https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019


Geoscientific Model Development Discussions
Open Access
EGU

# Current status on the need for improved accessibility to climate change models

Juan A. Añel[1], Michael García-Rodríguez[1,2], and Javier Rodeiro[2]
[1]EPhysLab, Ed. Campus da Auga, Campus As Lagoas, 32004, Ourense, Galicia (SPAIN)
[2]School of Computer Sciences, Campus As Lagoas, 32004, Ourense, Galicia (SPAIN)
**Correspondence:** Juan A. Añel (j.anhel@uvigo.es)

**Abstract.** Over the past few years, increasing attention has been focused on the need to publish computer code as an integral part of the research process. This has been reflected in improved policies on publication in scientific journals, including key related issues such as repositories and licensing. We explore the state-of-the-art of code availability and sharing of climate models, using as a testbed the models from the Climate Model Intercomparison Project (CMIP5), and we include some particular reflections on this case. Our results show that there are many limitations in terms of access to the codes of these climate models, and that the climate modelling community needs to improve its code sharing practice in order to comply with best practice in this regard, and the most recent editorial publishing policies.

## 1 Introduction

Reproducibility of results is essential in scientific research. Because so much scientific output today relies on the use of computers, there are new requirements in terms of the description of any experiments performed, in order to assure computational scientific reproducibility (CSR). This is widely known (Añel, 2011) and was recently discussed in a Sackler Colloquium on "Reproducibility of Research: Issues and Proposed Remedies" (Allison et al., 2018). The complexity of the problem, where in some cases scientists may be unaware of some of the determinants, or may make subjective judgements that have little to do with the most appropriate from a scientific point of view (Joppa et al., 2013), or may even fail to make the correct assessment, makes it necessary to consider a range of issues (Añel, 2017), including legal aspects. Recent examples have revealed some very low levels of CSR (Allison et al., 2018; Stodden et al., 2018). Steps are being taken to improve matters, e.g., an increasing number of journals now have computer-code policies (Stodden et al., 2013; GMD Executive Editors, 2015; Nature, 2018), and recommendations have been made to ensure greater reproducibility of results (Wilson et al., 2017).

The study of climate change relies heavily on the use of large computer simulations with geoscientific models of varying levels of complexity. In projects involving the intercomparison of climate models and in some research papers, it has become increasingly common to provide details of the simulations performed, including initial configurations, which are generally clear, accessible and formalised, in related outputs with digital object identifiers (DOIs) (e.g. Eyring et al. (2016); Morgenstern et al. (2017)). However, it is somewhat perplexing that the codes of the underlying models are not always made available or at best they are shared informally, using links, repositories without any security regarding long-term availability or access, or email addresses via which it is claimed that the code will be delivered after contact. Especially in a field where heated debates occasionally arise following the publication of results, it seems odd that this core element of the research is not made more widely accessible.

There are other reasons that justify the need for access to the codes of climate models used in scientific research. One of the most important is to prevent the loss of knowledge on the cycles of development of these models. Some of them nowadays rely on 'legacy' code that was written up to five decades ago, and new developers must understand why some decisions on implementation were undertaken so long ago. There is both an educational and practical dimension to this issue. In some cases, different models share sections of code but its development remains fairly obscure (Knutti et al., 2013). It could be said that adequate sharing and documentation is not necessary if the code used in the models includes appropriate comments, but it is generally the case that climate models do not comply with what would be the ideal level of programming practice. Indeed, the incidence of comments throughout the code is very low, and programmers have tended to perform very badly in this regard in particular (García-Rodríguez et al., 2019). The replicability of results in different computing environments can also be difficult and should not be expected by default (Easterbrook, 2014), even where the same model is used (Massonnet et al., 2019).

Some informal efforts have been made to document accessibility for some climate models (Easterbrook, 2009; RealClimate.org, 2009) and others more formally to check their quality (e.g., Pipitone and Easterbrook (2012); García-Rodríguez et al. (2019)). In light of these efforts, in this study our intention was to test the current status of accessibility to the most commonly used global climate models, in particular those that have contributed to the Climate Model Intercomparison Project (CMIP). In the sections that follow, we describe our efforts to gain access to these models, the procedures we followed, and a classification of the models according to some metrics related to accessibility, and we also provide a discussion containing reflections on the state-of-the-art.

