Geosci. Model Dev. Discuss., 6, 2967–3001, 2013
www.geosci-model-dev-discuss.net/6/2967/2013/
doi:10.5194/gmdd-6-2967-2013

Geoscientific Model Development Discussions
Open Access

Development and exploitation of a controlled vocabulary in support of climate modelling

M.-P. Moine[1], S. Valcke[1], B. N. Lawrence[2], C. Pascoe[2], R. W. Ford[3], A. Alias[4], V. Balaji[5], P. Bentley[6], G. Devine[7], and E. Guilyardi[7,8]

[1]CNRS/INSU, Sciences de l'Univers au CERFACS (CERFACS), URA1875, France
[2]National Center for Atmospheric Science, British Atmospheric Data Centre, STFC Rutherford Appleton Laboratory, UK
[3]STFC Daresbury Laboratory, UK
[4]Centre National de Recherches Météorologiques, Meteo-France, France
[5]NOAA Geophysical Fluid Dynamics Laboratory and University of Princeton, USA
[6]Met Office Hadley Centre, UK
[7]National Center for Atmospheric Science, University of Reading, UK
[8]Institut Pierre Simon Laplace, CNRS, France

Received: 12 April 2013 – Accepted: 21 April 2013 – Published: 23 May 2013

Correspondence to: M.-P. Moine (marie-pierre.moine@cerfacs.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

There are three key components for developing a metadata system: a container structure laying out the key semantic issues of interest and their relationships; an extensible controlled vocabulary providing possible content; and tools to create and manipulate that content. While metadata systems must allow users to enter their own information, the use of a controlled vocabulary both imposes consistency of definition and ensures comparability of the objects described. Here we describe the controlled vocabulary (CV) and metadata creation tool built by the METAFOR project for use in the context of describing the climate models, simulations and experiments of the fifth Coupled Model Intercomparison Project (CMIP5). The CV and resulting tool chain introduced here is designed for extensibility and re-use and should find applicability in many more projects.

1 Introduction

Climate models have experienced outstanding evolution in the last 20 yr driven by scientific improvements and increase in computing capabilities. Additional components of the earth system are being represented with increasing number of physical processes taken into account. Higher spatial resolution is supported thanks to emergence of high performance computing platforms. In addition, more and more research centers engage in climate modelling, which increases the number of models involved. One important consequence is the growth of the volume of data produced. Climate Model Intercomparison Projects (CMIP) initiated and supervised by the World Climate Research Programme (WCRP) are an academic exercise on which climate projection assessment is based. Higher complexity of numerical models, explosion in the volume of data produced, growing number of contributing modelling groups require a dedicated and expert infrastructure for data quality control, data documentation, data storage and access.

The technological part of this infrastructure is in CMIP5 ensured by ESGF (Earth System Grid Federation): several distributed data centers host the data produced by the modelling groups around the world, some of them (PCMDI, BADC, WDCC) being gateways for data publication and download (Williams et al., 2011). It became clear during the set-up of this infrastructure that the definition and adoption of standard metadata (that is data describing the data), is essential for aware data mining and unambiguous data interpretation by end-users (Guilyardi et al., 2011) – even outside the climate modelling community itself, for example by the environment and health impact community. Furthermore, climate metadata must describe both the data content and the model and simulations that produced this data. The CMIP5 metadata standardization effort exploited work conducted jointly by the CURATOR project in the US (Dunlap et al., 2008) and the METAFOR project funded by the European Commission (Callaghan et al., 2010). The approach we followed in METAFOR was to define three key metadata components: a conceptual container to store and organize the information (the CIM, Common Information Model); the possible content (the controlled vocabulary) and a methodology to harvest a specific content, i.e. an instance of metadata (the so-called "CMIP5 Questionnaire"). The CIM is introduced in Lawrence et al. (2012). Here we concentrate on the controlled vocabulary (CV) and the specific harvesting tool developed for CMIP5. We begin by setting the context of earth system models and simulations so to appreciate the challenge raised by climate metadata. We then present a brief inventory of existing metadata systems in climate area pointing out gaps and incompleteness and advocating for a unique and encompassing standard. We describe the methodology applied to build the METAFOR CV and its resulting structure based on key elements. Finally, we explain how this CV was used to construct the "CMPI5 questionnaire".

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◄◄ | ►►
◄ | ►
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

2 Picture of a climate model and climate experiments

Climate study is a highly interdisciplinary science that historically emerged with the convergence of scientific expertise in the research areas related to the earth system such as oceanography, atmospheric physics, sea ice dynamics, hydrology, etc. As a result, a climate model is a composition of models (here after referred to as “components”, some of which map onto “realms” using the nomenclature of Taylor et al. (2011a), each one being devoted to a specific domain of the climate system. These models are generally assembled by coupling software (see Valcke et al. (2012) for a review). The role of the coupler is to exchange coupling fields at the interface of the component domains (for example, wind stress and radiative fluxes are transmitted from the atmosphere to the ocean, sea surface temperature and currents from the ocean to the atmosphere), performing the spatial remapping from the grid of one component to the other. The resulting global model, including components and the coupler, is therefore referred to as a “coupled model”.

A given model can be run and integrated in time (i.e. a climate simulation can be performed) in a large number of different ways, depending on the temporal and dynamical schemes used, and according to the physical parameterizations turned on to model subgrid phenomena within each physical scheme of each component. Initial conditions and external forcing that influence the climate system must be prescribed, e.g. green house gases, volcanoes, aerosols types and concentrations, land-use changes. By adjusting model parameters such as orbital parameters or solar irradiance, by applying appropriate forcing and initial conditions, climate models can be run for various time durations (seasonal, decadal, centennial, millennium) and reproduce different climatic periods (paleo, present and future).

One particular model configuration is usually targeted at a specific scientific question: for example, to understand the sensitivity of a climate process to horizontal resolution or to provide a projection of future climate under a specific emission scenario. Hence, it is important to document not only the particular configuration, but also why that configuration was chosen. The purpose of an experimental protocol like CMIP5 (Taylor et al., 2011a) is to provide guidance for the set-up of models and simulations, so that the different modelling groups address the same questions in a comparable way with their own model. It is clear that the way the model is scientifically configured (including model parameterizations, initial conditions and forcing) and how it conforms to the experimental requirements is crucial information to interpret and compare results. It is therefore utterly important to preserve this information along with the data.

3 Existing Metadata for Weather Forecast and Climate

To ensure interoperability of geo-referenced and weather forecast data products, international organisations like the Open Geospatial Consortium (OGC) and the World Meteorological Organization (WMO) promote adoption of standards. These standards are currently used by national meteorology institutes and production centres of remote sensing and in situ observations all over the world.

One important standard widely adopted by the climate modelling community is the Climate Forecast (CF) convention (http://cf-pcmdi.llnl.gov/). Integrated within the self-documented NetCDF format it forms the CF-NetCDF data format. The CF convention provides a set of standard names for climate variables associated with a precise scientific definition and units. In the CMIP5 framework, CF-NetCDF is the compulsory format for output data set. Futhermore, CMIP5 output metadata are constrained by the CMIP5 tables which impose, among other, short names and units and ensures correspondence with the CF standard names, both for dimensional and physical variables (Taylor et Doutriaux, 2010). Additional low-level discovery metadata (i.e. describing what the data is) are included in the output files as global attributes, for example *experiment_id* or *model_id* that respectively identify the CMIP5 experiment and the coupled model that produced that data set, according to terms defined in the Data Reference Syntax document (DRS) (Taylor et al., 2011b).

Several previous projects, among which NMM (Numerical Model Source Metadata, University of Reading) and NumSim (Numerical Simulation Discovery Metadata, BADC/NCAS, http://proj.badc.rl.ac.uk/ndg/wiki/NumSim) have tried to address the higher level metadata issue, i.e. not only describing "what" are the data produced but also "how" they were produced (the model and simulation details). NMM and NumSim identified some key-terms (e.g. *genealogy, boundary condition type, initial condition type, ensemble type, model component, model category*) and used ISO standards where relevant. Other specific metadata systems have addressed more technical aspects of climate modelling like the configuration of coupling exchanges between earth system components (BFG, Ford and Riley (2011); OASIS4, Redler et al. (2010)) or the grids on with climate model data is discretised (gridSpec, Balaji, 2007). However, no one integrated high-level metadata system able to encompass the whole "climate modelling" process emerged from these projects, leaving only pieces of metadata, often disconnected. In the previous CMIP phase 3, this resulted in asking scientists to provide additional information about models and simulations in unconstrained text-based documents (the CMIP3 questionnaire, see an extract in Appendix A).

