Key: Blue Text=Reviewer. Black Text=Author Response. Purple Text=Alterations to the manuscript.

Reviewer #2: Brandon Reichl, GFDL

This paper proposes the framework for ocean and sea-ice focused data output standards for CMIP7 models. The text provides a summary of this project's origination, structure, and recommendations. The objective of this manuscript is to document this committee's task to create a high-level overview for the recommended marine based CMIP7 data requests and summarize the process. This objective is achieved. Since the paper lays out an important set of recommendations for the CMIP7 community, it is appropriate for publishing. My comments can be viewed as suggestions that the authors may consider for revision rather than concerns that need to be addressed to warrant publication.

We thank Dr. Reichl for his time and effort in helping with this process.

General Comments

The authors clearly devoted considerable time and effort into generating this data request, and it is a monumental feat to achieve this summary. This paper could easily have been 2-3x its current length depending on the granularity of presentation of the opportunities and the discussion of output variables associated with each. Distilling it down to something more manageable is a positive outcome, but this does sacrifice the ability of this paper to be self-contained. As such, there is much jargon and detail that assumes a reader is already intimately familiar with CMIP and the CMIP7 data requests. The reader is also assumed to know where to find the details that were omitted. That is probably fine for the intended audience and for the purpose of this work to fit within a collection of related manuscripts. It is a very different paper from the OMIP protocol and diagnostics paper (Griffies et al., 2016), which seems fine given a more high-level purpose.

We appreciate the thoughtfulness of this remark and appreciate your suggestions for improvements where possible.

It took me a considerable amount of time to figure out how to find the data variables for this request. I found the Airtable website (https://airtable.com/appOcSa4gXyzHThmm/shrkayKObes58Zu45/tbljoSaMlK7m0DunX/viw0evRBr0vqp658c) after some exploring through the provided Zenodo link, github, and the binder software implementation, and eventually realized this was all organized on a website (https://wcrp-cmip.org/cmip7-data-request-v1-2-2/). Is the website address intentionally omitted from this text? The Airtable seemed to me the most direct way to scan through the details of the data request without installing special software and spending some time learning to navigate the data structures. Maybe it was omitted because it is not a permanent resource? Could some of the information be exported to PDF supplementary materials?

We intend to offer such permanent links alongside the paper. We will check with the CMIP-IPO and find out the planned implementation, and make it clear in the final version.

I do not notice any obvious omissions of ocean and sea-ice model output that would prevent CMIP7 models from serving their primary purposes. There are a number of non-baseline variables included whose necessity could be debated at length, and probably already was debated within this large author group. The discussion about why these variables are needed is laid out well scientifically. However, I do find the data justification of this request lacking in detail, focusing on these science questions that can always be better addressed with more data, but not reporting much detail on the consideration of specific data demands. Of course, judging the appropriateness of certain data volume requests will vary significantly depending on one's specific interests and the resources available to a specific modeling institution. If possible, providing some quantitative estimates related to the added data cost of various opportunities, variable groups, etc. would be helpful. This would especially help convey that this real (and significant) implication of the cost to institutions of serving the data request was weighed against the science it will support.

This is an important point. There are estimates of the data volume that can be extracted from the Airtable. We will find these estimates and include them in the paper. There are some companion papers in progress, such as the OMIP protocol paper, which will add more detail relevant to that narrower audience.

Prioritization of outputs (including variables or frequency) could also have been discussed more, if a goal is to ensure more institutions provide certain specific model output. There are some "high", "medium", and "low" priority cases, but what gives a variable its priority is not discussed in detail. Is it expected that modeling centers under numerous pressures from deadlines will be able to save and publish the "low" priority output? My takeaway from Section 4.1 is that a proper prioritization process was very difficult and likely demanded more resources than were available.

The prioritization was indeed an overwhelming task. The basic approach was to begin with the prioritization from the CMIP6 data request (Griffies et al., 2016; Notz et al., 2016) and then modify based on assuming a lower median level to begin, then raising and lowering the priorities based on our discussions and the download and upload statistics from CMIP6 archives. Prioritization is very subjective, and one may argue that many variables featured in a low-priority category may actually be high priority for a specific scientific purpose. We provide one perspective on variable priority, acknowledging that this is not the only point of view. Modelling centers will then have to decide what they can produce, depending on feasibility.