## 2 Methods

In our attempt to better understand the current status of CSR and the availability of climate models, we used as a testbed the models of CMIP5 (Taylor et al., 2012) in view of their extensive use in climate research over the last five years. These served as a key tool for the last IPCC AR5 (IPCC, 2013), and given the ongoing development of CMIP6, groups of modellers should now be more open to sharing the code, due to the possible depreciation of the earlier version. We followed a standard procedure to obtain the code of each model, firstly by checking the information available on each model in the webpage of CMIP5, and contacting research groups where necessary using a variety of different approaches, ranging from non-disclosure of ourselves as climate scientists to full explanation of our interest in studying the code. These approaches are detailed in the following sections.

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019



### 2.1 Survey methods

Using a systematic methodology, we attempted to obtain the codes of all the climate models involved. This procedure included a first step using the web addresses indicated on the CMIP5 webpage. Where this was not enough, we searched through the internet and institutional web pages for open repositories. In a few cases this was sufficient (see Table 1), but in others we had to proceed by making contact with development teams at different levels (emails (see Appendix) in English and French (with folow-up emails two weeks after the first contact).) For the NASA-GMAO model, after failing to get a reply from contact via email, we discussed it at a conference and the development team granted us access.

For those cases where we needed to establish contact via email, we provide details in the Table 1 of the different replies that we received. The first email was always sent from his student email address (under the domain esei.uvigo.es) by Michael García-Rodríguez, who had had no previous involvement in the activities of the international climate modelling community. The idea behind this was to check whether after it had become obvious that the models were not available easily, institutions and researchers would then share them with someone from the general public. In the end, to assure CSR and accessibility, details of experiments must be open to everybody, not just to peers or other scientists. In the final email sent, we identified ourselves and our team, to make clear that we were indeed climate scientists, and thus to check whether we thus had a better chance of obtaining the code. Where access to the code was denied, we sent a survey with a few questions to better understand the reasons for this. All emails sent followed the same template as that given in the Appendix A.

## 3 Results

After all attempts and several months, we successfully gained access to 10 out of 26 models (27 out of the 61 model versions or configurations) contributing to CMIP5. Table 2 provides a summary of the details of the replies obtained from these centres, teams or contacts that allowed access to the code. In terms of research centres or groups contributing to the CMIP5 project, this also represents 10 out of a possible 26. We found a strong regional bias in terms of the countries where models were made accessible. The USA, Germany and Norway stood out as the best contributors in that we obtained the code for all their models (though Norway only contributes one). Together they represent 38% of the research centres or CMIP5 models and 44% of all the versions. For France, we gained access to one of two models (three out of five versions). This analysis is relevant, because in some cases the decision on whether to share the code of the models could have been due to national or regional regulations on software copyright, intellectual property, etc. Figure 1 shows the percentage of models obtained from a global perspective, with specific plots for Europe and Asia. This makes it easier to visualise the rather narrow distribution of the regions on the maps and because different countries could apply different national laws in order to share the codes of the models.

In some cases, a great number of email exchanges were required over periods of time longer than one week to receive a reply or the code. In six cases, there was no obvious way to contact the development teams, in five cases we received no reply at all, and in seven research centres (corresponding to eighteen models) they replied that they did not share the codes of their models. We decided to include in this final group EC-Earth, for which the code is said to be available to a given group of users, but in practice the procedure to access it makes it completely unfeasible for non-members of the regular team involved in its

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019

Geoscientific Model Development
Discussions
Open Access

development. In no case did we receive a response to the questionnaire sent asking for the reasons why they did not want to share the code. For the models obtained, we performed a ranking as shown in Table 3, taking into account licensing issues and availability for reuse by third parties, among other factors. We considered the level of requirements introduced by the GPLv3 license (https://www.gnu.org/licenses/gpl-3.0.en.html) as the ideal case, or a license under which the model can be shared, modified and used without restriction. This is in line with the recent updates to the policy on code availability published by Geoscientific Model Development (GMD Executive Editors, 2019).

We also addressed other issues relevant for running the models. In some ways, accessiblity or ability to gain access to the code means nothing if adequate documentation for the model, a description of its components, instructions on how to compile or run it, and basic examples are not provided. This is in line with recommendations contained in the literature (Lee, 2018). The results are shown in Table 4 and it can be seen that almost all the models obtained comply with all these criteria, with the exception of NICAM.09, which only includes a 'Readme' and a 'Makefile'. For the IPSL, although the link to access the documentation does not work, it is possible to gain access to it by performing an internet search.