4 The Metafor Controlled Vocabulary

Given that the metadata have to address all stages of the modelling process and given that they should serve data discovery and access tools, the prime objective of the METAFOR project was to design a conceptual metadata scheme and develop the associated hosting structure, the Common Information Model (CIM). The CIM defines objects, classes, and their relationships (Lawrence et al., 2012). Through specialized UML (Unified Modelling Language, www.uml.org) packages, the CIM addresses the description of the constitutive elements of climate modelling: the "activity" package includes the experimental context and simulations; the "software" package covers the climate model itself; the final data objects produced by simulations and their inputs are described by the "data" package and the numerical grids of the models by the "grid" package; finally, a “shared” package of reusable elements supports some “orphan” classes such as quality control records and platform descriptions.

To be operational, each individual CIM package needs an associated Controlled Vocabulary (CV) that defines sets of allowed attributes (name/value pairs). For the “data”, “grid” and “shared” packages, the CV was mainly based on a list of already existing terms, respectively the CF standard, gridSpec and some ISO standards. Vocabularies for the “activity” and “software” packages did not exist, and were developed from scratch. In the following we present the resulting “Model Controlled Vocabulary” and the “Simulations and Experiments Controlled Vocabulary”, used in support of CMIP5 to populate the “software” and “activity” packages respectively.

4.1 The Model Controlled Vocabulary

The Model Controlled Vocabulary describes the heart of the climate data production chain that is the numerical model itself. This work to define the Model CV started from scratch and had to go through the early steps of a classical CV building process:

1. Identify the relevant and discriminating information (about the climate model components);
2. Set an ensemble of appropriate terms (meaningful and non ambiguous) to synthetically and faithfully express the information;
3. Organise these terms hierarchically, with possible inter-dependencies;
4. Attach a definition to each term;
5. Identify allowed/possible values for each term.

Following the CMIP5 protocol (Taylor et al., 2011a), the first level decomposition of a coupled climate model was mapped onto eight identified realm components: *ocean, atmosphere, land surface, land ice, sea ice, atmospheric chemistry, aerosol and ocean* *biogeochemistry*. Each realm component is in its turn made of sub-components, one per main physical or dynamical process. Here the components are logical descriptions of the model, not descriptions of the actual software - it is important that users of these CVs understand the distinction, since with the version of the CIM used there is not necessarily a direct mapping between the description of the components and the actual layout in software components.

The way of organising the CV was both driven by typical structure of the numerical models themselves and by the scientific rationale for gathering ideas into main themes, the two being obviously closely related. The current CV granularity is a compromise driven by intercomparison concern: reach a level of details sufficient to be meaningful and discriminating across the various climate models but avoid overloading and too specific information.

The CV could not be established ad-hoc by exploring the model literature alone. The compromise reached is the result of a wide consultation of number of climate modellers led by one dedicated person in METAFOR. The resulting collaboration of a significant number of scientists from the international climate community, each working with different climate models, was a key part of the CV development. More than 35 experts from 13 research centres, representing 6 countries contributed (see list of contributors in Appendix 7), each bringing important scientific expertise to help in identifying the model characteristics important to capture and document for intercomparison. During face meetings or through audio screen-sharing sessions, modellers were asked to tell us about the science and algorithms of the climate model component they developed. The discussions were captured using mindmaps (Freemind software, http://freemind.sourceforge.net/wiki/index.php/Main_Page), one for each realm, which proved to be very appropriate for capturing structured information and feedback on the fly.

The interviewing and reviewing procedure is illustrated in Fig. 1: following a first round interview with one realm expert (step 1), revision processes were launched with others scientists from other research centres (step 5). We integrated the feedback in a structured way (steps 2, 3), capturing their precise meaning, getting confirmation when necessary, working out possible conflicting views (step 6), and taking care not to introduce inconsistencies with previously collected CV. Following this consultation process, several iterations led to a consensus among the modellers interviewed. The resulting CV can be seen as the product of a converging process giving at end, both the content and the granularity of that content. For instance, the case of CV for atmospheric chemistry and aerosol modelling raised some debate within the scientific community since the CMIP5 steering committee had decided to separate them into two different realms. Intensive and rich scientific discussions and exchanges of views were necessary to raise a consensus.

The resulting scientific CV for climate models has three main categories:

1. The CV for the model realm components, including details of the numerical schemes deployed for dynamical processes (advection, diffusion, transport), for time integration and key information about the parameterizations used to model sub-grid scale physical processes (e.g. precipitation and clouds in the atmosphere realm; soil hydrology in land surface realm, gas phase processes in atmospheric chemistry realm); this is the heart of the model CV;
2. The CV associated with the numerical grids used by the models for spatial discretization;
3. The CV for describing the way components are coupled together exchanging coupling fields, including selected terms for spatial regridding and time transformation of these fields; these latter have been derived from vocabulary used for standard configuration of couplers;

4.1.1 Model realm component CV

The complete set of CV for realm components addresses more than 570 leaf parameters over 8 realms[1].

The CV schema adopted for describing the model components has a hierarchical structure we illustrate with the *SeaIce* realm component (Fig. 2). The CV is made of possibly embedded elements: single "leaf parameters" (name/value pairs; e.g. *SchemeType/snow-aging* in Fig. 2) are gathered within "parameter groups" containers (e.g. *Snow* to follow the same example in Fig. 2) themselves gathered within "components" (e.g. *SeaIce_Thermodynamics*). Some groups of parameters are "conditional parameter groups" (e.g. *if VerticalDiffusion is multi-layer* in Fig. 2) depending on the value taken by another parameter (here *VerticalDiffusion*). The tree structure of these different container families define the allowable embedding of the controlled vocabularies and their relationships.

The CV forms a semantic data-base (the possible content) for building a metadata instance (an actual content recorded as a CIM document) for a given model and related simulation. A suite of tools were developed to exploit this semantic data-base in an automatic way so to feed downstream tools such as the CMIP5-Questionnaire (see Sect. 5.1). To that end, coding rules were added to the mindmaps. We defined a set of formal typographic rules, e.g. different font formats and icons to distinguish the different types of CV containers and the different types of choice (exclusive or not) among possible parameter values for parameters or to define the type of expected value (numeric or string). These rules are illustrated in Fig. 2 and detailed in the legend of this figure. Definition of parameters is provided (as attached note, not shown in Fig. 2) and units are prescribed where numeric values are expected.

---

[1] see http://METAFORclimate.eu/trac/browser/controlled_vocabularies/branches/cmip5/Software

4.1.2 Model grid CV

With the model grid CV, METAFOR describes the computational grids of the model components. These grids may differ from the grid the data is expressed on, which, according the CMIP5 guidance, should be described following the gridSpec standard (Balaji, 2007) The model numerical grid CV has to provide information about the horizontal and vertical coordinate system, the vertical coordinate used, the number of levels in the mixed layer and boundary layer, for ocean and atmosphere respectively, etc. A systematic comparison with gridSpec vocabulary was conducted prior to establishing the numerical grid CV so to make re-use of terms when possible. A part of this model grid CV, dealing with the vertical coordinate system, is shown in Fig. 3: according to the value of the *VerticalCoordinateType* leaf parameter, different values for vertical coordinate (e.g.: *sigma* coordinate is proposed only if the type of vertical coordinate is *terrain following*).

4.1.3 Coupling exchanges CV

The CV defined in METAFOR to describe the coupling exchanges between the component models should be considered as an elementary first step. For each exchange, the source and target components are identified, and the coupling CV covers the coupling software used, the type of the spatial regridding and time transformation of the fields (if any). As one can see, the coupling exchange CV is currently quite limited.

4.1.4 Climate model CV evolution and preservation

Even though frozen in the context of CMIP5, we expect that, with usage, this Climate Model CV will evolve, improve and be reused in other scientific projects. Thus, we will have to manage the evolution and ensure the preservation of this CV, which is the first one encompassing all components of a coupled climate model. To that end, it is planned to set up an international governance committee under the auspices of ENES2, the EU-FP7 project that follows on IS-ENES.

