Baseline Climate Variables for Earth System Modelling - revisions

Martin Juckes[1], Karl E. Taylor[5], Fabrizio Antonio[8], David Brayshaw[2], Carlo Buontempo[3], Jian Cao[4], Paul J. Durack[5], Michio Kawamiya[6], Hyungjun Kim[7], Tomas Lovato[8], Chloe Mackallah[9], Matthew Mizielinski[10], Alessandra Nuzzo[8], Martina Stockhause[11], Daniele Visioni[12], Jeremy Walton[10], Briony Turner[13], Eleanor O'Rourke[13], Beth Dingley[13].

[1]Univesrity of Oxford, and UKRI STFC, UK

[2]University of Reading, UK

[3]ECMWF, Bonn, Germany

[4]Earth System Modeling Center, Nanjing University of Information Science and Technology, Nanjing, China

[5]PCMDI, Lawrence Livermore National Laboratory, Livermore, USA

[6]JAMSTEC, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

[7]Korea Advanced Institute of Science and Technology, Korea

[8]CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy

[9]CSIRO, Climate Science Centre, Australia

[10]Met Office Hadley Centre, UK

[11]DKRZ, Germany

[12]Department of Earth and Atmospheric Sciences, Cornell University, USA

[13]CMIP International Project Office, UK

Correspondence to: Martin Juckes (martin.juckes@physics.ox.ac.uk)

Table of Contents

General Response

We are grateful for the broad and constructive response. The interest generated is an indication of the utility of the list as a focus for discussion.

The comments and questions provide valuable feedback and will help us to clarify the role and context of the list.

Several of the comments refer to the details of vertical levels of atmospheric and oceanic data. We can provide more detail, but we also wish to allow flexibility. The variable list has now been adopted as a starting point for the for the AR7 Fast Track request (https://wcrp-cmip.org/cmip7/cmip7-data-request/public-consultation/). Decisions about the extension of the request to include more variables have been delegated to author teams. More details about this process will be published separately. Version 1.0 of the request contains 1813 variables compared to 132 in the BCV list and 2062 in the CMIP6 data request. There is also ongoing discussion about changing the precise levels used. Those author teams have more detailed knowledge of the scientific requirements in the relevant domains so it appears appropriate to let them make the final (for the AR7 cycle) decision and report on their decision process.

There are a number of suggestions for additions and changes. Where backed by clear new evidence a change has been accepted, but changes which are only backed by a brief comment are not considered as grounds for modifying the outcomes of the process that we have been through. As Reviewer 1 commented, the process itself is important. We will clarify that when MIPs request data they are expected to add more variables. Additional variables

which are requested by many MIPs are likely to be strong candidates for inclusion in the next version of the list.

While it is clear that a well designed governance structure and an associated timeline for revision would be highly desirable, we are not able to provide this at present. The community is focussed on the immediate challenges of developing a full request for IPCC AR7 Fast Track simulation. The question of next steps cannot be fully addressed until this heavy-workload stage of this process is behind us. The publication of this list will be an initial step in an iterative process.

Reviews and Editorial Feedback

RC1: Claire Macintosh

Reviewers Comments

General comments

This paper represents a substantial and important step forward for the CMIP community looking towards CMIP7. The presented BCV list will form the core of the CMIP7 data request, with the underlying groundwork and philosophy having wide ranging implications across the ESM community.

There is some tension in the paper between the concept of a BCV list as it applies to the WCRP modelling multiverse generally, and the specific implementation of this list as the core of the CMIP7 Data Request and its associated tight timescale. I have tried to make clear in this review which aspect is being addressed by each comment.

In addition to the carefully considered results presented here, the author team should be acknowledged for their approach to the transparency of process in the development of the BCVs, which is an excellent example of good practice in the field.

Specific comments

Please note that I have been asked to give this review in part to provide perspective from the observational community. Some comments reflect that request.

Table 1. Stakeholders of the CMIP DR. Row 1. Examples of “communities studying the global climate” is currently restricted to MIP communities. Other direct users of CMIP data also exist outside of the MIP framework, not least a large number of scientific researchers using CMIP to elucidate specific processes or aspects of the climate system outside a specific MIP.

Section 3 Line 288, Line 303 – see Section 5 comment.

Section 3.4 Role from the data user’s perspective

The BCV list as a whole is aimed primarily at modelling centres. However, the manuscript would benefit from more careful consideration of the wider CMIP user and associated observational communities.

The example of the need of some users for high temporal resolution data presented here is important, but by no means the only consideration from the perspective of the wider CMIP user community.

