephrase the 'explore' sentence, and at the end of the introduction section, where we explain the structure of the paper. We will also add a dedicated 'Methods' section (see our reply in 'major points' for the suggested text).

L106 - Please try to be specific. Does 'many' mean 5, 50% or something different? It should be possible to give numbers for papers in your session, or you might randomly sample the conference in a desk-based exercise.

We agree with the reviewer and throughout the paper we will rephrase text to be more specific on our method of data collection.

General - Is there scope for a table of key points, or graphic to present the most important findings? I am a bit ambivalent about saying this as us readers shouldn't be lazy, but this could usefully highlight the key detailed points. Example of how this could be done - each co-author gets 3 votes, and size of coloured blob relates to number of votes in the graphic.

We thank the reviewer for this great suggestion. At the end of the paper, we will provide a list of action points or conclusions that are described in the different sections of the paper.