4.1.5 Controlled vocabulary for simulations and experiments

Although the Model CV discussed above is valid for any climate model, the vocabulary necessary to describe an experimental framework depends on the experiment context and aims. Unlike for the model description, METAFOR was not asked to define a specific vocabulary for experiments and simulations, this latter being extensively defined in the CMPI5 experiment design document (Taylor et al., 2011a). This document addresses two main sets of experiments, long-term and near-term, further subdivided according to distinct scientific purposes: study of a particular time period (e.g. mid-Holocene, last glacial maximum or 20th century long-term experiments), analysis of the climate response to a given forcing scenario (e.g. volcanic eruptions, anthropogenic aerosols) or evaluation of model errors and statistical significance (e.g. atmosphere-only -AMIP- experiment to identify biases due to coupled mode). Each experiment type is characterised by a set of compulsory requirements and additional recommendations. But even among mandatory requirements, some flexibility remains in their concrete implementation. METAFOR work consisted firstly in encoding CMIP5 defined experiment and simulation vocabulary as specific CV-XML documents so to become machine readable. Secondly, it aimed at capturing the characteristics of a simulation that is left to the person configuring the simulation. Thirdly, it proposed a way to tell how the simulation described meets the experiment requirements it pretends to fit; this is ensured by introduction of the “Conformance” concept.

The Experiment CV-XML documents containing the specific CMIP5 experiment and simulation CV[2] were fixed once for all and cannot be modified by the climate modellers; they are ready for ingestion into the CMIP5 Questionnaire (see next section)

---

[2] see http://METAFORclimate.eu/trac/browser/controlled_vocabularies/branches/cmip5/Activity

and conform to the CIM "activity package" class structure. In such CV, Experiments are identified by a label, a title, a description and an associated list of requirements. Taking the pre-industrial control experiment as an example, *3.1_pi-Control* stands for the experiment label; *Pre-Industrial Control: control experiment against which perturbations are compared* for the experiment title and *Pre-Industrial coupled atmosphere/ocean control run. Imposes non-evolving pre-industrial conditions* for the experiment description. Each requirement has in turn a label, a type and a description attached. To continue with the same example: the requirement with label *3.1.bc.CO2_conc* has *BoundaryCondition* as requirement type and *Prescribed atmospheric concentrations of pre-industrial well mixed gas: Carbon Dioxide* as requirement description.

One or several simulation(s) may support the realization of one particular experiment. Each simulation is identified by a short name, a long name, a description, its DRS member name ("rip" values standing for "realization – initialization method – physics" identifier, see Taylor et al. (2011b)), the name of the model used, the hardware platform on which it has been executed, the start date, time extent, or end date . Among these attributes only model name and the DRS member name is controlled vocabulary (defined within the CMIP5 experiment protocol, as mentioned above). When an experiment requires ensemble runs, one simulation is in its turn described as composed of one or several simulation members, each one being unambiguously identified by its DRS member name ("rip" value). Ensemble type (with the following possible values: *Experiment Driven, Initial Condition, Perturbed boundary Conditions, Perturbed Physics* or *Mixed*) is an additional attribute important for capturing in a standard way the perturbation applied to the ensemble members.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

5 From Controlled Vocabulary to Metadata

5.1 Creating instances of CMIP5 metadata

To collect metadata for CMIP5 numerical models, simulations and experiments, METAFOR has constructed what was initially intended to be a “simple questionnaire”. However, it rapidly became clear that a traditional questionnaire based on a linear collection of information would be completely inappropriate for the task given the amount of information to be collected and given that much of this information would have to be shared and compared, for instance across 2 simulations descriptions. Moreover, a simple-linear text-base questionnaire would have required a huge effort of “by hand” treatment in order to translate information harvested into CIM-instances that at end feed the CMIP5 metadata database (see Sect. 6 for details on information workflow). Thus a more complex tool was needed, and clearly that tool had to be based on the controlled vocabularies defined for CMIP5 and described in Sect. 4. The name has remained, but the “CMIP5 Questionnaire” should be thought of as a complex metadata entry tool, reproducing the CIM syntax structure and syntax and able to make links between metadata objects referring each other.

The resulting Questionnaire provides support for harvesting all aspects a modeler controls when he performs a CMIP5 experiment (see Fig. 4): the model(s) used (included in the coupling system), its associated grids, the computational platform it has be run on, the different simulations performed and the experiment they are related to, the input data files and optionally the CF standard names of the variables in the file used as a model component input. It allows users to interactively produce CIM metadata documents (see Lawrence et al. (2012), for an explanation of the term “document” in this context) without any knowledge of CIM structures. The CMIP5 questionnaire has been built using the python Django web framework (http://www.djangoproject.com/), deployed at the British Atmospheric Data Centre (BADC) and is available on line, at http://q.cmip5.ceda.ac.uk/.

An illustration of how the Model CV is exploited to build the CMIP5 Questionnaire pages is shown in Fig. 5a. The end result is that the structure of the model component pages in the Questionnaire – in terms of, for example, hierarchy presented, order of the parameters asked about – is completely controlled by the originating CV mindmap. This flexibility has of course been crucial in the development of the Questionnaire.

Figure 5a shows the page corresponding to *SeaIceThermodynamics* component taken as example when discussing the Model CV definition process (Fig. 2). The navigation tree on the left provides a hierarchical view of the possible component structure of an earth system model. It strictly reflects the CV structure of the 8 realm components as fixed in the mindmaps. The 3 first frames (from the top of the page) are for generic questions, common to all components (either realm or child): user-defined component names (the component type, here *SeaIceTermodynamics* being fixed) and which grid is used by the current component. The next 3 frames, zoomed in Fig. 5b, contain questions entirely driven by the CV for that component. For example, the 5th frame that asks a question about the SchemeType (*snow-aging*, *snow-ice* or *Other*) mirrors the *Snow* parameter group.

As explained above, the CMIP5 questionnaire helps the modellers to describe their model using the CV. But the Questionnaire is also extensible, offering the possibility for the user to define parameter-value attributes for each component, and indeed arbitrary additional component structures. Obviously, such flexibility is not in line with the current main scope of standardization. Nevertheless, we considered it was important to allow the user to add information that has not been anticipated by the METAFOR CV. Moreover, additional user inputs can help identifying parts of the CV that will need to be completed or changed in an after-CMIP5 perspective.

The questionnaire also uses the specific CV defined for the simulation descriptions. The way a given simulation meets the CMIP5 requirements of an experiment is described by a so-called “Conformance” (see Sect. 4.2). Conformance can be reached via modifications of model inputs, changes in the model parameters, slight modifications of the code itself, or via combination of those. A simulation may even not fully conform to its experiment (for instance when the data producer realizes afterwards, when checking the long list of requirements, that his simulation missed one of them). In this later case “not conformant” is the minimal information to provide. Figure 6 illustrates how the conformance of a simulation named *PICTL* to requirements of the *Pre industrial Control* experiment is capture by the Questionnaire.

6 The information pipeline

It is clear that the METAFOR CV has been built with the intent to go beyond the simple vocabulary collection usage. Indeed, it is targeted for automatic ingestion by downstream tools (the CMIP5 Questionnaire -discussed in previous Sect. 5.1) and for inclusion into OWL (Web Ontology Language) ontologies, e.g. as used in the ESG/CURATOR portal then in use. The Experiment and Simulation CV were fixed by the CMIP5 protocol and are not likely to evolve in the CMIP5 timeframe (hence it was created and stored directly in XML without extra tooling). The CV built for Model description is, on the other hand, intentionally managed in a different way (i.e. in mindmaps, see Sect. 4), independently from the software tools using them. The objective is to ensure separation of concerns between building and usage so that the semantic data base (the Model CV) and the tools using them (the Questionnaire) or hosting them (the CIM) can evolve on their own time line. However, the mindmap format cannot directly feed these downstream tools: format conversion into a machine-readable format was required. To that end, we developed the software to support the information pipeline illustrated on Fig. 7. This tool chain can be found on the METAFOR SVN repository at http://metaforclimate.eu/trac.

To satisfy CMIP5 Questionnaire needs, a simple XML-CV structure was defined to encode the Model CV based on the mindmap rules and constraints described earlier. A mindmap Validator (top-left grey box in Fig. 7), written in XSLT and invoked by Python, was implemented to check that a specific mindmap (top-right red box in Fig. 7) conforms to the defined encoding rules (see Sect. 4.1). If a feature in the mindmap missed a rule (e.g. an element coded as leaf parameter having a child element) the person responsible for the CV mindmap is asked to make appropriate corrections. Once the validation step passed, a mindmap translator (top-right grey box in Fig. 7) re-writes the mindmap information into an XML file (middle-right red box in Fig. 7), suitable for ingestion in the Questionnaire (middle orange box in Fig. 7). These CV-XML documents are then imported into Django tables and are used to automatically build the Questionnaire graphical interface part related to the Model description. Once filled-up, the questionnaire supports three levels of validation (validate in the middle-left, Fig. 7): *(i)* the CV constraints are directly enforced while filling the component description (e.g. a page cannot be saved if a text is provided where a numeric value is expected); *(ii)* when documents are exported in XML files, a validation against the CIM XSD (http://www.w3.org/TR/xmlschema11-1/) is automatically enforced; *(iii)* a Schematron (http://www.schematron.com) based validation is performed to check deeper level of coherency between the different parameters.