## 4 Conclusions

It is widely acknowledged that some scientists are reluctant to share code because of the perceived potential damage to their reputations. Given that many scientists have no formal training as programmers, it may be presumed that they consider that their code may not comply with the standards of excellence that they usually pursue in their main fields of knowledge. Indeed, it has been clearly documented that some climate scientists acknowledge that imperfections in climate models exist, and they simply address them through continuous improvement without paying too much attention to the normal techniques of software development (Easterbrook and Johns, 2009). Nevertheless, all scientists must believe that their code is good enough (Barnes, 2010) and that there are thus no reasons not to publish it (LeVeque, 2013). Barriers to code-sharing through licensing, imposed by e.g., government bodies, cannot be an excuse and when contributing to scientific studies and international efforts where collaboration and trust are critical, such practice is not acceptable. For cases where we obtained the code of a given model, we were not provided with a reason for the license behind it. In fact, in some cases despite getting the code we did not see a license explaining clearly the terms of use.

It is a matter of some regret that we obtained straightforward access to just 3 of the 26 models (7 of the 61 versions) in CMIP5 and that for 16 (34 versions) we were not able to obtain the code at all. For all others, some interaction was required, from email exchanges to personal discussions at workshops. Indeed, we did not get access at all to the codes for more than half of all the versions used in the CMIP5 despite identifying ourselves as research peers. Therefore, we have to report the very poor status of accessibility to climate models, which could generate serious doubts for the reproducibility of the scientific results produced by them. While there is no reason to doubt the validity of the results of the study of climate change obtained using the CMIP5 models (in a similar way to findings for other disciplines (Fanelli, 2018)), we encourage all model developers to improve the availability of the codes of climate models and their CSR practices. Previous work has already shown that there is room for significant improvement in the structure of the codes of the models, which is in some cases very poor (García-

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019


Geoscientific Model Development Discussions
Open Access
EGU

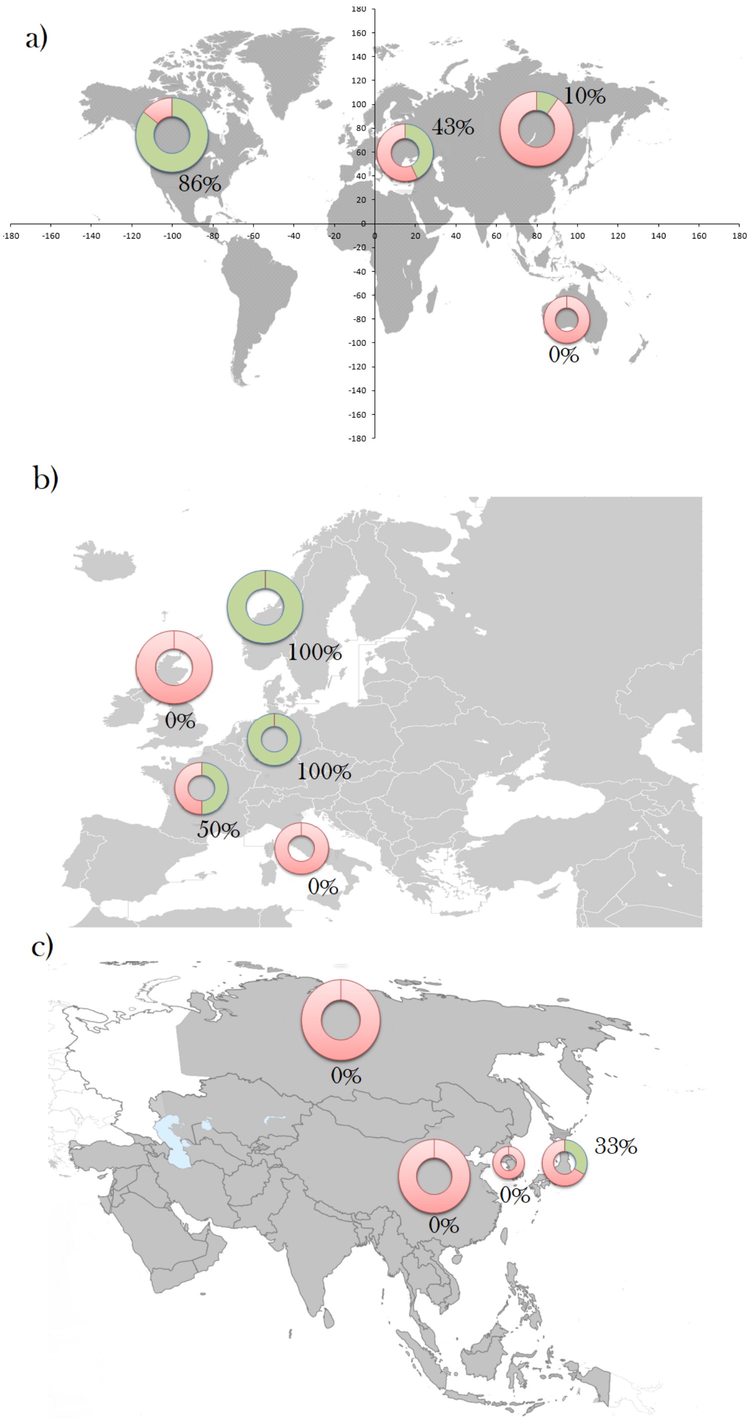
Rodríguez et al., 2019), and sharing it could help to alleviate this situation. We suggest the possibility of making available 'frozen' versions of the codes used for research studies in open repositories. Similarly frozen versions of the climate models later used to support the results discussed in international reports on climate change should be made accessible along with the outputs from simulations in official data portals. Moreover, as shown for other fields of software development, this could help to improve the development process of climate models and how these may be expected to work for the scientific research community (Boulanger, 2005). It is clear that funding should be allocated by agencies and relevant bodies to support such efforts, notwithstanding that the whole framework of science faces new challenges but at the same time presents opportunities for improvement in such a sensitive field as climate science.

