

HARMONISED THEMATIC VARIABLES

Public Consultation Guidance

Five thematic author teams have been set up to develop the controlled list of high priority variables through an IPO-supported and Data Request Task Team coordinated paper writing process, with co-authors representing major stakeholders of each theme.

Document contents

<u>QUICK LINKS</u>	2
<u>BACKGROUND</u>	3
<u>A new structure for the CMIP7 Data Request</u>	3
<u>Harmonised Thematic Variables process and author teams</u>	4
<u>Author teams</u>	4
<u>Public consultation</u>	4
<u>TIMEFRAME</u>	5
<u>Timeline</u>	5
<u>Consultation phases</u>	5
<u>Phase 1</u>	5
<u>Phase 2</u>	5
<u>Phase 3</u>	6
<u>DEVELOPING THE LIST OF HIGH PRIORITY VARIABLES</u>	7
<u>Proposing new content</u>	7
<u>Step-by-step guidance on submitting new proposals</u>	8
<u>Step 1</u>	8
<u>Step 2</u>	8
<u>Commenting on existing and proposed new content</u>	11
<u>Making a decision</u>	12
<u>VIEWING THE DATA AND NAVIGATING AIRTABLE</u>	13
<u>What you need to know:</u>	13
<u>Variables and Physical Parameters</u>	14
<u>What you might notice</u>	14
<u>Navigating Airtable bases</u>	15

QUICK LINKS

- CMIP7 Data Request webpage <https://wcrp-cmip.org/cmip7/cmip7-data-request/>
- Public consultation webpage <https://wcrp-cmip.org/cmip7/cmip7-data-request/public-consultation>
- View glossary <https://bit.ly/CMIP-DR-Glossary>
- Add item to glossary <https://bit.ly/CMIP-DR-Glossary-Add-Item>

Consultation - viewing information in the Data Request bases:

- Data Request Opportunities: <https://bit.ly/CMIP-DR-Opportunities>
- Data Request Physical parameters: <https://bit.ly/CMIP-DR-physical-parameters>
- Data Request Variables: <https://bit.ly/CMIP-DR-Variables>

Consultation -propose new content forms

- **Propose an opportunity** <https://bit.ly/CMIP-propose-opportunity>
- **Propose a variable group** <https://bit.ly/CMIP-propose-variable-group>
- **Propose an experiment group** <https://bit.ly/CMIP-propose-experiment-group>
- **Propose a variable** <https://bit.ly/CMIP-propose-variable>
If there is no existing physical parameter for your variable, you will need to propose this first and allow time for this to go through the spam check
- **Propose a physical parameter** <https://bit.ly/CMIP-propose-physical-parameter>
- **Submitting a reference document** for any of the proposed new content <https://bit.ly/CMIP-new-document>

Consultation -comment on content in the bases -including new content

- **Comment on an Opportunity, variable group or experiment group** <https://bit.ly/CMIP-comment-opportunities>
- **Comment on a variable** <https://bit.ly/CMIP-variable-comment>
- **Comment on a physical parameter** <https://bit.ly/CMIP-physical-parameter-comment>
- **Public discussion via GitHub:** <https://github.com/CMIP-Data-Request>

BACKGROUND

In responses to the [CMIP6 Survey](#), many users raised that they found the CMIP6 Data Request to be too complex with too little documentation. However, the value in a consistent Data Request was also commonly highlighted. The survey raised significant tensions, including:

- Data users wanting more variables but data producers struggling with the size of request, and
- Those wanting fewer updates to the Data Request with others wanting more regular version releases.