To ensure usage by the ESGF gateway interfaces and faceted browsing (Williams et al., 2011), a tool was developed to convert the METAFOR Model CV into an OWL ontology (bottom-right red box in Fig. 7). This ontology was also used to guide the mapping tool which allowed the conversion of CIM documents into gateway RDF (Resource Description Framework,http://www.w3.org/TR/rdf-mt/) triple stores (Lawrence et al., 2012). The conversion of Model CV into OWL was then the decisive step for final adoption of METAFOR CV as CMIP5 metadata for models and simulations (bottom-right yellow box in Fig. 7). At end, CIM-compliant documents, conforming to CMIP5 DRS were broadcasted as “atom feeds”, and the corresponding metadata were ready to be included in the CMIP5 metadata catalogue deployed on ESG portal.

Since the original tool chain was developed, a new toolchain has been deployed. The CIM-compliant XML documents are now stored in a database, and extracted and displayed in client portals via javascript code which loads the documents across the net, and then displays them.

7 Summary and further work

Given that CMIP phase 5 was conducted by 20 modelling groups that produced about 90 000 yr of simulation for a total volume of several petabytes, a CMIP5 climate data user is faced with a large amount and large diversity of data sets archived in CMIP5 data-node centers. In this context, the METAFOR mission was to provide a metadata system to support data preservation, data reuse (both in time and by different research communities), data readability and discovery, and that guarantees the data quality (or conformity). Until now, such an integrated metadata system for climate modelling was missing. The controlled vocabulary for model and simulation should be considered as necessary raw material for such a system.

This paper introduced the controlled vocabulary developed both for generic description of earth system models and as input for the tool developed to collect this description for CMIP5 models and simulations (the CMIP5 questionnaire). The mindmap technology used facilitated the CV development, ensuring a wide engagement of the scientific community in this process, hiding away the complexity of the underlying ontological concepts (the CIM). The metadata pipeline, which starts from the mindmaps, serves both the metadata entry tool (the CMIP5 Questionnaire) and metadata catalogues such as the ESGF gateways. The cornerstone of the METAFOR CV was indisputably the engagement of large number modellers from the climate community since the early stage of the CV elaboration process. The CV collection raised at the end of the METAFOR project gathers thousand of terms, which hierarchical arrangement is not less important than the terms themselves. Even though perfectible, METAFOR CV is the first one addressing the whole climate modelling chain. Available in CMIP5 metadata catalogues and supporting data discovery tools, it is hoped to provide essential services to climate data user.

There are two significant pieces of work yet to be done before the CV can be easily governed and maintained. Firstly, a conversion tool taking the CV XML back to the mindmap format would support a complete round-tripping. This tool would allow use of the CV XML as the primary preservation and governance artifact, generating mindmaps from those XML instances for websites and human mediated discussions for example. Secondly, we need to formalize an http interface to the CV following appropriate standards (see Leadbetter et al. (2011)). Secondly, the maintenance and governance of the controlled vocabulary and of the associated metadata pipeline, needs addressing. Gathering feedbacks from the questionnaire users and finding ways to benefit from these feedbacks to make the CV evolve should also be strongly considered. In minded focus is the set-up of a real standard, which requires a governance committee to emerge (ass planned in the framework of EU-FP7 ENES2 project). For now, the METAFOR work is extended within the UK JISC-funded PIMMS project (Portable Infrastructure for the METAFOR Metadata System).

While the application to CMIP5 has dominated most of the development thus far, the questionnaire is already being deployed with a simpler version of the CV to describe the models and simulations used in the ENSEMBLES EU project (http://www.ensembles-eu.org). Initiated during the METAFOR project, specific CV for statistical and dynamical downscaling methods used in regional climate studies is currently being developed jointly by the US NCPP project (National Climate Predictions and Projections, http://earthsystemcog.org/projects/ncpp/) and EU EURO-CORDEX (Coordinated Downscaling Experiment – European Domain, http://www.euro-cordex.net). One can expect that the METAFOR CV for global climate models will be re-used in recent EU FP7 initiatives dedicated to climate services as SPECS project (Seasonal-to-decadal climate Prediction for the improvement of European Climate).

Appendix A

List of climate scientist involved in the METAFOR consultation process

The METAFOR project members would like to express their sincere thanks to all of the climate scientists who contributed in a significant way to METAFOR controlled vocabularies elaboration process, sharing their knowledge without restriction and providing excellent guidance and recommendations (in alphabetical order):

Abrahams, L., UKCA, UK

Balaji, V., GFDL, USA

Boone, A., CNRM, France

Bopp, L., LSCE-IPSL, France

Braesicke, P., UKCA, UK

Bruehl, C., UKCA, UK

Buja, L., NCAR, USA

Decharme, B. CNRM, France

Déqué, M. CNRM, France

Elkington, M., MetOffice, UK

Fichefet, T., UCL-LLN, Belgium

Gibelin, A.-L., CNRM, France

Goosse, H., UCL-LLN, Belgium

Griffies, S., GFDL, USA

Guilyardi, E., LOCEAN-IPSL, France

Hagemann, S., MPI, Germany

Horowitz, L., GFDL, USA

Hourdin, F., LMD-IPSL, France

Kageyama, M., LSCE-IPSL, France

Khodry, M., IPSL, France

Krinner, G., LGGE, France

Lawrence, B., NCAS-BADC, UK

Madec, G., LOCEAN-IPSL, France

Malyshev, S., GFDL, USA

Mann, G., Univ. of Leeds

Marti, O., LSCE-IPSL, France

Peuch, V.-H., CNRM, France

Polcher, J., LMD-IPSL, France

Ritz, C., LGGE, France

Salas, Y. M. D., CNRM, France

Slawitch, R., Univ. Maryland, UK

Strand, G., NCAR, USA

Van Velthoven, P., KNMI, Netherland

Vancoppenolle, M., UCL-LLN, Belgium

Wyman, B., GFDL, USA

Appendix B

CMPI3 text-based questionnaire

Model Information of Potential Use to the IPCC Lead Authors and the AR4. CNRM-CM3 (version used for IPCC AR4) 2 August 2005

Model identity:

A. Institution, sponsoring agency, country: Centre National de Recherches Météorologiques, Météo France, France

B. Model name (and names of component atmospheric, ocean, sea ice, etc. models): CNRM-CM3

Atmosphere: ARPEGE-Climat version 3

Ocean: OPA 8.1

Sea-ice: GELATO 2

C. Vintage (i.e., yr that model version was first used in a published application): 2004

D. General published references and web pages:

http://www.cnrm.meteo.fr/scenario2004/references_eng.html

E. References that document changes over the last 5 yrs (i.e., since the IPCC TAR) in the coupled model or its components. We are specifically looking for references that document changes in some aspect(s) of model performance.

– descriptions of previous versions of the ARPEGE-Climat model can be found in the following publications:

–Déqué et al., 1994

–Déqué and Piedelièvre, 1995

–Royer et al. , 2002

F. IPCC model version's global climate sensitivity (KW-1m2) to increase in CO2 and how it was determined (slab ocean expt., transient expt–Gregory method, ±2 K Cess expt., etc.): not yet available

G. Contacts (name and email addresses), as appropriate, for:

1. coupled model: David Salas y Melia, David.Salas@meteo.fr
2. atmosphere : Michel Déqué, Michel.Deque@meteo.fr
3. ocean : David Salas y Melia, David.Salas@meteo.fr
4. sea ice: David Salas y Melia, David.Salas@meteo.fr
5. land surface: Hervé Douville, Herve.Douville@meteo.fr
6. vegetation: Hervé Douville, Herve.Douville@meteo.fr
7. other?

Besides atmosphere, ocean, sea ice, and prescription of land/vegetated surface, what can be included (interactively) and was it active in the model version that produced output stored in the PCMDI database?

A. Atmospheric chemistry?

– Ozone transport with simplified chemistry as described in Cariolle and Déqué (1986) and Cariolle et al. (1990).