For context: a search of Scopus lists 4189 papers containing “CMIP6” in the title or abstract. Of these, 1371 (33%) also contain a least one observational keyword (observations OR satellite OR in-situ OR reanalysis)[1], increasing to 1720 (41%) if the word “evaluation” is also included. This inexhaustive list of keywords represents therefore a lower bound on the

fraction of the CMIP6 community that is using at least one auxiliary dataset alongside CMIP6 data.

The implications for the BCV list are clear. Given that more than a third of the CMIP community is using some kind of observational data, a key role of the BCV list must be not only that it is common across CMIP modelling centres, but also that it provides enough information to downstream users for observational comparisons and evaluation to be possible. This includes e.g. information on pressure levels, variable names that are consistent across the ECV-BCV boundary, variable choices that are suitable for observational evaluation, considerations of relevant observing resolutions, and clear information on methodological choices to generate BCVs. In short, it must ensure that it is externally facing such that it is sufficient for these analyses.

Some discussion of implications and additional requirements for the BCV list for external communities would be beneficial –

- In the general case: What are the implications for exploitation of the BCVs with and without coordination/interoperability with equivalent observational parameter lists (ECVs, GCIs etc.).
- Do these differ for direct vs indirect users (the latter being more likely to be using derived metrics, where the original form and nuance of both the BCV DR and observational data may be obscured).
- In this phase of CMIP: What input is needed from observational or other auxiliary data communities to maximise the interoperability aims of the BCV DR (for example, development of variables that are more directly comparable with model output – e.g. trivially skin vs layer temperature - or techniques and documentation where comparisons are nuanced e.g. vertical integration to a small number of layers vs observing resolution, pitfalls for regional analysis).
- What actions can the BCV DR take aimed at maximising the uptake of BCV DR across this interface and therefore achieving the overarching aims of the exercise, both for this phase and beyond.
- What gaps exist at the interface that should be filled?
- How might future iterations of the list more systematically address the widespread use of auxiliary datasets in analysis of CMIP or ESM MIP data? What is needed in the longer term?

Section 5 Conclusions: The BCV list has wide implications for ESM MIPs generally, but will also in the near future form the core of the CMIP7 data request. Given that there will be immediate and substantial CMIP community interest in the practical implementation of the list, and that numerous downstream communities will begin to make decisions on their respective implementations in preparation for CMIP7, Section 5 would benefit from some discussion on immediate and future next steps, and an aggregation and expansion of relevant issues identified elsewhere in the paper.

Please note that it is not for the authors to necessarily answer in detail to all aspects of the implementation phase, but rather to highlight in this paper issues that must be addressed by next steps, any potential pitfalls, and further community engagement that is needed in order to maximally exploit the careful and detailed work presented here. For example-

- Governance: how will the list be managed and updated? What issues must be addressed?
    - For this phase of CMIP: How will any updates or amendments be transparently curated, deployed and communicated to the community.

    - For this phase of CMIP: How will this list interact with the wider CMIP7 data request. For example the passing on of specific variable requests to the wider DR communities, where they are assessed as not part of the BCVDR. What action is needed from within and without the BCV community.
    - For future phases: Line 303. How might new or emerging variables be fairly and transparently assessed for inclusion (e.g. new land surface or biosphere variables, that may be disproportionately important in the climate services and impacts communities, but do not appear prominently in the CMIP6 data request, or variables that have an easily assessable observational counterpart but may not be essential for model intercomparisons). How might user groups such as those illustrated by the high-resolution example in Sec 3.4 be identified systematically, rather than ad-hoc[2]?
    - Future phases: By definition, the existence of this list will create a feedback effect on the most downloaded variables, a core component of its initial derivation. What are the implications for the methodology to update the list going forward? What other issues must be addressed for evolution of the list in the longer term.

- Curation of the list for this phase of CMIP

    - What auxiliary information that is not described in this paper is needed for the full implementation of the BCV list. Where will it be available?
    - e.g. Table A2, A3 details on pressure levels if needed, any other methodological details required for derivation of BCVs. (I would also strongly suggest some version control and numbering).
    - Line 288 Section 3.1 How will new naming conventions be developed and disseminated to the community, or what is needed to address this. Does this need to happen before the AR7 Fast Track runs begin.

- Implications from external/adjacent communities on maximum exploitation of the BCVDR in this phase of CMIP
- Modelling centres and working groups: Are there issues arising from e.g. methodological choices of modelling centres, that are not the responsibility of the BCV list, but that may directly impact its utility (for example, do definitions of mixed layer depth affect how these variables can be intercompared). Are there additional engagement and documentation needs directly relating to BCVs, are there implications for the BCVs from a lack of this engagement, and how can the respective communities collaborate including across the wider CMIP7DR
- Observational community – addressed in earlier comment
- Other neighbouring auxiliary data communities – e.g. downstream modelling exercises, communities using CMIP as boundary conditions, etc. As for observational community, what is needed in terms of engagement on both sides to maximally exploit the BCVDR.