A new structure for the CMIP7 Data Request

The [Data Request Task Team](#) have developed an activity that works with community representatives to devise a controlled list of high priority variables that facilitate the majority of user needs. They proposed the following structure of the next Data Request to the CMIP Panel and WGCM Infrastructure Panel (WIP). [The Panel and WIP have both approved the structure.](#)

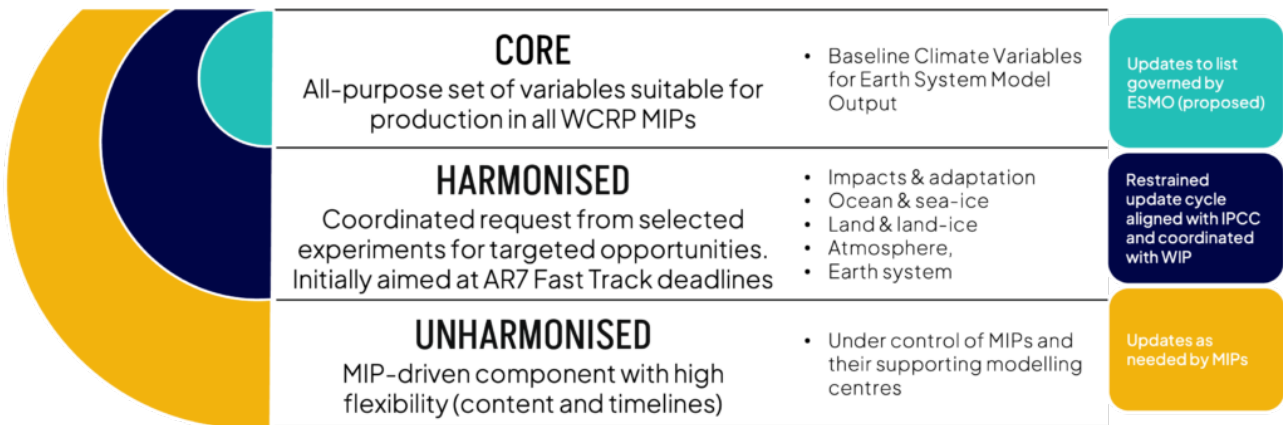


Figure 1 The new Data Request structure consisting of three parts: the Core, Harmonised, and Unharmonised request

CORE

The Baseline Climate Variables list. These variables will be requested as standard across all of CMIP7, forming a base of fundamental variables that are ideally produced by all models for all experiments. A paper outlining the variables, and the process followed to identify the variables will be published in GMD soon.

HARMONISED

A controlled list of high priority variables that facilitate the majority of user needs, while keeping the request as small as possible. This component complements the Core by providing a consistent list of variables across five thematic areas (ocean & sea-ice, land & land-ice, atmosphere, earth system, and impacts & adaptation), which can be mapped onto specific groups of experiments according to the scientific objectives and downstream needs of each case. Five thematic author teams have been set up to develop the controlled list of high priority variables through an IPO-supported and Data Request Task Team coordinated paper writing process similar to that undertaken for the Core list, with co-authors representing major stakeholders of each theme.

UNHARMONISED

This is the component of the CMIP7 Data Request that allows MIPs and other community activities to exploit data request services without being restricted by the deadlines and engagement requirements of the harmonisation process.

Harmonised Thematic Variables process and author teams

Papers across the identified five themes will define a collection of harmonised variables to support the high impact analysis of climate model output and provide a consistent and robust set of parameters on which to build the CMIP7 archive. Each will be a collaboration between the CMIP Data Request Task Team and the MIPs and communities relevant to each theme.

The Data Request will include variable definitions and the mapping of variables against the justification for the request expressed in terms of the opportunities that the data will generate. This will take the form of a controlled list of high priority variables that serve both the majority of user needs and create a harmonised set of data requests which balance scientific demand for data against modelling centre and infrastructure capacity.

The [AR7 Fast Track experiments](#) include a range of dual use experiments which will both provide input for AR7 and serve a wider role as control experiments for a wide range of investigations by community MIPs, many of them running on independent time lines. The Data Request will include data requirements covering both use cases so that those modelling centres wishing to support community MIPs can generate the data that they request efficiently.

This process for developing the Data Request is designed to enable a more transparent and effective approach to variable and request harmonisation and prioritisation. The planned timeline for this process is in Figure 2.

