Getting Started Tutorial for Users Seeking CMIP5 Model Output
(revised 3 March 2011)

1. **Access to data.**

   Anyone can browse the CMIP5 model output archive catalogue, but in order to download data you will be required to register (i.e. open an account) (see step 2 below). Your account will be valid for all data (CMIP5 plus much, much more) served by the Earth System Grid Federation (ESGF). Because different restrictions are placed on different datasets served by ESGF, you must also enroll in one of the two "CMIP5 groups" and agree to the terms of use established for CMIP5 (see step 3 below).

2. **Obtaining an ESGF account.**

   a. All CMIP5 Model output can be accessed through any one of several Gateway websites operated by ESGF. Although nearly a dozen gateways currently exist, the gateways with primary responsibility for CMIP5 are:
      PCMDI   http://pcmdi3.llnl.gov
      BADC   http://cmip-gw.badc.rl.ac.uk
      DKRZ   http://ipcc-ar5.dkrz.de
      The other gateways are listed on the home page of any of the above sites (in a box at the right of the page).
   b. Who should register?
      Users who previously registered at PCMDI for the purpose of obtaining CMIP3 output will not be able to access CMIP5 output under that CMIP3 account. Those users plus new users who don't currently have an account at any of the ESGF gateways will need to register.
   c. To open an account, you must first select a "home gateway." The choice is not critical, but if your interest is primarily in CMIP5 output, there may be some advantages in selecting a data center (located near you) that has plans to host a substantial fraction of the total CMIP5 output (currently, PCMDI, BADC, DKRZ). Once registered, you are free to access data through any of the gateways in the federation.
   d. Gateways have registration procedures that are not all identical but are all quite straightforward. The following registration instructions apply specifically to the PCMDI Gateway.
      i. If you choose PCMDI as your home gateway, go to http://pcmdi3.llnl.gov and click on the "Login" tab at the top of the page.
      ii. Click on "Create new Account" and provide the information requested.
      iii. On completing your input, an email will immediately be sent to you, which will include a link (URL) that you must click on to complete your registration. Note that registration will enable you

to login to the ESG, but (as described in step 3 below) you'll have to enroll in one of the two CMIP5 user groups before you'll be able to download data.

iv. The email sent to you will also contain your permanent OpenID, which you should record for future reference. You will need your OpenID when you download any data not served by your home gateway.

3. **Enroll in the appropriate CMIP5 group.**

a. In order to enroll in a group that has permission to access CMIP5 model output, you must apply at the PCMDI ESGF Gateway (since PCMDI controls access to CMIP5 output).

b. Go to http://pcmdi3.llnl.gov and click on the "Login" tab at the top of the page.

c. If PCMDI is your home Gateway, log in with your username and password; otherwise enter your OpenID in the "External OpenID Login" box provided. When you press "Go", your browser will be redirected to your home gateway, and you will need to enter your password (and for some sites your username). Once you have done so, your browser will be redirected back to PCMDI.

d. The next step is to determine which of the two CMIP5 user groups you should belong to: the "research" group or the "commercial" group. If you subscribe to the "research" group, you will have access to all the CMIP5 data, but you may use it only for non-commercial educational and research purposes, as spelled out in the terms of use. If you plan to use the data for other purposes, you should join the "commercial" group, and then you will be given access to output from a subset (about two-thirds) of the models, but your use of it is unrestricted. The terms of use governing the CMIP5 data are spelled out at:
   http://cmip-pcmdi.llnl.gov/cmip5/terms.html?submenuheader=2.

e. To declare your intended use of the CMIP5 output, click on the "Account" tab at the top of the web page. Click on "Apply for Group Membership". There are a number of ESG member groups that will be listed. Click on either "CMIP5 Research" or "CMIP5 Commercial", depending on your intent (as described in 3d above), and then click on "subscribe".

f. You next will be asked to read and agree to the terms of use. If you agree, *be sure to click on "I accept" before proceeding.* You also must enter a brief statement of work. The kinds of questions you might wish to answer in providing your "statement of work" include:
   - Will you be studying a particular phenomenon (e.g., ENSO, MJO, dry spells)?
   - Will you be assessing model skill (comparing with observations) or will you be using the output to study impacts or drive other models?
   - Will you focus on a specific region?

- Will you focus on particular experiments?
g. After entering your statement of work, *double check that you have clicked on "I accept" before clicking on "subscribe"; otherwise your "statement of work" will have to be re-entered*. Once you have completed the subscription procedure, you will be authorized to download the CMIP5 data.