- Immediate next steps of the BCV community

Technical/minor comments

- Ln125: “from”-> “to”?
- Footnote 3 on ECVs. The GCOS ECVs span all observation types including in-situ observations, they are not restricted to Earth Observation.
- Section 3 title: Second “and” should be “of”?
- Table A3 Omon.masscello is missing descriptors in its row

With thanks to the author team,

Claire Macintosh, ESA.

[1] Equivalent numbers from Dimensions.ai (free to access): 5944 articles mention CMIP6 in the title + abstract, of which 2099 (35%) include an observational keyword. This increases to 2538 (43%) if the word ‘evaluation ‘is included, which typically implies some kind of auxiliary data source. Searches conducted 16-Sept-24.

[2] For illustration: Dimensions.ai search “CORDEX” returns 2364 title + abstract results, “CMIP5” returns 6439, but the overlap (CMIP5 AND CORDEX) is only 262, as the majority of the CORDEX community are indirect users of CMIP data. This community will not show up in the methodology as described but is very large and currently not accounted for except via user engagement surveys. The principle of assessment of indirect users is more widely applicable to the BCV concept.

RC1: Author Response

Thank you for the detailed and constructive review.

* Table 1: Yes, we will add reference to "research teams and individual researchers at all career stages" to avoid the unintended suggestion that it only applies to MIPs. We note that neither the usage statistics used as an initial guide nor the process of selecting authors relied on the MIP framework. [Revision RC1.1]

* Implications for exploiting BCVs for comparison with ECVs: This is a big topic. The drafting of this paper has run in parallel with a revision of the GCOS ECVs. At this stage the two processes are independent apart from some communication at an individual level between those taking part. The question of scientific comparison between models and observations is not picked up in this paper. The reviewer makes a valid point about the fact the climate models and observations are increasingly used together, but the engagement approach used to construct our list, which has been praised by the reviewer, necessarily looked at general metrics of utility and did not go into analysis of scientific use of each variable. The standardising of pressure levels has not been discussed by this author team. There is such a discussion within the CMIP AR7 Fast Track data request author team for the atmospheric theme, and also in the impacts and adaptation theme. There is a similar, but independent, discussion about a set of standard ocean layers. This is perhaps a good example of how

more domain specific issues can be better handled in more specialised settings. In this case the more specialised settings are ad hoc groups set up to create the AR7 Fast Track Data Request. [An appendix listing the levels has been added with a reference to the CMIP AR7 Fast Track process: Revision RC1.2]

* One the interests of indirect users: The indirect users have a significant interest in the objective of greater consistency within and among multi-model ensembles. Inconsistencies on primary outputs create problems in generation of products for indirect users. This is reflected on in the discussion of figure 4. Indirect users often use products which depend on multiple parameters, and lack of consistency in the selection of parameters provided can make it difficult to carry out such multi-parameter calculations consistently across many models. [the phrase "indirect users" has been added to discussion of Figure 4 in section 1.3: Revision RC1.3]

* On the link to observational datasets: This is a good question. One aim of this paper is to facilitate discussion of such issues by having the baseline list available ahead of finalisation of the more complete request and divorced from the complexity of specifications about differing experiments and usages. This has not been fully successful, as the Version 1.0 of the data request will be published this November, but the early discussion of this list has raised visibility of some issues. The current structure of CMIP requires community groups to come forward with requests for variables to tackle science goals. These are not explicitly expressed in terms of specific observational variables. Generally, the description of science goals which is provided in the data request does not go into sufficient detail to identify specific observational datasets. [An additional reference to GCOS ECVs has been added: Revision RC1.4'

* To maximise uptake: we need to continue advocacy, both with WCRP and in the broader community, of the list and the role it can play in enhancing consistency and interoperability.

* Gaps in the interface between ECVs and ESM-BCVs: there are many, both at technical, scientific, and governance levels (or in terms of axiology, ontology, and epistemology). for example, at the technical/ontological level there is a need to agree a common, interoperable, syntax for recording our lists. A full analysis is out of scope here.

* The question of future iterations of the list is very open at the moment, and goes beyond the mandate of this author team. [The issues around the process are discussed briefly in a new subsection: Revision RC1.5]

Governance questions.

* For CMIP, the CMIP request will be a larger list which takes this as a starting point. The list provides a reference point. The DR communities need to decide on details of implementation, such as specific advice about high-volume variables. Action from the BCV community is not foreseen (there is some overlap of individuals).

* On revisions: as noted above, this is beyond the remit of this author team.