B. Interactive biogeochemistry?

– no

C. What aerosols and are indirect effects modeled?

– The distributions of marine, desertic, urban aerosols, sulfate aerosols are specified. Marine and desertic aerosols are constant in all experiments. Urban aerosols vary according to estimates between 1860 and 2000. Sulfate aerosols are specified in all experiments according to Boucher and Pham (2002) data, see http://www-loa.univ-lille1.fr/~boucher/sres/ for more details. Note that only the direct effect of anthropogenic sulfate aerosols was taken into account.

D. Dynamic vegetation?

– no

E. Ice-sheets?

– fixed

[...]

Component model characteristics (of current IPCC model version):

A. Atmosphere

1. Resolution: triangular truncation T63 with "linear" reduced Gaussian grid equivalent to T42 quadratic grid

2. Numerical scheme/grid (advective and time-stepping schemes; model top; vertical coordinate and number of layers above 200 hPa and below 850 hPa):

   – semi-lagrangian semi-implicit time integration with 30 mn time-step, 3 hour time-step for radiative transfer;

   – top layer 0.05 hPa, progressive hybrid sigma-pressure vertical coordinate with 45 layers, 23 layers above 200 hPa, usually 7 layers below 850 hPa (less in regions of high orography)

3. List of prognostic variables (be sure to include, as appropriate, liquid water, chemical species, ice, etc.). Model output variable names are not needed, just a generic descriptive name (e.g., temperature, northward and eastward wind components, etc.)

– temperature, northward and eastward wind components, specific humidity, ozone concentration, surface pressure

4. Name, terse descriptions, and references (journal articles, web pages) for all major parameterizations. Include, as appropriate, descriptions of:

a. Clouds:

– statistical cloud scheme for stratiform clouds based on Ricard and Royer (1993). Convective cloud cover based on the mass-flux transport

b. Convection

– mass-flux convective scheme with Kuo-type closure based on Bougeault (1985) boundary layer based on Louis et al. (1982) with modifications by Mascart et al. (1995). SW, LW radiation based on Fouquart and Morcrette parameterizations implemented in a former version of the ECMWF model (Morcrette, 1990, 1991)

c. any special handling of wind and temperature at top of model:

– relaxation of temperature, linear (Rayleigh) friction for wind

Simulation Details (report separately for each IPCC simulation contributed to database at PCMDI)

Picntrl/Run_1

This preindustrial control simulation was initialized from a coupled simulation of a previous version of CNRM coupled model initialized an ocean at rest with temperature and salinity profiles specified from Levitus (1982) climatology, integrated for 30 yrs with a relaxation of surface temperature to the monthly mean Reynolds climatology for 1950. The CNRM-CM3 version was then integrated for 70 yr with preindustrial 1860 greenhouse gases concentrations as a spin-up. After this spin-up period results have been stored from nominal yrs 1930 to 2429.

Acknowledgements. METAFOR was funded by the EU 7th Framework Programme as an e-infrastructure (project #211753). The support of the EU FP7 IS-ENES (project #228203) is also acknowledged. This work benefited significantly from the engagement of other METAFOR members and colleagues from the US Earth System Curator project. We also appreciated guidance from the METAFOR advisory committee, in particular Wilco Hazeleger and Karl Taylor.

The publication of this article is financed by CNRS-INSU.

References

Balaji, V.: Gridspec: A standard for the description of grids used in Earth System models http://www.gfdl.noaa.gov/~vb/gridstd/gridstd.html, GFDL 2007. 2972, 2977

Callaghan, S. A., Treshansky, A., Moine, M.-P., Guilyardi, E., Alias, A., Balaji, V., Bojariu, R., Cofiño, A. S., Denvil, S. Elkington, M., Ford, R., Kolaninski, M., Lautenschlager, M., Lawrence, B. N., Steenman-Clark, L., and Valcke, S.: The METAFOR project: preserving data through metadata standards for climate models and simulations, in INTL-DPIF '10 Proceedings of the 1st International Digital Preservation Interoperability Framework Symposium, No. 6, doi:10.1145/2039263.2039269, 2010. 2969

Dunlap, R., Mark, L., Rugaber, S., Balaji, V., Chastang, J., Cinquini, L., DeLuca, C., Middleton, D., and Murphy, S.: Earth system curator: metadata infrastructure for climate modeling, Earth Sci. Inf., 1, 131–149, doi:10.1007/s12145-008-0016-1, 2008. 2969

Ford, R. W. and Riley, G. D.: The Bespoke Framework Generator, in Earth System Modelling, Volume 3 Coupling Software and Strategies, Series: SpringerBriefs in Earth System Sciences, ISBN 978-3-642-23359-3, 2011. 2972

Guilyardi, E., Balaji, V., Callaghan, S., DeLuca, C., Devine, G., Denvil, S., Ford, R., Pascoe, C., Lautenschlager, M., Lawrence, B. N., Steenman-Clark, L., and Valcke, S.: The CMIP5 model and simulation documentation: a new standard for climate modeling metadata, CLIVAR Exchanges, 16, 42–46, 2011. 2969

Lawrence, B. N., Balaji, V., Bentley, P., Callaghan, S., DeLuca, C., Denvil, S., Devine, G., Elkington, M., Ford, R. W., Guilyardi, E., Lautenschlager, M., Morgan, M., Moine, M.-P., Murphy, S., Pascoe, C., Ramthun, H., Slavin, P., Steenman-Clark, L., Toussaint, F., Treshansky, A., and Valcke, S.: Describing Earth system simulations with the Metafor CIM, Geosci. Model Dev., 5, 1493–1500, doi:10.5194/gmd-5-1493-2012, 2012. 2969, 2972, 2980, 2983

Leadbetter, A., Clements, O., and Lowry, R.: Emerging standards in vocabulary server access methods, Geophys. Res. Abstr., 13, EGU General Assembly 2011 2985

Redler, R., Valcke, S., and Ritzdorf, H.: OASIS4 – a coupling software for next generation earth system modelling, Geosci. Model Dev., 3, 87–104, doi:10.5194/gmd-3-87-2010, 2010. 2972

Taylor, K. E. and Doutriaux, C.: CMIP5 Model Output Requirements: File Contents and Format, Data Structure and Metadata, available at:http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5_output_metadata_requirements.pdf and http://pcmdi-cmip.llnl.gov/cmip5/docs/standard_output.pdf 2971

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, B. Am. Meteorol. Soc., 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2011. 2970, 2971, 2973, 2978

Taylor, K. E., Balaji, V., Hankin, S., Juckes, M., Lawrence, B.N. and Pascoe, S.: CMPI5 Data Reference Syntax (DRS) and Controlled Vocabularies, http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf, 2011 2971, 2979

Valcke, S., Balaji, V., Craig, A., DeLuca, C., Dunlap, R., Ford, R. W., Jacob, R., Larson, J., O'Kuinghttons, R., Riley, G. D., and Vertenstein, M.: Coupling technologies for Earth System Modelling, Geosci. Model Dev., 5, 1589–1596, doi:10.5194/gmd-5-1589-2012, 2012. 2970

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: Description and basic evaluation, Clim. Dynam. Special Issue, doi:10.1007/s00382-011-1259-y, 2011. 2997

Williams, D. N., Lawrence, B. N., Lautenschlager, M., Middleton, D., and Balaji, V.: The Earth System Grid Federation: Delivering globally accessible petascale data for CMIP5, in: Proceedings of the 32nd Asia-Pacific Advanced Network Meeting, 121-130, New Delhi, doi:10.7125/APAN.32.15, http://usymposia.upm.my/index.php/APANProceedings/32ndAPAN/paper/view/155, 2011. 2969, 2983