Author teams

An open call was conducted for authors and reviewers for the five thematic papers:

- [Impacts & adaptation](#) (call closed on 30 November)
- [Ocean & sea-ice](#) (call closed on 01 March 2024)
- [Atmosphere](#) (call closed on 08 March 2024)
- [Earth system](#) (call closed on 08 March 2024)
- [Land & land-ice](#) (call closed on 08 March 2024)

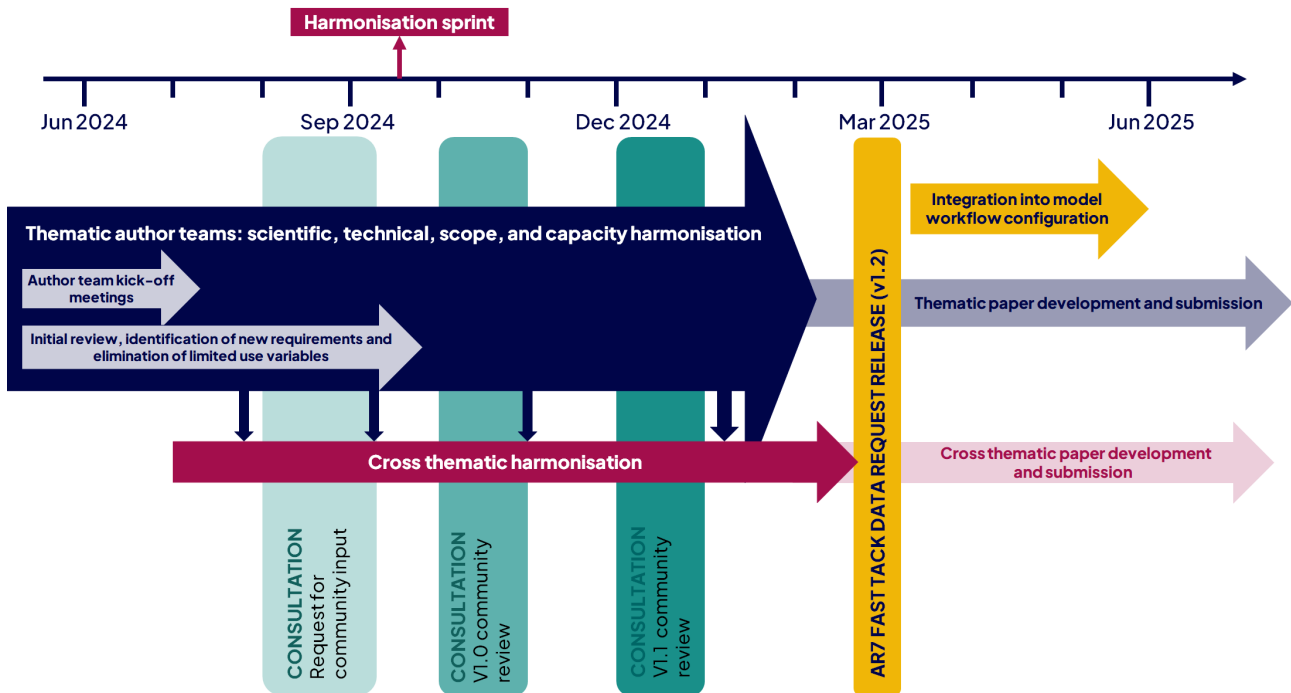
Appointed authors can be viewed [here](#).

Public consultation

In addition to the author teams' work, there will also be three phases of consultation. See more details in the [Consultation phases](#) section below.

TIMEFRAME

Timeline



Consultation phases

There will be three phases of consultation.

Phase 1

Building the data request - community members can engage directly with author teams, as well as input directly using the Airtable forms. There are two key ways to input - submitting proposals and adding comments. See the dedicated [Developing the list of high priority variables](#) section below for more details.

This is an *in real time* consultation. Community members with input are advised to do so as early as possible.