* Questions about exactly what will be needed for implementation in CMIP7, or any other activity, need to be picked up in CMIP. CMIP is itself an evolving process dependent on many community inputs. We cannot specify a precise set of requirements at this point.

* There are issues around consistency of approach to calculation of BCVs. We hope to encourage more consistent gathering of information both about approaches used and concerns of the community.

* Neighbour communities tend to be involved in the data request. For instance there is an active discussion about sea ice variables for comparison with GCOS sea ice variables.

[These issues are picked up in the new subsection on revisions, including a reference to the fact that new communities are being engaged through the fast track process: Revision RC1.5]

RC1: Manuscript Revisions

Revision RC1.1: We have added a comment about research teams and individual researchers at all career stages in the final column of the first row of table 1.

Revision RC1.2: New appendix 5 added: "Pressure levels for atmospheric variables" (also responding to comments CC2 and CC7 below).

Revision RC1.3: reference to "indirect users" added in Section 1.3, line 119.

Revision RC1.4: brief note on link to GCOS ECVs in lines 470-474, also referring to the wider changes going on in this area.

Revision RC1.5: a new Sub-section "Limitations, extensions, and revision" has been added in Section 4. to discuss future revisions.

RC2: Young Ho Kim

Reviewer Comments

This paper proposes a list of Baseline Climate Variables for Earth System Modelling (ESM-BCV), aimed at enhancing consistency across various modeling projects. With 132 variables derived from the most frequently used elements in the CMIP6 data request, this list promotes the evaluation and utilization of climate simulations, supporting data consistency in future modeling projects, including CMIP7. This paper offers a valuable resource for the climate modeling community and strengthens data consistency. However, it could benefit from additional detail on the selection criteria, weighting, and the importance of high-volume variables. Such additions would enhance the list’s practicality and scope of application. With these revisions, this paper could serve as an essential tool for Earth system modeling research and policy-making. My detailed comments are as follows:

Comments in Detail

- While the paper explains the process for selecting the 132 variables, providing more detail on why other significant climate variables were excluded and outlining criteria for future updates would be beneficial. This additional clarity would assist researchers in expanding or adapting the list.

- For example, including 10m surface eastward and northward winds, 2m air temperature, and 2m specific humidity in the 3-hourly data provides valuable meteorological parameters essential for analyzing near-surface dynamics. However, the absence of downwelling shortwave radiation and cloud fraction in this dataset limits the ability to comprehensively assess ocean-atmosphere interactions. Both downwelling shortwave radiation and cloud fraction are critical for understanding

surface energy fluxes and cloud-mediated radiation effects, which directly impact sea surface temperatures and mixed-layer dynamics. Including these parameters would significantly enhance the utility of the 3-hourly data for accurately evaluating heat exchange processes and cloud-related feedbacks in ocean-atmosphere interactions, providing a more complete picture of the surface energy budget.

- Additionally, including ocean mixed layer thickness in the dataset would greatly enhance the ability to analyze ocean-atmosphere interactions. Mixed layer thickness is a key parameter that influences and responds to surface heat fluxes, wind forcing, and freshwater input, all of which are essential for understanding energy and momentum exchange between the ocean and atmosphere. This variable is also crucial for interpreting subsurface thermal dynamics and stratification changes that affect upper-ocean mixing and biogeochemical processes. Adding mixed layer thickness to the dataset would provide a more comprehensive framework for evaluating how surface conditions drive ocean responses, thereby supporting a holistic approach to studying coupled ocean-atmosphere processes.

- The lack of weighting or prioritization criteria for each selection indicator is noted. Providing specifics on how each criterion influenced the final list would support researchers in developing similar data requests.

- While this list has the potential to enhance interoperability across models, discussing plans to expand it with additional variables necessary for regional modeling or high-resolution climate predictions would be helpful.

- Some variables are marked as "high volume" and can be selectively produced based on available resources. Providing more insight into the critical importance of these high-volume variables would guide users in determining when to prioritize these variables.

Response

1. Plans for expansion are, in the short term, covered by the CMIP AR7 Fast Track Data Request (https://wcrp-cmip.org/cmip7/cmip7-data-request/public-consultation/.) [Covered in Revision RC1.5]
2. We accept that this list does not support comprehensive analysis in many areas: supporting comprehensive analysis will need different lists tailored for different topics. That is the role of the fast track request referred to in item 1.
3. Mixed layer thickness is clearly an interesting parameter, but it is not at this point clear that it is sufficiently well defined to be a priority for model output. At this point the evidence does not point to strong demand for this parameter as a model diagnostic.
4. The selection process has been clarified [Revision RC2.1]

Revision

Revision RC2.1: Clarification of how scores are combined added in lines 219, 220.