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

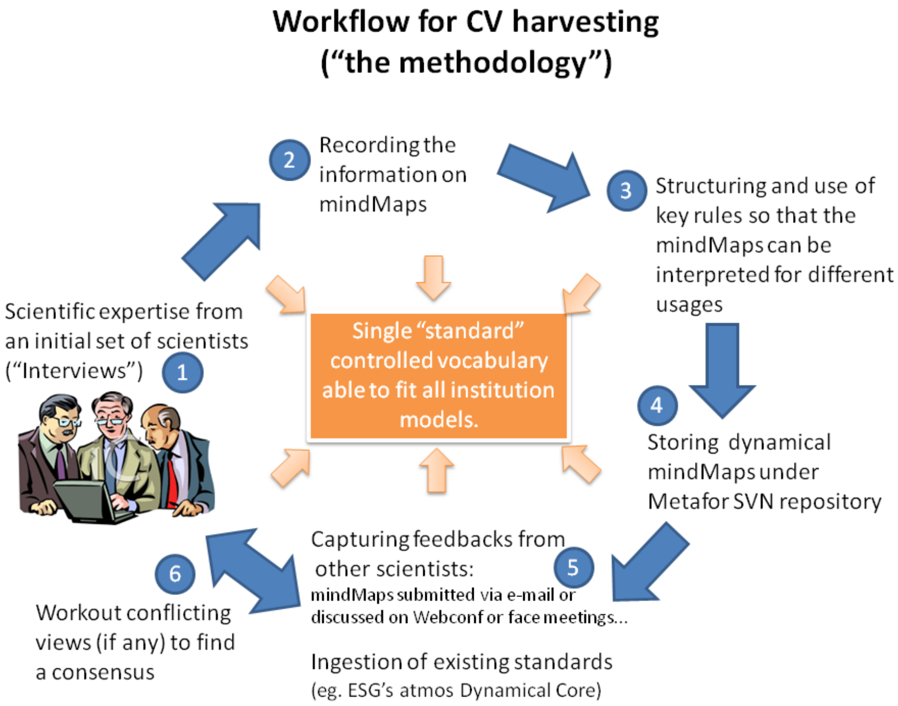


Fig. 1. Consultation process with scientists to define the CV for climate model description.

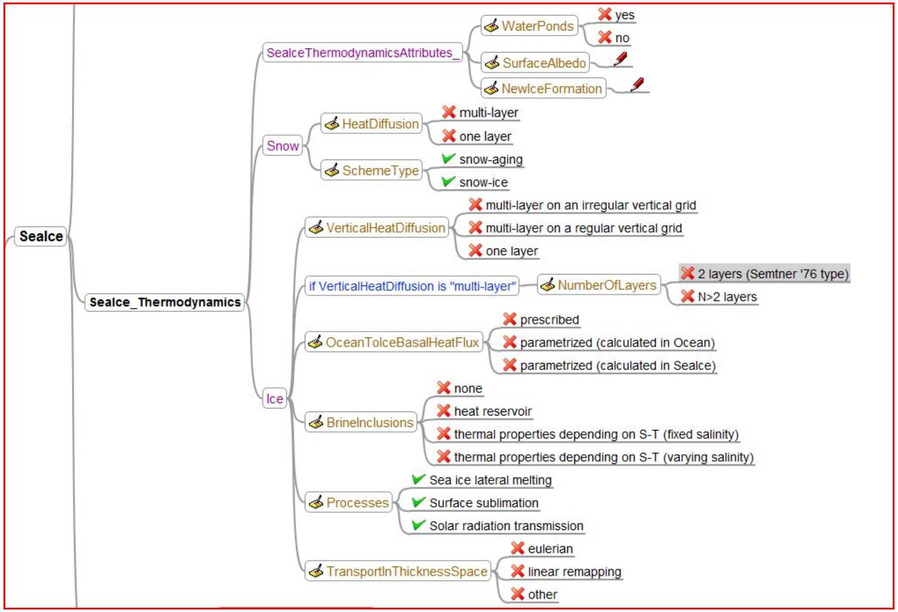


Fig. 2. A portion of the sea ice CV, showing the *SeaIceThermodynamics* sub-component. Black bold font denotes model components; purple is for parameter groups; blue is for conditional parameter groups; brown is for leaf parameters expecting values; black is for possible values for the leaf parameters; red cross icons mark single choice (XOR), green tick mark icons symbolise multiple choice (OR); pencil icons is for free text entry (numeric entries are also possible (not shown); note-book icons ahead of a leaf parameter indicate that a definition is attached as a foot-note.

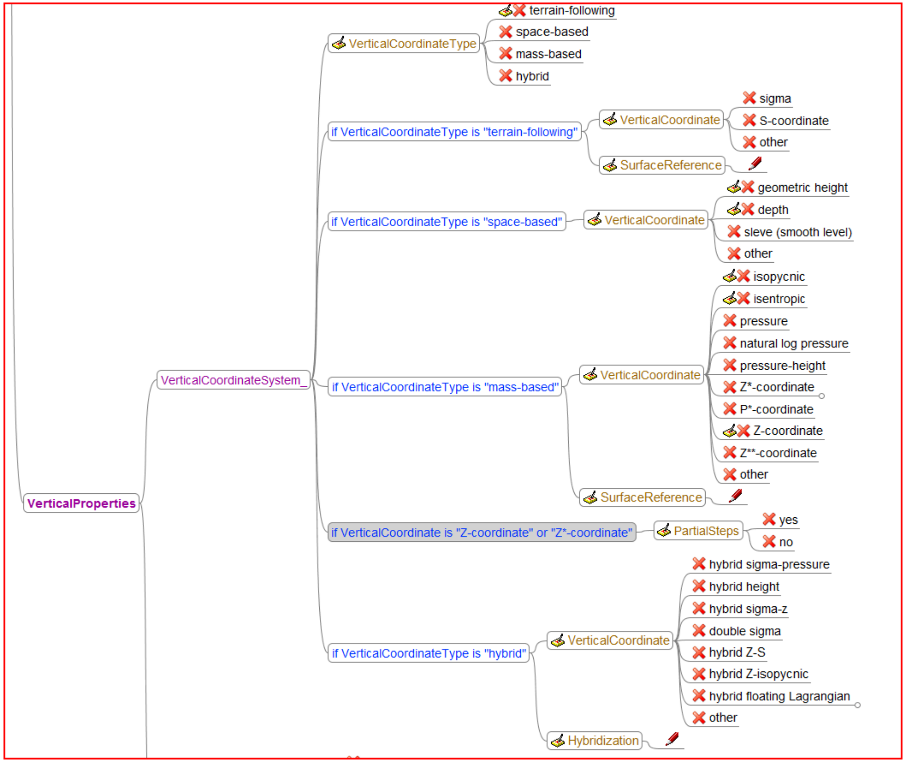
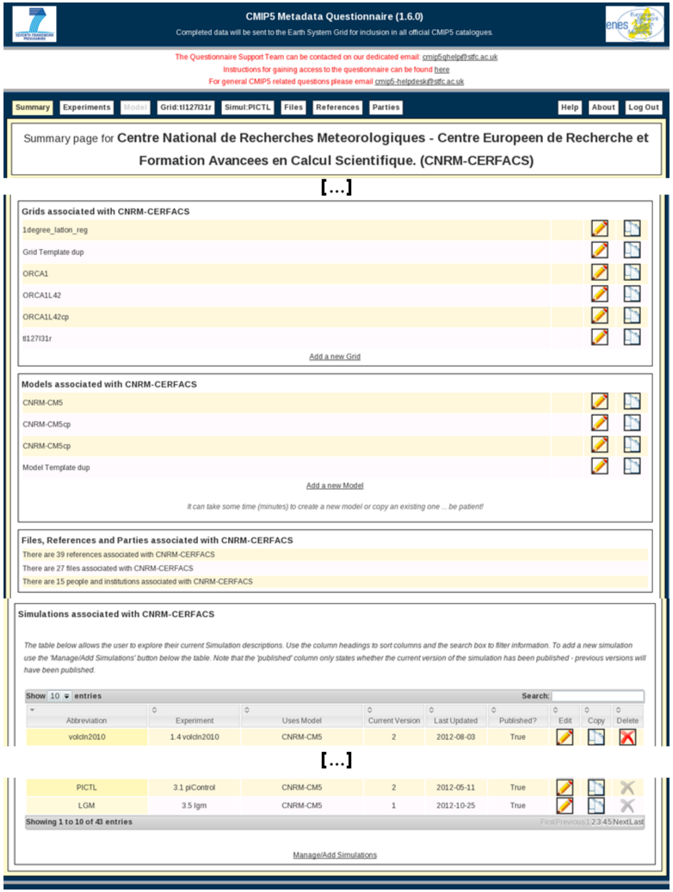


Fig. 3. A portion of the model grid CV. Same typographic rules as in 2 are applied. The parameter group *VerticalCoordinateSystem* gathers information about the vertical coordinate system used by the model.



CMIP5 Metadata Questionnaire (1.6.0)

Completed data will be sent to the Earth System Grid for inclusion in all official CMIP5 catalogues.

The Questionnaire Support Team can be contacted on our dedicated email: cmip5qhelp@stfc.ac.uk
Instructions for gaining access to the questionnaire can be found here
For general CMIP5 related questions please email cmip5-helpdesk@stfc.ac.uk

Summary page for **Centre National de Recherches Meteorologiques - Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique. (CNRM-CERFACS)**

[...]