- Real time consultation will remain open until 8th September,
- Author meetings will take place during the week of 9th September to finalise decisions
- Sprint in mid-September - finalise all cross-thematic decisions and possible trigger of further author team prioritisation on decisions if the list is still too long.
- Within 2 weeks of the end of the sprint - finalise and issue the first version of the Data Request.

Phase 2

Version 1.0

ISSUED: 31 JULY 2024

Review of the first version of the controlled list of high priority variables –there will be an opportunity during this phase for MIPs to propose additional experiments and any additional input.

Phase 3

Review of the refined version of the controlled list of high priority variables. This is the final opportunity for any additional input.

DEVELOPING THE LIST OF HIGH PRIORITY VARIABLES

Each thematic author team is using the CMIP6 request as a starting point, eliminating variables with limited use and adding new variables to cover areas of emerging interest or new model capabilities. Alongside discussing individual variables, author teams have been asked to define and review proposals for:

- **Experiment groups:** non-exclusive grouping of experiments (e.g. 'AR7 Fast Track', 'DECK', 'Scenarios' etc.) –experiments can belong to more than one group
- **Variable groups:** non-exclusive grouping of variables (e.g. monthly time slices of the baseline variables) –variables can belong to more than one group.
- **Opportunities:** intended use-case/justification for one or multiple variable groups. Opportunities are linked to relevant experiment groups. Identifying opportunities helps to provide a structure to map variables against requirements. Each opportunity description will convey why this combination of variables and experiments is important and how they contribute to impact.

Proposing new content

During the public consultation, the proposal forms can be used to propose:

- **New Variables:** defining a new variable may involve simple adjustments to existing metadata, it may require a lengthy discussion on the CF discussion forum, or complex balancing of priorities. For new variable proposals, the first step is to propose a **new physical parameter** if it does not already exist before you can submit the new variable proposal form.
 - What is the different between a variable and physical parameter? Read the [Variables and Physical Parameters](#) section to understand.
- **New physical parameters:** If an existing CF Standard Name does not exist to describe your parameter, you should make a [new name proposal](#) to the CF Standard Names discussion list (see [guidance on new standard name proposals here](#) and [rules for CF standard names here](#)). Please add 'CMIP7' at the start of you GitHub issue title. If you are submitting multiple proposals, please group together similar ones to reduce the workload for the CF community. The new standard name proposal process through CF could take up to 2-3 months, depending on the complexity of the proposal. We recommend you propose any required new standard names as soon as possible to give time for any required discussions to take place before the Data Request deadline in October.
- **New Variable Groups:** the variable group includes justification for each set of variables. This can include a MIP set of variables.
- **New Experiment Groups** –propose new groups of experiments –there will be an opportunity to add experiments not in the AR7 Fast Track later during Phase 2, for inclusion in the v1.1 release in December. During Phase 1 proposers can only link to the AR7 Fast Track experiments, but there is the chance to indicate that additional experiments outside the Fast Track will need to be added to the group.

- **New Opportunities:** Identifying opportunities helps to provide a structure to map variables against requirements. Each opportunity description will convey why this combination of variables and experiments is important and how variables and experiments contribute to impact. Each opportunity includes:
 - A high-level description of the science and/or societal use and impact
 - A time slice to specify a block of years for which data is needed. As with the priority and ensemble size, this applies to all data in this opportunity (optional).
 - Experiments (maybe tens) and variables (maybe hundreds) that link to the opportunity. Some experiments and variables will be linked to more than one opportunity.

Each opportunity is intended to cover a broad area of work spanning multiple MIPs and community activities.

There are 5 forms you can use to propose new content:

- **Propose a physical parameter variable** <https://bit.ly/CMIP-propose-physical-parameter>
- **Propose a variable** <https://bit.ly/CMIP-propose-variable>
If there is no existing physical parameter for your variable, you will need to propose this first
- **Propose a variable group** <https://bit.ly/CMIP-propose-variable-group>
- **Propose an experiment group** <https://bit.ly/CMIP-propose-experiment-group>
- **Propose an opportunity** <https://bit.ly/CMIP-propose-opportunity>

There is also a form to **submit a reference document** for any of the proposed new content <https://bit.ly/CMIP-new-document>.