Editorial Comment (references)

Please note that your reference list has not been compiled according to our standards. Please consider adjusting your reference list with the next revision of your manuscript. The

manuscript preparation guidelines can be seen at: https://www.geoscientific-model-development.net/for_authors/manuscript_preparation.html

Revision

The reference list has been revised. In addition to formatting revisions, the following corrections were made:

- l86: reference to Paris Agreement corrected.
- l144: WMO expanded and reference list corrected.
- l149: reference to defunct online CMIP6 information removed.
- l197-199: reference to WGCM meeting clarified.
- l189: reference to announcement text corrected
- l266, 395: reference for CMIP3 variable list corrected.
- l408, : references to GCOS and WMO report clarified.
- The list has been expanded in response to comment CC1 (see below)

Discussion Comments

CC1 Treguier

Comment

Dear authors,

Congratulations for this manuscript! I would like to share a suggestion. Some of the variables proposed in the list are not simple physical parameters like temperature. An example is Omon.mlotst, the ocean mixed layer depth. Its computation requires making nontrivial choices. It would be useful to add for each variable a reference to the paper that documents the method, for example Griffies et al., 2016 https://doi.org/10.5194/gmd-9-3231-2016 for ocean variables. If a change in method is decided relative to the existing reference, this change should also be documented and referenced (this may be the case for Omon.mlotst).

Best regards,

Anne Marie Treguier

Response

Accepted. The submitted text cites the immediate source (the CMIP6 request) and the CMIP5 standard output. The point that the intellectual author should also be cited, at least for variables added in CMIP6 should also be cited. In addition to Griffies et al., 2016, this would include CMIP6 GMD papers for DynVarMIP, HighResMIP, Lmon, LS3MIP, SIMIP and VIACSAB.

Revision CC1.1: An additional "Provenance" subsection has been added in "Results".

CC2 Isla Simpson

I have just a minor comment as I was using this paper as I prepared some opportunities for the CMIP7 data request and I couldn't find the information on what the pressure levels actually are for the various options i.e., the 19, 8 and 3 pressure level options for the

atmosphere. I think it would be helpful to have what those pressure levels are actually listed so that this could be a stand alone resource for people to find out about the options available to them from these baseline variables. Sorry if I've missed it somewhere.

Response

The levels may be revised in CMIP7. The way forward may be to name these and then allow for the fact that the high-resolution version might have more levels in the future. See Revision RC1.2 above.

CC3: Alistair Adcroft

I'm surprised to see Oday.sos (surface salinity) but not Oday.zos (table A6). I'm unclear what the purpose of sos is at such high frequency? I believe daily zos (and zostoga) would be more widely used (e.g for local sea-level analysis, mesoscale activity, ...) and should be a baseline variable.

Response

Oday.sos along with Oday.tos was requested in CMIP6 for model evaluation, i.e. comparing with observations. It is included in the ESM-BCV list because it is "considered to be of high importance for characterising the ocean state." Daily zos was not requested in the CMIP6 data request, so we do not have evidence to compare level of interest against sos and tos.

In the monthly CMIP6 data (within which zos, sos and tos are present) the tos field is by far the most popular.

Taking into account the additional sources cited in comment CC4 below we have added Oday.zos.

CC4: Baylor Fox-Kemper

This is a critically important topic, and it will inform all of the CMIP7 results. I have two suggestions (at this moment) for alterations.

1) Omon.zos should be converted to Oday.zos. The daily sea level is important for extreme event diagnosis (as tos and sos are). This variable is a critical one for both impacts and input for downscaling, and is particularly revealing in showing the *failures* of coarse resolution models to reproduce SSH variance as high resolution models do (see Fig. 9.12 of AR6 WGI, panels g-i, which had to be created using resources outside of CMIP6 ones because 0day.zos was not included).

2) There is an issue with only collecting bigthetao, in that most ocean models do not use the TEOS-10 equation of state. McDougall et al. made a recommendation to address this point (https://doi.org/10.5194/gmd-14-6445-202), but the data details for thetao and bigthetao presently do not allow this option (use whichever is the model "native" variable to calculate OHCA.). Thus, at a bare minimum *either* bigthetao or thetao, whichever is the model native, should be in Omon here. Furthermore, there is an ongoing assessment within the OMIP team noting that bigthetao cannot be compared to observations easily, which are presently mostly categorized in observational climatologies via thetao. Thus, even if a model is using bigthetao, a comparison to observations (e.g., AR6 many figures comparing temperature and OHCA to observed temperatures and OHCA in Chps 2, 7, 9, 10, 11...).

Response [accept]

1: Accepted. As noted above in response to CC3, we have added Oday.zos.

2: Both Omon.thetao and Omon.bigthetao are included. A note on the need to provide Omon.thetao in addition to Omon.bigthetao has been added in Table A4.