A discussion on spatial coarsening of model output could be useful somewhere. Some variables (e.g., heat content, certain scalars, MOC) can likely be coarsened with the savings in data serving demands outweighing the loss in information (e.g., from a ¼ degree grid to a 1 degree grid). But other data should presumably not be coarsened beyond the model native resolution if possible (e.g., extremes; winds if the purpose is for driving a wave model).

This is a good point, and in the CMIP6 archives, model variables were categorized into those that should be on the native grid (i.e., those primarily intended for budgets to close) and those that can be safely interpolated onto a 1 degree grid. However, there was no easy way within the

data request protocol to decimate data in the horizontal. Decimation in time and in the vertical were more easily done and were utilized by our team.

I wonder if “Waves” belong as their own topic alongside ocean and sea-ice? E.g., the title could be “ocean, sea-ice, and wave priorities and opportunities”, given the attention to waves in this data request. This emphasis on waves would then better justify why two of the more data intensive opportunities are associated with waves parameters.

Given that waves are inherently part of the ocean realm, we believe the current title is sufficient to describe the extent of this data request. It is true that 2 out of the 7 opportunities relate to waves. This was particularly interesting to our team and not of our own design. A few of our authors were involved in wave modeling efforts for CMIP6 (including Fox-Kemper), but those opportunities were contributed from outside of our team. The fact that there are two opportunities reflects the two ways that waves are simulated in CMIP, offline via the COWCLIP team and its descendants and online via the efforts at FIO and NCAR during CMIP6 and additional teams joining in CMIP7. Because these two applications differ substantially in which data is requested (forcing vs. forcing and response, essentially), and the fact that only some modeling centers have online wave capabilities, while all modeling centers can choose to provide forcing data for offline wave calculations, we found it easier to organize them separately.

Specific Comments

Introduction

L77: Check this sentence, maybe replace “has” with “and”.

Fixed.

L82: Multiple time [and spatial] scales. (is frequency needed here?)

Fixed.

L86: The terms in italics could be defined somewhere (perhaps in a table). E.g., after reading this paper a few times I’m still not entirely clear on what is meant by “opportunities”.

These are terms of art that extend across the entire data request. However, we have confirmed that definitions are provided on first use.

L107: “The accompanying tables…” This reference seems to only include a small spreadsheet of the REF variables. Is it meant to be to the full opportunities/groups/variables discussed in this text?

As mentioned above, we will add links that point to a permanent archives of the full data lists.

Approach and Methodology

This essentially reads like a technical report, there is not much to comment on regarding the methodology.

Unavoidably so, unfortunately. We experimented with different language styles but none were major improvements.

Table 1: I wonder if there is a way to add more granularity to the breakdown of the variables. E.g., ID 47 has 240 variables, but if they are 2d monthly variables it would be a completely different request than if they are 240 3d daily mean variables. Grouping by time period and/or 2d vs 3d could help clarify the data demand. If data volumes were part of the decision making, it could also help to explicitly include some quantitative data estimates (e.g., many data volume estimates are given by Juckes et al., 2024). I eventually figured out that this information is contained in the AirTable (if that is reliable?), but without significant experience using AirTable that wasn't obvious to me and so could be summarized here.

The airtables are the correct repository, and sorting as you suggest is possible there. We will point to the permanent archives of these tables from the paper and note the capability for grouping and sorting there. We have produced a csv formatted version of the archive, which should be robust enough to avoid issues from future software changes in python, etc.

Table 1: "Experiment Groups" might be useful to define.

Fixed.

3.1, Ocean Changes, Drivers and Impacts

ENSO is briefly mentioned, but the science questions referenced here largely neglect tropical topics. Maybe that reflects the state of the scientific interests now, but perhaps some ENSO implications could be mentioned?

Added.

No glaring variable omissions. The total estimate of data volume on AirTable is about 23 TB, which seems manageable. The ocean_mesoscale addition is a fairly substantial fraction of this, I think that their cost could be acknowledged.