There is a recommended order for proposing new content. Please see the section on [Step-by-step guidance on submitting new proposals](#) below.

Step-by-step guidance on submitting new proposals

Due to dependencies across the different tables in Airtable, there is a recommended order for you to submit new proposals. You may not need to follow every step of this guide (for example, you may not need to submit a new variable). Where this is the case, just skip that step and move ahead to the next step.

Step 1

Check all three bases so you can answer:

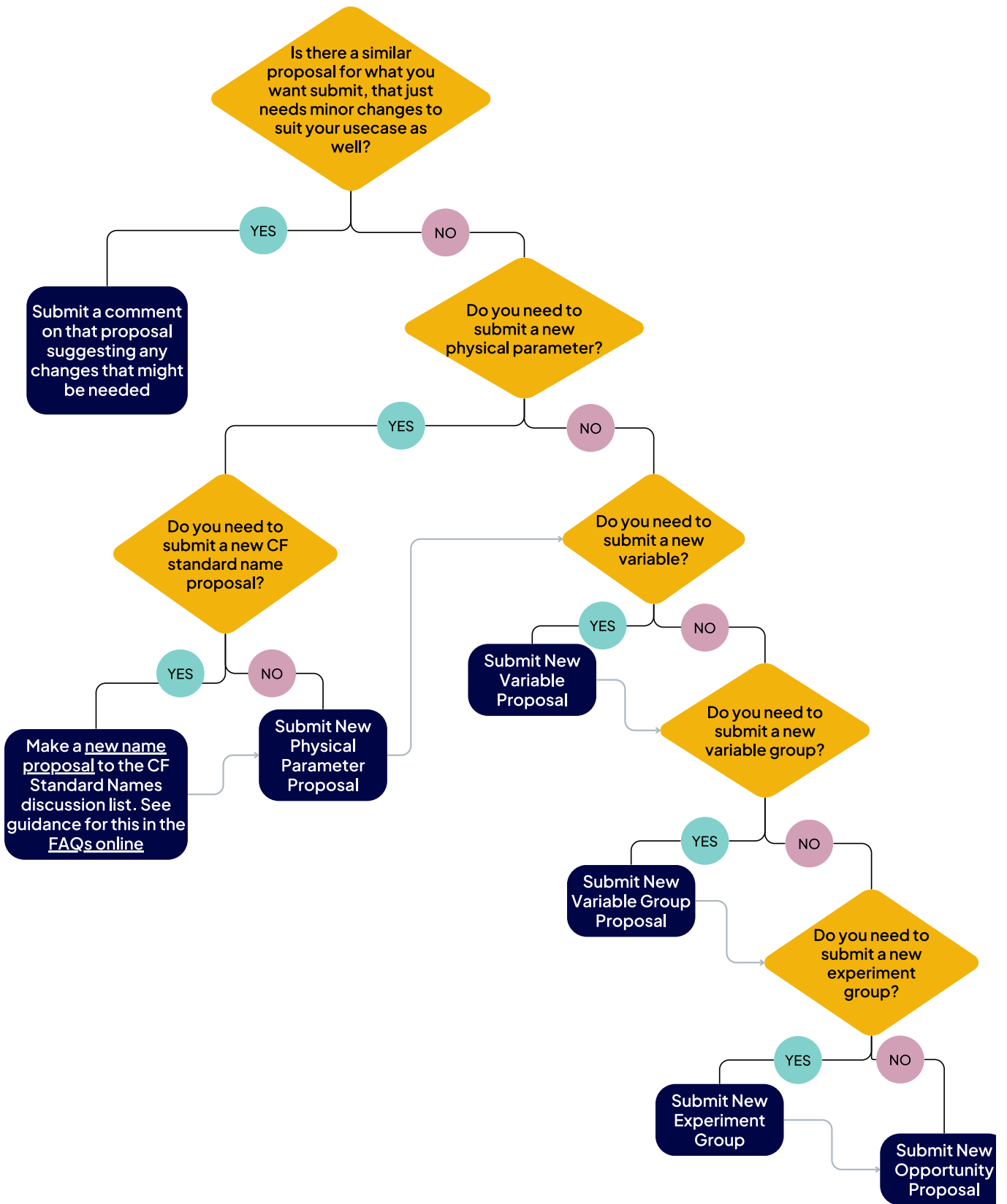
1. Does the Physical Parameters base contain all the physical parameters I need?
 - a. If I need to propose a new physical parameter, is there an existing CF standard name for this, or do I also need to propose a new CF standard name?
2. Does the Variables base contain all the variables I need?
3. Does the Opportunities base contain all the experiment groups I need?
4. Does the Opportunities base contain all the variable groups I need?
5. Are there any submitted opportunities/variable groups/experiment groups that would suit what I need with some minor changes?