CC5: Nathan Gillett

Congratulations on this manuscript! I have one suggestion. We expect that emissions-driven simulations will play a bigger role in CMIP7, and expect that more models in CMIP7 will include coupled carbon cycles than in CMIP6. If groups submit emissions-driven simulations, it will be essential to know the simulated CO2 concentration in order to interpret the results; and if groups submit concentration-driven simulations it would be very helpful to be able to diagnose compatible CO2 emissions, for example to calculate remaining carbon budgets. Also, a calculation of compatible emissions in the 1pctCO2 simulations would be needed to diagnose Transient Climate Response to Emissions (TCRE). These calculations would require monthly mean atmosphere-ocean CO2 flux, atmosphere-land CO2 flux, and atmospheric CO2 concentration or mass. These variables are included in the “Constructing a Global Carbon Budget” opportunity, but that opportunity includes a large number of other variables, and it is possible that some modelling centres would decide not to output these variables. I suggest adding this minimal set of carbon cycle variables to the baseline – with the understanding that of course these can only be provided for models with a carbon cycle.

Response

The growing importance of emission driven simulations was largely discussed at length within the authors team, as well as the fact that the identification of all relevant variables to analyse and constrain carbon cycles is an ongoing process. Beside the data request proposed in “Constructing a Global Carbon Budget” opportunity, there are a number of requests dealing with each Earth system realm that will certainly enforce the need to produce at least the variables related to the carbon fluxes. However, we agreed that it is still too early to clearly identify a minimal, consolidated set of carbon cycle variables.

The Baseline Climate Variables is not intended to be complete, but it provides a starting point based on prior evidence matured from past cycles of CMIP. In particular, the BCVs principally deal with physical metrics that enable for a large interoperability across different models and were widely used/requested in previous climate studies.

We did not attempt to pre-judge the broader and emerging requirements of the community and the fact that AR7 Fast Track is requesting additional variables is in line with the intention here. Certainly, the next revision of the baseline will be able to evaluate a range of variables needed to diagnose emission driven runs which are becoming more established in CMIP7.

Addition reference to the ongoing CMIP AR7 Fast Track process have been added to the manuscript in the "2.4 Shortlist Revision and Consequences" section [Revision RC1.5 above]

CC6: Christopher Danek

Hi

Thanks a lot for your efforts! Please see the following comments.

13

1) In 2.2 it's not clear to me how r1 and r2 are defined, i.e. how "downloads" are measured. It is certainly possible to count the number of download-clicks in a browser or the number of wget-scripts generated via a browser. But what about direct data usage via ssh access to an ESGF node, which I assume a lot of scientific users have? That cannot be counted I guess? Also, can the (successful) execution of a wget command be counted? If yes, that means I could tweak the download statistics by running a trillion wget-cronjobs of an unpopular variable? I would like to see a sentence more about this technical aspect (I could not find any details on this in the two given references Fiore et al. 2021 and the ESGF dashboard).

2) In my view it would make sense to add seawater density to the baseline variables. Its an important variable but does not get much attention in the literature, at least this is my impression. At the same time its rather cumbersome to post-process seawater density. 1) Downloading the two high volume 4D variables thetao and so is time consuming. 2) Utilizing a seawater equation software (e.g. gsw from TEOS10) on this large amount of data is time consuming as well. 3) Some ocean model output is not provided on its native grid (`gn`) but horizontally and/or vertically interpolated (`gr`). Hence, if I post-process seawater density from such interpolated thetao and so, the obtained result is a less accurate (?) representation of the actual density during ocean model runtime. I am aware that seawater density would yield a high volume 4D variable but I wonder if its worth to include it due to the above points.

3) I would find it useful to add global averages/sums of important variables to the baseline variables (e.g. tosga, sosga, siarean, siareas, siextentn, siextents, sivoln, sivols) as they are 1) easy to compute for the modeling centers but not for the user (downloading a lot of data is necessary) and 2) only need a tiny amount of resources.

4) I would find it user-friendly if the utilized potential density threshold and reference level were added to the title and/or CF standard name of the mixed layer depth (mlotst), e.g. "... Defined by Sigma T of 0.03 kg m-3 wrt to model level closest to 10 m depth" or such.

5) In the appendix tables, why is "Radiation" a realm and what means "Weighted Time-Mean" (e.g. SImon.siconc)?

Thanks a lot and cheers,
Chris

Response

1. The download statistics are from the server log files which record successful responses to requests received over HTTP, including requests from scripts and from browsers, therefore the wget commands are counted as well.
    - More specifically, these log files are sent in near real-time from each data node to the statistics service which is in charge of processing and aggregating the information.
    - However, currently, there is no mechanism for filtering repeated requests that may come from cronjobs or other tweaking processes: indeed, from the data

statistics perspective, all logs with http status 200 are considered for the metrics calculation, with no ways to control the users' behaviour.