Grids associated with CNRM-CERFACS

1degree_latlon_reg	✎	⧉
Grid Template dup	✎	⧉
ORCA1	✎	⧉
ORCA1L42	✎	⧉
ORCA1L42cp	✎	⧉
tl127l31r	✎	⧉

Add a new Grid

Models associated with CNRM-CERFACS

CNRM-CM5	✎	⧉
CNRM-CM5cp	✎	⧉
CNRM-CM5cp	✎	⧉
Model Template dup	✎	⧉

Add a new Model

It can take some time (minutes) to create a new model or copy an existing one ... be patient!

Files, References and Parties associated with CNRM-CERFACS

There are 39 references associated with CNRM-CERFACS
There are 27 files associated with CNRM-CERFACS
There are 15 people and institutions associated with CNRM-CERFACS

Simulations associated with CNRM-CERFACS

The table below allows the user to explore their current Simulation descriptions. Use the column headings to sort columns and the search box to filter information. To add a new simulation use the 'Manage/Add Simulations' button below the table. Note that the 'published' column only states whether the current version of the simulation has been published - previous versions will have been published.

Show 10 **entries** **Search:**

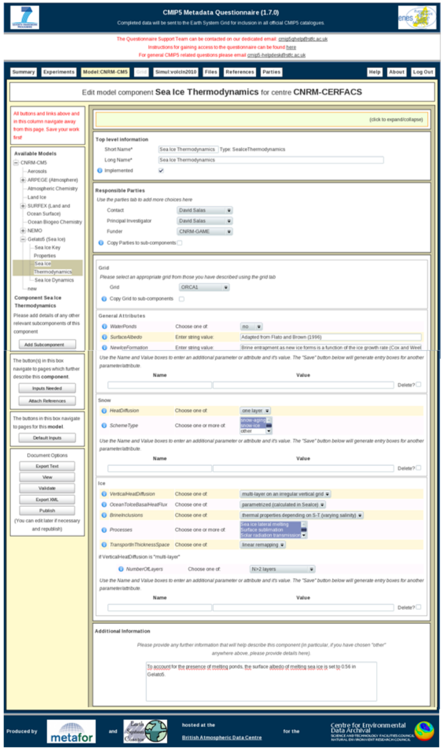
Abbreviation	Experiment	Uses Model	Current Version	Last Updated	Published?	Edit	Copy	Delete
voldn2010	1.4 voldn2010	CNRM-CM5	2	2012-08-03	True	✎	⧉	✗
[...]								
PICTL	3.1 piControl	CNRM-CM5	2	2012-05-11	True	✎	⧉	✗
LGM	3.5 lgm	CNRM-CM5	1	2012-10-25	True	✎	⧉	✗

Showing 1 to 10 of 43 entries

Manage/Add Simulations

Fig. 4. Partial view of CMPI5 questionnaire summary page for CNRM-CERFACS modelling group and their CNRM-CM5 coupled model (Voldoire et al., 2011).

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀◀
▶▶
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion



CMIP5 Metadata Questionnaire (1.7.0)

Completed data will be sent to the Earth System Grid for inclusion in all official CMIP5 catalogues.

The Questionnaire Support Team can be contacted on our dedicated email: cmip5qhelp@stfc.ac.uk
Instructions for gaining access to the questionnaire can be found here
For general CMIP5 related questions please email cmip5-helpdesk@stfc.ac.uk

Edit model component **Sea Ice Thermodynamics** for centre **CNRM-CERFACS**

All buttons and links above and in this column navigate away from this page. Save your work first!

Available Models

- CNRM-CM5
  - Aerosols
  - ARPEGE (Atmosphere)
  - Atmospheric Chemistry
  - Land Ice
  - SURFEX (Land and Ocean Surface)
  - Ocean Biogeo Chemistry
  - NEMO
  - Gelato5 (Sea Ice)
    - Sea Ice Key Properties
    - Sea Ice Thermodynamics
    - Sea Ice Dynamics
  - new

Component Sea Ice Thermodynamics

Please add details of any other relevant subcomponents of this component

The button(s) in this box navigate to pages which further describe this **component**.

The buttons in this box navigate to pages for this **model**.

Document Options

(You can edit later if necessary and republish)

Top level information

Short Name* Sea Ice Thermodynamics Type: SealceThermodynamics
Long Name* Sea Ice Thermodynamics
Implemented ☑

Responsible Parties

Use the parties tab to add more choices here

Contact: David Salas
Principal Investigator: David Salas
Funder: CNRM-GAME
Copy Parties to sub-components ☐

Grid

Please select an appropriate grid from those you have described using the grid tab

Grid: ORCA1
Copy Grid to sub-components ☐

General Attributes

WaterPonds Choose one of: no
SurfaceAlbedo Enter string value: Adapted from Flato and Brown (1996)
NewIceFormation Enter string value: Brine entrapment as new ice forms is a function of the ice growth rate (Cox and Weel

Use the Name and Value boxes to enter an additional parameter or attribute and it's value. The "Save" button below will generate entry boxes for another parameter/attribute.

Name **Value** Delete? ☐

Snow

HeatDiffusion Choose one of: one layer
SchemeType Choose one or more of: snow-aging, snow-ice, other

Use the Name and Value boxes to enter an additional parameter or attribute and it's value. The "Save" button below will generate entry boxes for another parameter/attribute.

Name **Value** Delete? ☐

Ice

VerticalHeatDiffusion Choose one of: multi-layer on an irregular vertical grid
OceanToIceBasalHeatFlux Choose one of: parametrized (calculated in SeaIce)
BrineInclusions Choose one of: thermal properties depending on S-T (varying salinity)
Processes Choose one or more of: Sea ice lateral melting, Surface sublimation, Solar radiation transmission
TransportInThicknessSpace Choose one of: linear remapping

If VerticalHeatDiffusion is "multi-layer"

NumberOfLayers Choose one of: N>2 layers

Use the Name and Value boxes to enter an additional parameter or attribute and it's value. The "Save" button below will generate entry boxes for another parameter/attribute.

Name **Value** Delete? ☐

Additional Information

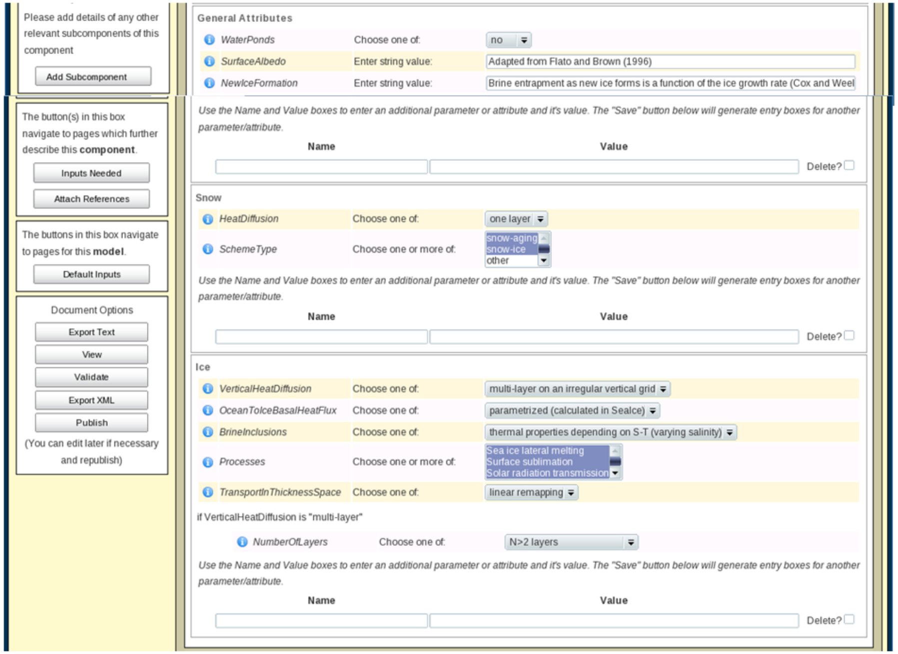
Please provide any further information that will help describe this component (in particular, if you have chosen "other" anywhere above, please provide details here).

To account for the presence of melting ponds, the surface albedo of melting sea ice is set to 0.56 in Gelato5.