Step 2

Once you can answer all the questions above, you can follow the decision tree below to decide which form(s) you need to fill out and in what order. If you need to fill out the comment forms, you can read more information about those in the [Commenting on existing proposed materials](#) section below.

Remember we are aiming for a consolidated, concise list. When proposing materials always ask, ***Is this essential?***

- **What** data is requested, including CF metadata?
- **Why** is it needed and why it is a priority?
- **Who** will make use of it?
- **How** it will be used?



Commenting on existing and proposed new content

There are 3 forms you can use to comment and a GitHub repository to open a point raised in a comment into a public discussion.

Author teams may wish to respond to comments (but they do not need to) if they have collectively identified action that needs to be taken. The action will appear in the Airtable once inserted by either a Data Request liaison or IPO staff member. Anyone looking at the element the comment relates to, will see related comments and any actions noted by thematic teams.

<input type="checkbox"/>	A Title	a protection cons...	A Your name	✉ Your email	🗨 Data Request Actions
1	Incorrect cell methods				Ocean: correction of label required. replace area sum with area mean. Land: agree with ocean, add xxx should also be added to label

There are three forms author teams and consultees can use to comment on existing and proposed content:

- Comment on an Opportunity, variable group or experiment group <https://bit.ly/CMIP-comment-opportunities>
- Comment on a variable <https://bit.ly/CMIP-variable-comment>
- Comment on a physical parameter <https://bit.ly/CMIP-physical-parameter-comment>

There is also a form for submitting suggestions for new glossary items

- View glossary <https://bit.ly/CMIP-DR-Glossary>
- Add item to the glossary <https://bit.ly/CMIP-DR-Glossary-Add-Item>

ACTIVATING A PUBLIC DISCUSSION

Author teams can decide to open a comment for public discussion. To do this, they will create a GitHub issue, which will then appear in the relevant phase of consultation repository on GitHub: <https://github.com/CMIP-Data-Request>. If you submitted the original comment, you will receive an email to let you know a GitHub issue has been opened.

Author team members can either discuss the comment on GitHub individually or as a team. Personal comments will come from their personal GitHub accounts, while team comments will come via official Data Request accounts.

SETTING UP YOUR GITHUB PROFILE

We encourage all community members interacting on GitHub, including authors, to complete your name on your GitHub profile. We also recommend you connect your OrcidID to your GitHub profile.

For author team members it would be helpful if you could notify the IPO of your GitHub username so that we can update our records.

Making a decision

This is a real-time consultation. Author teams can take a decision at any point during Phase 1. If new information becomes available, perhaps through a comment, that decision could change. If you believe an incorrect decision has been made due to insufficient information being available to the author team, you are welcome to submit a new comment to let them know.

Decisions will be finalised in the September 2024 Sprint.