- In addition, multiple access to files held on shared file servers are not included, neither are downloads made using Globus. It is an index of usage which is repeatable and well defined from a provenance perspective, but it is not an accurate measure of usage.
- We have added text to clarify this. [Revision CC6.1]

2. On density, we are, essentially, following Griffies et al. who specified the approach to be taken for CMIP6. While that is unlikely to be the last word on this topic scientifically, it is the most recent review completed for CMIP.
3. While this is plausible, it is not clear that there is widespread demand.
4. We have added to the variable long name (the CF Standard Name is intentionally flexible in allowing arbitrary value of Sigma T)..
5. The qualification "Weighted" is not needed here and has been removed. Radiation: The radiation category is added to distinguish between the properties of the atmosphere itself and the electromagnetic radiation passing through the atmosphere. It is not the same as the cmip modelling realm, so perhaps the heading "realm" could be improved on, but we have not found a better term.

Revision

Revision CC6.1: Clarification of scope of download statistics added in footnote 8 on page 11.

Revision CC6.2: "of 0.03 kg m-3" added to the long name of mlotst.

Revision CC6.3 "Weighted" removed from structure description for SImon.siconc.

CC7: Gaëlle Rigoudy

Congratulations for this reference paper and the impressive work behind!

Here are suggestions from people from the CNRM-Cerfacs modeling group for some adjustments to BCVs list:

- add sfcWind at 3hr along with uas, uas for var association coherency
- add hurs at 3hr along with huss, tas for var association coherency
- remove hurs at 6hr frequency (since now added at 3hr - see previous point)
- useful to have ta daily on P19 along with ua, va, zg, hus for var association coherency
- remove hus, ua, va, ta at daily frequency, on P8 as it is redundant to have them both on P19 and P8 (P8 included on P19)
- add monthly msftyz (MOC) since it is a basic variable not easy to compute offline
- remove pr at 3hr frequency since already requested at 1hr frequency

- add od550aer at monthly frequency to have minimum information about aerosols (integrated content for all species, important to estimate aerosol radiative forcing) at a low cost (2D monthly variable)
- add hus and zg at 6hrPt (along with ta, ua, va), useful to feed the RCM statistical emulators ; provide them on P7h instead of P3 (to have 950 hPa and 700 hPa)

And a general comment: Would be useful to have a table with the list of pressure levels for each pressure level set.

Response

- Omission of Eday.ta appears to have been an oversight. This has been added.
- 3hr sfcWInd was discussed and omitted in order to avoid duplication of high volume elements of the list.
- hurs has lower usage, perhaps because it is a variable which is difficult to compute reliably in models and has been plagued with problems. Until we have evidence that it can be computed reliably it is unlikely to be widely used. The 6hr version has been included despite being well down in the usage ranks, but promoting it further does not appear to be justified.
- Daily ta will be added for consistency. This appears to have been overlooked in past reviews.
- The list provides for both the 8 and 19 level versions. The redundancy can be exploited or avoided by MIPs exploiting the list to create a request. Similarly for the hourly and 3-hourly redundancy. A clarification has been added to the text [Revision CC7.1].
- msftyz and od550aer are important variables in their domain, but have not passed the test of broad interest. Similarly for hus and zg at 6hrPt: the specific use case should be dealt with by a specific request, not through inclusion in the baseline.

Pressure levels: yes, we have listed these in an appendix [Revision RC1.2 above]

Revision

Revision CC7.1: A sentence on redundancy has been added at the end of section 3.1, line 269 to 271 and another sentence at line 427-429 in the new subsection 4.2.

Revision CC7.2: Three variables have been added. Daily zos [addressing CC3 & CC4], daily ta [addressing comments raised here], and monthly rluscs for consistency with other radiation variables [picked up while reviewing anc verifying the two changes above]

CC8 Gavin Schmidt

Comment

I am very conscious of the work that goes into defining these variables and the struggle to keep everyone as happy as possible. Nonetheless, I think there are some important 'meta' considerations that should be informing these choices a little more strongly. These principles come from the notions that a) we are trying to inter-compare models, and b) (where

possible) we should be able to compare to observations on a like-for-like basis. At minimum, the authors need to address how these considerations inform the choices, and if they want to continue with these variables (due to inertia, or other reasons) these should be stated. These principles lead to a number of consequences:

First, diagnostics that are specific to a single model should be discarded. They are (by definition) not comparable to other models or observations. Things I would include here are cloud variables (or really anything) defined on atmospheric model levels - since each model has different levels, these are incommensurate (without doing a lot of work, which might be impossible to do correctly post-hoc). This goes as well for ocean variables on model levels - these should be defined on fixed depths (or more technically) fixed pressure levels.