British Atmospheric Data Centre

Fig. 5a. How the Model pages of the CMIP5 Questionnaire automatically inherits from the CV mindmap organization. Components hierarchy (realm and child components) designs the model navigation tree (left column); each model component mindmap provides the content of the corresponding questionnaire page and parameter groups in the component mindmap determine the frames in the page; mindmap leaf parameters define the requested information lines in the frames; list of possible CV values for a given parameter forms the content dropdown menus.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion



Please add details of any other relevant subcomponents of this component

The button(s) in this box navigate to pages which further describe this **component**.

The buttons in this box navigate to pages for this **model**.

(You can edit later if necessary and republish)

General Attributes

WaterPonds	Choose one of:	no
SurfaceAlbedo	Enter string value:	Adapted from Flato and Brown (1996)
NewIceFormation	Enter string value:	Brine entrapment as new ice forms is a function of the ice growth rate (Cox and Weel

Use the Name and Value boxes to enter an additional parameter or attribute and it's value. The "Save" button below will generate entry boxes for another parameter/attribute.

Name	Value	
		Delete? ☐

Snow

HeatDiffusion	Choose one of:	one layer
SchemeType	Choose one or more of:	snow-aging snow-ice other

Use the Name and Value boxes to enter an additional parameter or attribute and it's value. The "Save" button below will generate entry boxes for another parameter/attribute.

Name	Value	
		Delete? ☐

Ice

VerticalHeatDiffusion	Choose one of:	multi-layer on an irregular vertical grid
OceanToIceBasalHeatFlux	Choose one of:	parametrized (calculated in SeaIce)
BrineInclusions	Choose one of:	thermal properties depending on S-T (varying salinity)
Processes	Choose one or more of:	Sea ice lateral melting Surface sublimation Solar radiation transmission
TransportInThicknessSpace	Choose one of:	linear remapping

if VerticalHeatDiffusion is "multi-layer"

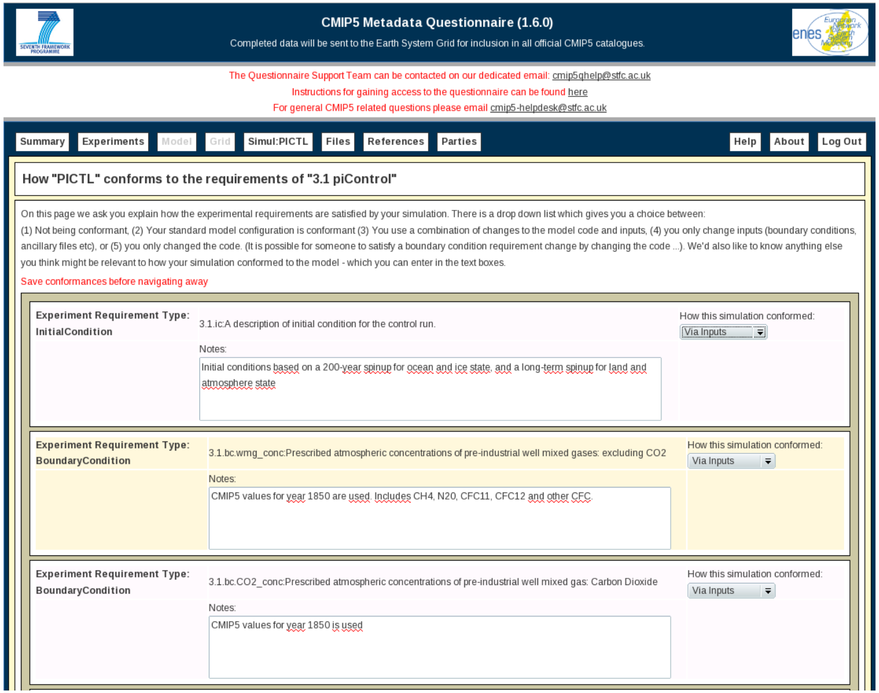
NumberOfLayers	Choose one of:	N>2 layers

Use the Name and Value boxes to enter an additional parameter or attribute and it's value. The "Save" button below will generate entry boxes for another parameter/attribute.

Name	Value	
		Delete? ☐

Fig. 5b. Zoom of the 4th, 5th and 6th frames of Fig. 5a.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion



CMIP5 Metadata Questionnaire (1.6.0)

Completed data will be sent to the Earth System Grid for inclusion in all official CMIP5 catalogues.

The Questionnaire Support Team can be contacted on our dedicated email: cmip5qhelp@stfc.ac.uk
Instructions for gaining access to the questionnaire can be found here
For general CMIP5 related questions please email cmip5-helpdesk@stfc.ac.uk

How "PICTL" conforms to the requirements of "3.1 piControl"

On this page we ask you explain how the experimental requirements are satisfied by your simulation. There is a drop down list which gives you a choice between:
(1) Not being conformant, (2) Your standard model configuration is conformant (3) You use a combination of changes to the model code and inputs, (4) you only change inputs (boundary conditions, ancillary files etc), or (5) you only changed the code. (It is possible for someone to satisfy a boundary condition requirement change by changing the code ...). We'd also like to know anything else you think might be relevant to how your simulation conformed to the model - which you can enter in the text boxes.

Save conformances before navigating away

Experiment Requirement Type: InitialCondition
3.1.ic:A description of initial condition for the control run.
How this simulation conformed: Via Inputs
Notes: Initial conditions based on a 200-year spinup for ocean and ice state, and a long-term spinup for land and atmosphere state

Experiment Requirement Type: BoundaryCondition
3.1.bc.wmg_conc:Prescribed atmospheric concentrations of pre-industrial well mixed gases: excluding CO2
How this simulation conformed: Via Inputs
Notes: CMIP5 values for year 1850 are used. Includes CH4, N20, CFC11, CFC12 and other CFC.

Experiment Requirement Type: BoundaryCondition
3.1.bc.CO2_conc:Prescribed atmospheric concentrations of pre-industrial well mixed gas: Carbon Dioxide
How this simulation conformed: Via Inputs
Notes: CMIP5 values for year 1850 is used

Fig. 6. Illustration of "Conformance" concept in the case of a *PICTL* simulation performed with CNRM-CM5 model in the framework of *3.1 piControl" CMIP5 experiment*. For the three requirements shown, the conformance is ensured *via inputs*, which means that the input files used contain the forcings requested. Note that free-text place to enter additional details is always provided.

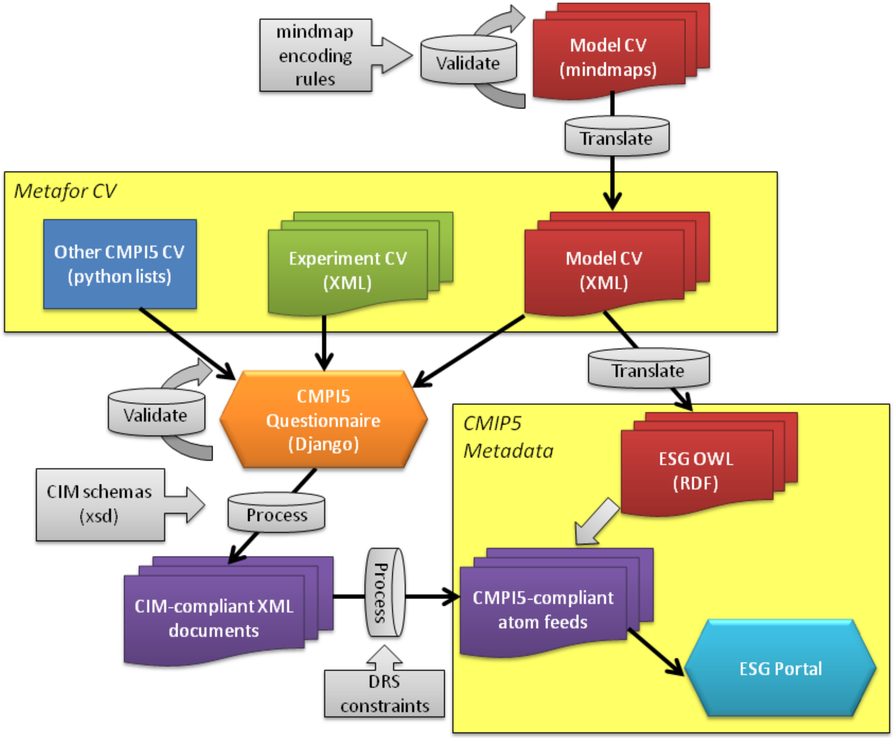


Fig. 7. Key components of the CV and information pipeline from METAFOR CV (top yellow box) to CMIP5 metadata (bottom-right yellow box).