Author teams will be assessing proposals against the following questions:

- **What** data is requested, including CF metadata?
- **Why** is it needed and why it is a priority?
- **Who** will make use of it?
- **How** it will be used?

DECISION TYPES

For proposed content relevant to your theme, author teams can decide to

- Accept
- Reject
- Defer to cross-thematic group review
- N/A – author team has decided the is not relevant for their team

Decisions will be logged in the Airtable next to your proposal (you may need to scroll right to view this) alongside a justification for decisions made.

A Compound Name	or t...	Atmosphere review c...	Cross-thematic team...	Ocean & sea-ice aut...	Land & land-ice auth.
6hrLev.ua			Defer to cross-thematic ...	Reject	Accept

VIEWING THE DATA AND NAVIGATING AIRTABLE

There are three bases (collections of tables) that are publicly available:

- Data Request Opportunities: <https://bit.ly/CMIP-DR-Opportunities>
- Data Request Physical Parameters: <https://bit.ly/CMIP-DR-physical-parameters>
- Data Request Variables: <https://bit.ly/CMIP-DR-Variables>

These are all linked/relational databases.

These bases enable you to view all data within the structure of the data request.

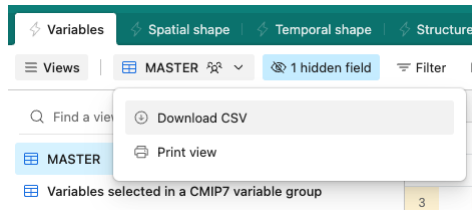
At the top of each base, you see a range of tabs showing tables of information.

- Within a table, you can view the record by clicking the double arrow which **expands** the record.
- Whether you are viewing the data in the table format fields that you can expand for further information have a pale blue square. If you have expanded the record, linked fields show as large rectangles with a white background, with some sub information showing. If in doubt, you can click on a linked record to bring up more information, if you can't click on it, then it is not linked.
- On the right you will see an **Extensions tab** - you can click to open it. At the top you see *Base schema*. You can expand this, and it provides you with a map of how the base is structured and how fields link to each other.

What you need to know:

- **Opportunities base** - propose opportunities and/or variable groups and leave comments on opportunities, variable groups and/or experiment groups.
- **Data Request Variables base** - This is where all CMIP6 variables and proposals for new variables will be appear.
 - In the Variables table are all of the CMIP6 CMOR variables and, as they are added, new variables proposed. You can see which are new variables and which are CMIP6 CMOR variables by checking the Status column.
 - If a variable is a baseline climate variable the column titled 'Variable is included in ESM-BCV v1.3' will contain '1 checked of 1'. Alternatively, you can view a list of the BCVs by selecting the 'BCV Filtered view (v1.3)' on the left-hand side.
 - No variables are assumed to be included in the Data Request. If you need a variable for an opportunity, you must make sure you it is included in a relevant variable group.
- **Data Request Physical Parameters base** - In this you can only see two tables. The Physical Parameters table has both the variables from CMIP6 - called Existing Variables in the *Status* column and newly proposed physical parameters - and the CF Standard Names.

You can export any of the tables to csv if that is useful for you. You can do this by clicking the down arrow next to the View name, just underneath the table tab:



Variables and Physical Parameters

Each Physical Parameter record defines a parameter name, associated with a CF Standard Name. The physical parameters were previously called 'MIP Variables' in CMIP6. The name has changed to 'Physical Parameters' to try and reduce confusion between them and Variables. In CMIP6, the physical parameters were collated in MIP Tables, which were used to organise the variables.

Defining a Variable starts with choosing the relevant physical parameter, then adding additional information/metadata to define more of its attributes.

For example, tas (Near-Surface Air Temperature) is the physical parameter, Amon.tas is the Variable (monthly-mean tas).

What you might notice

The Variables table appears in a number of bases, but you should view this in the Data Request Variables base as this is the MASTER table of variables and allows you to click through and view the additional metadata.