*Code and data availability.* There is no code or data relevant to this paper.

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019


Geoscientific Model Development
Discussions
Open Access

**Figure 1.** Geographical map with percentage of the models obtained for each country: a) worldwide; b) Europe (EC-Earth is only included in the worldwide view because it is developed as a consortium of sixteen european countries); c) Asia. Green colours and percentages represent the obtained models.



https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019

Geoscientific Model Development Discussions
Open Access
EGU

Table 1: CMIP5 model list, research centre responsible for each one and details on the procedure for accessing their code.

| Modeling center | Model | Free download | Answer Email 1 | Answer Email 2 | Comments/Answer |
|---|---|---|---|---|---|
| BCC | BCC-CSM1.1 | No | - | - | No email or contact phone is available. |
| | BCC-CSM1.1(m) | No | | | |
| CCCma | CanAM4 | No | Yes | | The code is not shared. |
| | CanCM4 | No | | | |
| | CanESM2 | No | | | |
| CMCC | CMCC-CESM | No | No | No | No answer. |
| | CMCC-CM | No | | | |
| | CMCC-CMS | No | | | |
| CNRM-CERFACS | CNRM-CM5 | No | No | Yes | The code is not shared. |
| | CNRM-CM5-2 | No | | | |
| COLA and NCEP | CFSv2-2011 | Yes | - | - | Code available from the official web site. |
| CSIRO-BOM | ACCESS1.0 | No | Yes | | The code is not shared. |
| | ACCESS1.3 | No | | | |
| CSIRO-QCCCE | CSIRO-Mk3.6.0 | No | - | - | No email or contact phone is available. |
| EC-EARTH | EC-EARTH | No | - | - | The code is not shared. |
| FIO | FIO-ESM | No | No | No | No answer. |
| GCESS | BNU-ESM | No | No | No | No answer. |
| INM | INM-CM4 | No | - | - | No email or contact phone is available. |
| IPSL | IPSL-CM5A-LR | Yes | Yes | | Available after email exchange. |
| | IPSL-CM5A-MR | Yes | | | |
| | IPSL-CM5B-LR | Yes | | | |
| LASG-CESS | FGOALS-g2 | No | No | No | No answer. |
| LASG-IAP | FGOALS-gl | No | - | - | No email or contact phone is available. |
| | FGOALS-s2 | No | | | |
| MIROC | MIROC4h | No | Yes | | The code is not shared. |
| | MIROC5 | No | | | |
| | MIROC-ESM | No | | | |
| | MIROC-ESM-CHEM | No | | | |
| MOHC | HadCM3 | No | No | Yes | The code is not shared. |
| | HadCM3Q | No | | | |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019