Noted the high-res implications and estimated the total.

L183: What is the practical benefit of this clustering? What is the benefit of a variable group?

Now defined more carefully on first use.

Section 3.2 (Sea Ice Changes, Drivers, and Impacts)

No comments, the data estimate from AirTable was 16.1 TB

Thank you.

Section 3.3 (Paleoclimate)

The data estimate from AirTable was 22.4 TB

Thank you. There is some overlap with the preceding opportunities' lists.

L297: Some elaboration could help here, the assumption now is that a reader is familiar with paleoclimate experiments and knows what the abrupt-127k simulation means (I had to look it up)

A more extensive definition is now provided. We link to relevant publications that define abrupt-127k and that illustrate the function of this simulation in the context of the wider CMIP7 framework.

Section 3.4 (Polar Amplification):

The data estimate from AirTable was 35.8 TB

Thank you. There is some overlap with the preceding opportunities' lists.

Section 3.5 (Extremes):

The data estimate from AirTable was 54 TB, this opportunity will likely be an important one from CMIP7 and will be well utilized by groups performing model analysis.

Thank you. We agree that this opportunity will provide important new insights.

L386: I'm surprised to see BGC data also included in this request, it seems somewhat out of the stated scope.

This is only a limited set, selected here only for the purpose of multi-component extremes. These variables are treated in a larger BGC context by the biogeochemistry data request team.

Section 3.6 (Wind waves)

The total size of this data request on AirTable is 64 TB, or almost 3x that of the Ocean Changes opportunity. That significant potential overhead is probably worth discussing, or at least explaining how much of the data is an additional cost vs atmospheric variables that would already have been stored as part of other opportunities.

Thank you. We re-evaluated the experiments list that this opportunity points to and considered options to reduce the data footprint.

L405: Those that do are often at rather coarse grid spacing (this is sort of mentioned in other places).

Noted.

L407: "high-resolution data" -> This is subjective, it could be elaborated what time frequency and spatial resolution are preferred.

Specified.

L409: I'm unclear what is meant by "independent of ESM outputs". My understanding was that the ESM outputs are to be used to drive the wave model?

Clarified.

L411: Should it be obvious why this offers computationally efficiency and spatial detail? I'm not sure that it is.

Clarified.

Section 3.7 (wave coupling)

This is the most expensive opportunity in this topic area according to AirTable at 97 TB, or roughly 4x the size estimate of Ocean Changes. If the AirTable estimates are reliable and these are not just associated with otherwise collected variables I think some discussion is warranted. This opportunity being much more data intensive than the extreme impacts opportunity does make it feel like a substantial new request.

Thank you. We re-evaluated the list of experiments and variables for this opportunity. The new set of wave opportunity variables has been reduced to two groups (3-hourly and monthly). These changes should reduce the overall storage estimate compared with the previous version. However, it is important to note that not all model experiments are selected for this opportunity, we anticipate only a limited number of modeling centers will have active wave components. We also will propose (optional) time slices to provide these variables, which will prioritize a few time windows most meaningful for comparison of wave coupling impacts.

Appendix B: I'm unsure what the need is for choosing this subset of variables to define more formally (aside from being new to CMIP7?). It is a rather long/unwieldy list, but is incomplete and thus requires intimate knowledge of CMIP6 variables. A lot of non ocean and cryosphere variables are also included, is that intentional? I generally like the inclusion of some key variable information that can be accessed without surveying the AirTable (could be this appendix, or supplementary materials).

Yes, this list is provided to add some clarity (to the previously initiated) about the new variables. Although this list is long, we felt that it was helpful as it differs from preceding data requests. The non-ocean variables are relevant to the sea ice and paleoclimate opportunity, which is in scope (note that this is the ocean AND sea ice data request).

Typos/Word choice:

L163: As already noted [in the introduction and references within]

Fixed.

208: resolution -> grid-spacing

Fixed.

208: higher -> finer

Fixed.

Citation: https://doi.org/10.5194/egusphere-2025-3083-RC2