Secondly, variables that are differently defined in different models and observations are just a recipe for confusion. I would include in this, cloud fraction, or cloud cover variables. In the observations, there are observational constraints that define a minimum optical depth that 'counts' for a cloud (that could be variable in space in time) that is not used in the models (or it might be, and might differ across different models too).

Finally, there should be a greater emphasis on derived variables (i.e. variables for which observations exist, but that aren't prognostic variables in the models).

More specific points:

- Cloud ice and cloud water: These are model conceptions that do not exist in the real world nor in the observations. Any observation of either of these quantities cannot distinguish in-cloud variables from falling precipitation. There is a real danger that naive comparisons of these variables with remotely sensed quantities can lead groups to 'overfit' to biased data which could have important consequences for cloud feedbacks and climate sensitivity. These variables need to be forward modeled using remote-sensing lenses (see next point).
- Cloud-related forward models: Consistent comparisons of cloud properties (ice/water content, fraction, etc.) should be performed using observation-based forward models such as the COSP package. Most groups have implemented this for CFMIP and this should now be standard for the CMIP variables. They have the benefit of standardizing the diagnostics across models for whatever experiment, and in the historical simulations they provide direct comparisons to the satellite record. This should be a no-brainer.
- AMSU/MSU/SSU atmospheric temperatures. These exist as climate data records since 1979, and yet comparisons with models is much harder than it needs to be. These diagnostics can be coded as relatively simple global weighting (with possibly some variation over land and ocean and high topography - but these are minor issues for the trends).
- Ocean heat content. Observations have been sufficient to provide time series over the top 700m and 2000m since the 1960s. These 2D fields should be added to the data request for easier comparison to the observations.

- Derived indices: whether this is done by the model groups, or automatically when the data is ingested, we need to have easy access to key indices (Nino3.4, NAO index, NAM/SAM, IOD., GMST etc.). These are a tiny amount of data compared to the rest of the request, and it's frankly ridiculous that these need to be calculated independently by any researcher.

Response

1. The meta-level view is important, but should include (c) make information more accessible to users from the climate impacts and adaptation communities. This aim can overlap with (a) model intercomparison and (b) comparing with observations, but also brings distinct new considerations. While data on model levels may not be the ideal way of comparing between models, it remains the best approach we have until a standard set of levels is defined and accepted. Such standardisation discussions take place within the atmospheric and oceanographic communities and are beyond the scope of this paper. There may be progress in CMIP7 which could be incorporated into an update of this list. [Reference to the CMIP7 process and ongoing discussion of standardisation of levels has been added: Revisions RC1.2 and 1.5 above]
2. The monthly mean cloud cover is a very highly used variable. There are a number of variables for which there are concerns both about how uniform the implementation is in models and about how strong the equivalence to nominally equivalent observational variables. We have added caveat in the conclusions, together with comments on the relation to observations requested by other reviews. [Revision CC8.1]
3. The issue of cloud ice and cloud water issue is being picked up in the author teams of the AR7 Fast Track data request, together with the related point on cloud forward models. See https://github.com/cf-convention/vocabularies/issues/52 . The variables included here, clivi and clwvi, are clearly not appropriate for comparison with observations. The fact that both are provided by a high number of CMIP6 models implies that they can be readily generated and as such can provide a good basis for model intercomparison. As for item 2, we need to warn against direct comparisons with observations. CMIP7 will include a more complete and clearly documented set of satellite based diagnostics (compatible with ISCCP and MODIS), but these variables have a more limited user community at this point. [See revisions RC1.2 and 1.5 above].
4. Defining new model diagnostics for direct evaluation against satellite radiance measurements may be relatively simple, but it would involve the specification of a new forward model which is beyond the scope of this work.
5. New ocean heat content diagnostics are being discussed by the more specialist team of authors convened for the CMIP AR7 Fast Track request. [Revision CC8.1]
6. Yes. However, calculation, or even enumeration, of such indices is beyond the scope of this paper, but it may be possible to indicate the potential. The challenge is the

need for standardisation and the proliferation of different definitions. There is related discussion here: https://wcrp-cmip.org/event/ref-project-launch/. We have added reference to ongoing work [Revision CC8.3].

Revision

- Revision CC8.1: Caveat about potential incompatibility between observational and superficially related variables added in conclusions, line 473-475.
- Revision CC8.2: Reference to ongoing work aimed at improving CMIP approach to ocean heat content diagnostics added in conclusions, l424-429.
- Revision CC8.3 Reference to WCRP work on indices and the REF work within CMIP7 have been added in the Conclusion section, lines 457-460.