| | | | | | |
|---|---|---|---|---|---|
| | HadGEM2-A | No | | | |
| | HadGEM2-CC | No | | | |
| | HadGEM2-ES | No | | | |
| MPI-M | MPI-ESM-LR | Yes | Yes | - | Available after email exchange. |
| | MPI-ESM-MR | Yes | | | |
| | MPI-ESM-P | Yes | | | |
| MRI | MRI-AGCM3.2H | No | - | - | No email or contact phone is available. |
| | MRI-AGCM3.2S | No | | | |
| | MRI-ESM1 | No | | | |
| | MRI-CGCM3 | No | | | |
| NASA-GISS | GISS-E2-H | Yes | No | No | Available after email exchange. |
| | GISS-E2-H-CC | Yes | | | |
| | GISS-E2-R | Yes | | | |
| | GISS-E2-R-CC | Yes | | | |
| NASA-GMAO | GEOS-5 | Yes | No | No | Available after meeting during a workshop. |
| NCAR | CCSM4 | Yes | - | - | Code available from the official web site. |
| NCC | NorESM1-M | Yes | Yes | | Available after email exchange. |
| | NorESM1-ME | Yes | | | |
| NICAM | NICAM.09 | No | No | Yes | Available after email exchange. |
| NIMR/KMA | HadGEM2-AO | No | No | Yes | The code is not shared. |
| NOAA-GFDL | GFDL-CM2.1 | Yes | Yes | Yes | Available after email exchange. |
| | GFDL-CM3 | Yes | | | |
| | GFDL-ESM2G | Yes | | | |
| | GFDL-ESM2M | Yes | | | |
| | GFDL-HIRAM-C180 | Yes | | | |
| | GFDL-HIRAM-C360 | Yes | | | |
| NSF-DOE-NCAR | CESM1(BGC) | Yes | - | - | Code available from the official web site. |
| | CESM1(CAM5) | Yes | | | |
| | CESM1(CAM5.1,FV2) | Yes | | | |
| | CESM1(FASTCHEM) | Yes | | | |
| | CESM1(WACCM) | Yes | | | |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019

Geoscientific Model Development Discussions
Open Access
EGU

Table 2: Summary of reasons behind granting us access to the source code of the models.

| Modeling center | Model | Process and reasons to access to the code |
|---|---|---|
| COLA and NCEP | CFSv2-2011 | A tarball with the source code can be easily accessed from the official web site explaining what the code does and how the climate model works. |
| IPSL | IPSL-CM5A-LR<br>IPSL-CM5A-MR<br>IPSL-CM5B-LR | M. García-Rodríguez identified himself and explained via email the purposes of this research. After a meeting of the developing team and additonal details on this research we were granted access to a tarball with the source code. |
| MPI-M | MPI-ESM-LR<br>MPI-ESM-MR<br>MPI-ESM-P | The access to a tarball with the source code was granted after registration as an user via a web page and approval, without any extra communication or reasoning. |
| NASA-GISS | GISS-E2-H<br>GISS-E2-H-CC<br>GISS-E2-R<br>GISS-E2-R-CC | After two weeks, we received the answer to our email. They have provided us directly with a link in which a tarball with the source code can be accessed with the snapshots of the model. |
| NASA-GMAO | GEOS-5 | Initially, they did not answer the emails that were sent to them. However, after Dr. Añel's contact with one of the team members during a workshop, we were put in contact with one of the coders. By contacting this person we obtained access, available as 4073 files in directories retrieved using 'wget'. |
| NCAR | CCSM4 | The code of the model is available through a web page. The dowload proccess is open to anyone but it is hard. Each file of the model has to be individually retrieved (2247 files in total, each in its respective sub-directory). |
| NCC | NorESM1-M<br>NorESM1-ME | First, we received a reply stating that the code of the model is not shared with anyone outside the NorESM-community, asking if we really needed it. After identifying ourselves and explaining our research, we were granted access to a tarball after registering as users in the 'noresm wiki'. |
| NICAM | NICAM.09 | Initially, they asked us questions about the purpose of achieving the code. Then, explaining the objectives of the project, they have given us access to a tarball with the code after registering in the nicam user group. |
| NOAA-GFDL | GFDL-CM2.1<br>GFDL-CM3<br>GFDL-ESM2G<br>GFDL-ESM2M<br>GFDL-HIRAM-C180 | We were granted access to a tarball with the source code in reply to our first request via email. |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019


Geoscientific Model Development
Discussions
Open Access

| | | |
|---|---|---|
| | GFDL-HIRAM-C360 | |
| NSF-DOE-NCAR | CESM1(BGC) | We had to register to access the Community Earth System Model. After that, we were able to download a tarball with the source code. |
| | CESM1(CAM5) | |
| | CESM1(CAM5.1,FVV2) | |
| | CESM1(FASTCHEM) | |
| | CESM1(WACCM) | |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019



Table 3: CMIP5 models with code obtained and scores of reproducibility. Maximum value is given to those models it is possible to access through the internet without restriction, with a license that allows full testing and evaluation of the model. The score was reduced by one star when failing for each one of the following criteria: if in order to gain access to the model we had to contact a research centre or development group, to sign license agreements, or if we gained access only after identifying ourselves as scientists undertaking climate research and according to the rights to evaluate and use the model as granted by the license (if applicable). A not-filled star means that the license of the model does not allow modification of the code.

| **Institution** | **Model** | **Score** |
|---|---|---|
| Cola & NCEP | CFSv-2011 | ★★☆ |
| IPSL | IPSL-CM5A-LR<br>IPSL-CM5A-MR<br>IPSL-CM5B-LR | ★★ |
| MPI-M | MPI-ESM-LR<br>MPI-ESM-MR<br>MPI-ESM-P | ★ |
| NASA GISS | GISS-E2-H<br>GISS-E2-H-CC<br>GISS-E2-R<br>GISS-E2-R-CC | ★☆ |
| NASA GMAO | GEOS-5 | ★☆ |
| NCAR | CCSM4 | ★★★ |
| NCC | NorESM1-M<br>NorESM1-ME | ★★ |
| NICAM | NICAM.09 | ★ |
| NOAA GFDL | GFDL-CM2.1<br>GFDL-CM3<br>GFDL-ESM2G<br>GFDL-ESM2M<br>GFDL-HIRAM-C180<br>GFDL-HIRAM-C360 | ★★ |
| NSF-DOE-NCAR | CESM1(BGC)<br>CESM1(CAM5)<br>CESM1(CAM5.1 FV2)<br>CESM1(FASTCHEM)<br>CESM1(WACCM) | ★★★ |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019



Table 4: Availability of detailed information provided with the source code of the models in order to run them. 'Documentation' refers to full documentation of the model (for IPSL models a web address/link was included to access the documentation but it did not work). 'Readme' corresponds to a file containing basic explanations on the files part of the model and basic instructions. 'Basic example' refers to whether an example to explain the model is included. 'Dependencies' refers to the basic information on libraries, compilers or any other software and its version needed to run the model. 'Makefile' refers to the existence of a single file that manages all the process of compilation and model run.

| **Modeling center** | **Model** | **Documentation** | **ReadMe** | **Basic example** | **Depencencies listed** | **Makefile** |
|---|---|---|---|---|---|---|
| COLA and NCEP | CFSv2-2011 | yes | yes | yes | yes | yes |
| IPSL | IPSL-CM5A-LR<br>IPSL-CM5A-MR<br>IPSL-CM5B-LR | no* | yes | yes | yes | yes |
| MPI-M | MPI-ESM-LR<br>MPI-ESM-MR<br>MPI-ESM-P | yes | yes | yes | yes | yes |
| NASA-GISS | GISS-E2-H<br>GISS-E2-H-CC<br>GISS-E2-R<br>GISS-E2-R-CC | yes | yes | yes | yes | yes |
| NASA-GMAO | GEOS-5 | yes | yes | yes | no | yes |
| NCAR | CCSM4 | yes | yes | no | yes | yes |
| NCC | NorESM1-M<br>NorESM1-ME | yes | yes | yes | yes | yes |
| NICAM | NICAM.09 | no | yes | no | no | yes |
| NOAA-GFDL | GFDL-CM2.1<br>GFDL-CM3<br>GFDL-ESM2G<br>GFDL-ESM2M<br>GFDL-HIRAM-C180<br>GFDL-HIRAM-C360 | yes | yes | yes | yes | yes |
| NSF-DOE-NCAR | CESM1(BGC)<br>CESM1(CAM5)<br>CESM1(CAM5.1,FVV2)<br>CESM1(FASTCHEM) | yes | yes | yes | yes | yes |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019

Geoscientific Model Development
Discussions

| | | | | | | |
|---|---|---|---|---|---|---|
| | CESM1(WACCM) | | | | | |

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019

Geoscientific Model Development Discussions
Open Access EGU

## Appendix A: Templates of emails used to contact the model development teams

### A1 First email

Dear Sir/Madame,

my name is Michael García Rodríguez and I am an MSc Student at the EPhysLab in the Universidade de Vigo, Spain (http://ephyslab.uvigo.es). I am developing my MSc Thesis on the study of qualitative issues of climate models, mostly related to scientific reproducibility and copyright issues.

In order to do it, I have focused my research project on the study of the models that contributed to the last CMIP5 report. For it, I am trying to get access to the code of all the models that reported results of this effort.

Therefore I would like kindly request access to the code of your model, $MODEL-NAME$, namely the version that you used to produce CMIP5 results. Therefore, could you say me how could I get access to it?

Many thanks in advance.

Best regards,

Michael García Rodríguez

EPhysLab

Universidade de Vigo

http://ephyslab.uvigo.es

========================

### A2 Second email

The second email was equal to the first one, but a second try:

Dear Sir/Madame,

my name is Michael García Rodríguez and I am an MSc Student at the EPhysLab in the Universidade de Vigo, Spain (http://ephyslab.uvigo.es). I am developing my MSc Thesis on the study of qualitative issues of climate models, mostly related to scientific reproducibility and copyright issues.

In order to do it, I have focused my research project on the study of the models that contributed to the last CMIP5 report. For it, I am trying to get access to the code of all the models that reported results of this effort.

Therefore I would like kindly request access to the code of your model, $MODEL-NAME$, namely the version that you used to produce CMIP5 results. Therefore, could you say me how could I get access to it?

Many thanks in advance.

Best regards,

Michael García Rodríguez

EPhysLab

Universidade de Vigo

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019

Geoscientific Model Development Discussions

http://ephyslab.uvigo.es
======================

### A3 Third email

Dear Sir/Madame,

We are not trying to understand the code but some way to make qualitative measurements of the code most the focus of copy-right use and scientific reproducibility, focusing on how easy it is to get access to the code of the models. Size and complexity of the code is not a problem from my point of view. What you describe is similar to how other models are but for example, the CESM team has provided a tarball file maybe you can do something similar. We are not going to try to understand the physics of the code. If the amount of work is so great that in fact, you can not to deal with it, could you explain me why? It would be of great help, in case of not being able to get the code of the model, know the answer. Please, if it's possible, mark with a cross one or more answers on below:

[ ] Because of restrictions imposed by the institution/s where the model is developed
[ ] Copyright issues (please, if you mark this choice, could you send me a copy of the licenses?)
[ ] Development team policy
[ ] Legal restrictions of your country
[ ] Others reasons (please specify):
——————————————————————————-

In this case, I will be able to write down the reasons why I was not allowed access to the code and I could document it in my MSc Thesis on the study of qualitative issues of climate models.

thank you for your time, I really appreciate it.
Best regards,
Michael García Rodríguez
EPhysLab
Universidade de Vigo
http://ephyslab.uvigo.es
======================

*Author contributions.* All the authors participated in the design of the study and writting of the text. MGR and JAA made the attemps to get access to the code of climate models.

*Competing interests.* We do not have competing interest.

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019


Geoscientific Model Development Discussions
Open Access

*Acknowledgements.* This research was partially supported by the ZEXMOD Project of the Goverment of Spain (CGL2015-71575-P) and the European Regional Development Fund (ERDF). Juan A. Añel is supported by a 'Ramón y Cajal' Grant funded by the Government of Spain (RYC-2013-14560). We would like to thank to Didier Roche, Julia Hargreaves and Rolf Sander for useful comments to improve this paper.

https://doi.org/10.5194/gmd-2019-191
Preprint. Discussion started: 3 September 2019




## References

Allison, D., Shiffrin, R., and Stodden, V.: Reproducibility of research: Issues and proposed remedies, Proc. Nat. Ac. Sci., 115, 2561–2562, https://doi.org/10.1073/pnas.1802324115, 2018.

Añel, J. A.: The importance of reviewing the code, Communications of the ACM, 54, 40–41, https://doi.org/10.1145/1941487.1941502, 2011.

Añel, J. A.: Comment on 'Most computational hydrology is not reproducible, so is it really science?' by Hutton et al. (in press), Water Resour. Res., 2017.

Barnes, N.: Publish your computer code: it is good enough, Nature, 467, 753, https://doi.org/10.1038/467753a, 2010.

Boulanger, A.: Open-source versus proprietary software: Is one more reliable and secure than other?, IBM Syst. J., 44, 239–248, https://doi.org/10.1147/sj.442.0239, 2005.

Easterbrook, S. M.: https://www.easterbrook.ca/steve/2009/06/getting-the-source-code-for-climate-models, 2009.

Easterbrook, S. M.: Open code for open science?, Nat. Geosci., 7, 779–781, https://doi.org/10.1038/ngeo2283, 2014.

Easterbrook, S. M. and Johns, T.: Engineering the Software for Understanding Climate Change, Comput. Sci. Eng., 11, 65–74, https://doi.org/10.1109/MCSE.2009.193, 2009.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Mod. Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fanelli, D.: Opinion: Is science really facing a reproducibility crisis, and do we need it to?, Proc. Nat. Ac. Sci., 115, 2628–2631, https://doi.org/10.1073/pnas.1708272114, 2018.

García-Rodríguez, M., Añel, J. A., Foujols, M.-A., and Rodeiro, J.: FortranAnalyser: a software tool to assess Fortran code quality, submitted to SoftwareX, 2019.

GMD Executive Editors: Editorial: The publication of geoscientific model developments v1.1, Geosci. Mod. Dev., 8, 3487–3495, https://doi.org/10.5194/gmd-8-3487-2015, 2015.

GMD Executive Editors: Editorial: The publication of geoscientific model developments v1.2, Geosci. Mod. Dev., 12, 2215–2225, https://doi.org/10.5194/gmd-12-2215-2019, 2019.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Joppa, L. N., McInerny, G., Harper, R., Salido, L., Takeda, K., O'Hara, K., Gavaghan, D., and Emmot, S.: Troubling Trends in Scientific Software Use, Science, 340, 814–815, https://doi.org/10.1126/science.1231535, 2013.

Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: Generation CMIP5 and how we got there, Geophys. Res. Lett., 40, 1194–1199, https://doi.org/10.1002/grl.50256, 2013.

Lee, B. D.: Ten simple rules for documenting scientific software, PLoS Comput. Biol., 14, e1006 561, https://doi.org/10.1371/journal.pcbi.1006561, 2018.

LeVeque, R. J.: Top Ten Reasons To Not Share Your Code (and why you should anyway), SIAM News, 46, 7–8, 2013.

Massonnet, F., Ménégoz, M., Acosta, M., Yepes-Arbós, X., Exarchou, E., and Doblas-Reyes, F. J.: Replicability of the EC-Earth3 Earth System Model under a change in computing environment (in review), Geosci. Mod. Dev. Discuss., https://doi.org/10.5194/gmd-2019-91, 2019.

Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marécal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K., and Zeng, G.: Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI) , Geosci. Mod. Dev., 10, 639–671, https://doi.org/10.5194/gmd-10-639-2017, 2017.

Nature: Does your code stand up to scrutiny?, Nature, 555, https://doi.org/10.1038/d41586-018-02741-4, 2018.

Pipitone, J. and Easterbrook, S.: Assessing climate model software quality: a defect density analysis of three models, Geosci. Mod. Dev., 5, 1009–1022, https://doi.org/10.5194/gmd-5-1009-2012, 2012.

RealClimate.org: http://www.realclimate.org/index.php/data-sources, 2009.

Stodden, V., Guo, P., and Ma, Z.: Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals, PLoS ONE, 8, e67 111, https://doi.org/10.1371/journal.pone.0067111, 2013.

Stodden, V., Seiler, J., and Ma, Z.: An empirical analysis of journal policy effectiveness for computational reproducibility, Proc. Nat. Ac. Sci., 115, 2584–2589, https://doi.org/10.1073/pnas.1708290115, 2018.

Taylor, K., Stouffer, R., and Meehl, G.: An Overview of CMIP5 and the Experiment Design, Bull. Amer. Meteorol. Soc., pp. 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., and Teal, T.: Good enough practices in scientific computing, PLoS Comput. Biol., 13, e1005 510, https://doi.org/10.1371/journal.pcbi.1005510, 2017.