4. **Find out what data are available.**

   a. On any ESG gateway, select either the "Home" tab or "Data" tab at the top of the page, and then in the boxed panel near the top of the page elect one of the following "Search" options:
      i. "Datasets" which provides information about datasets that can be downloaded, or
      ii. "Simulations" which provides documentation of the models and their simulations.
   b. You can refine your search by entering text in the box near the top of the page and/or by selecting items from the pull-down menus under "Search Categories" in the panel on the left side of the page. If you find "Project" listed there, you should initially limit your search to "CMIP5", otherwise your search and selection procedure may yield many datasets that you really have no interest in or have no permission to download.
   c. The results of your search depend on which search option ("Datasets" or "Simulations") you have selected.
      i. If you selected "Datasets" the result will be a list of "datasets". Each "dataset" comprises a subset of the output saved from a single model run. The subset shares a common data sampling frequency (e.g., monthly, daily, 3-hourly) and a common modeling realm (e.g., atmosphere, ocean, land). The list of variables found (encompassing the union of all your search results) can be seen by clicking on the "Variable" drop-down menu under "Search Categories". The variables included in a particular "dataset" can be seen by clicking on the dataset and then clicking on the "Variables" tab.
      ii. If you selected "Simulations", the result of your search will be a list of model runs. If you click on a run, you will obtain information about the model and the simulation conditions.
   d. If you are only interested in a single variable (e.g., "eastward wind"), then before proceeding it is best to select it from the "variable" pull-down menu in the panel on the left. This will usually reduce the number of point and clicks subsequently required to download the data.

5. **Download model output.**

   a. From the list of datasets returned by your search (completed in step 4 above), you should select the ones of interest by clicking on the Data

Center from which you want to download the data. [Currently only a single center will be listed for most datasets; Once the data have been replicated (in the coming months, as planned), you'll then be presented with a choice of data centers.]

b. You can now click on "Download Files" (in the panel at the right of the page).

    i. If you have included any files you are not authorized to download, your request will not be granted. Instead you'll learn how you might gain authorization by joining a new "users group". If you are a member of the CMIP5 "commercial" group, and if the message suggests that you join the CMIP5 "research" group, you should decline. Instead, you will have to eliminate from your download request any model datasets you are not authorized to use. If you do not know which models are only available for non-commercial educational and research purposes (i.e., the ones you must eliminate), see the table at the end of this document. Note that in the future there will be easier ways of sub-selecting only those files that you are authorized to download.

    ii. Similarly, if you attempt to access data from a model that has not yet been released for public analysis, you will be given an option to request permission to access that data (i.e., subscribe to a special group with access privileges). You should **not** attempt to subscribe, since your request will not be granted.

    iii. If the datasets you have requested are hosted by a different gateway, you may be asked to enter your OpenID, which you should do.

c. When you get to the download page, you will see all the files comprising the dataset (but if you followed the recommended search procedure, this list only includes the variable you are interested in). There are two options for downloading files:

Option 1: Downloading files one at a time.

    You can download *individual* files one at a time by clicking on the blue "download" links in the far right column of the table. The data you wish to download will often be hosted at a Data Node located elsewhere (i.e., not at the site where you have initiated your search). When you select the link for a file located at a remote Data Node, you may be asked to enter your OpenID, which you should do.

Option 2: Downloading several files using a script.

    If you have a Linux, Unix, or Mac OSX platform, you can download *all* the listed files by checking the small box on the left in the first line at the top of the table(s), or you can select a subset of the files by clicking on the boxes in the left column of the table(s). After selecting the files you want, click on the "Download All Selected Files" button at the top right of the window. You'll then be shown the number of files and the volume of data you have requested, and one or more "Download" arrows, one for each script needed to get

all the data you want. When you click on each "Download", you'll download a Wget script. In the usual case, scripts will *not* execute until you have obtained an appropriate certificate, as described in the next section. [If you need a certificate, the web page will indicate that you should obtain a certificate.] When you are ready to run your scripts, open a terminal window, check that the script files are designated as executable (if not, "chmod +x <filename>"), and run your scripts (e.g., "./wget-download.sh").

6. **Obtaining a certificate.**

To use wget scripts (described in the previous section) to download your selected files, you will usually need a certificate that attests to your right to access the data. The certificate will expire after a set period (currently 12 hours). If you are unable to download your data within this period, you will need to acquire a new certificate.

During this beta-testing phase, you may encounter difficulties in successfully obtaining a certificate. One requirement is that a recent version of Java should be installed on your system (version 1.5 or higher).   Please report other problems to the help desk listed under item 7 below.

a. The download webpage will indicate whether or not you need to obtain a MyProxy certificate, which you can usually get via your browser. Click on the "Launch MyProxyLogon" link (on the page that also contains the link for downloading the wget script). [You must be patient here. At some point you might be asked to reenter your OpenID and/or password. You may be asked to "allow" the application to be launched, and it may take some time for webstart to begin to execute.] Eventually, a window should display a text box with your username, the Hostname, and Port. You should enter your "Passphrase" (i.e., your ESG password), and then click on "Logon".   [This is a java application that may not work with all browsers -- if it does not open, skip to step b below. If you get as far as clicking on "download" and then nothing happens after at least a minute or two, then you may have a firewall that is preventing access. You should consult your systems administrator.] The MyProxyLogon application will assign and store a certificate in a standard location on your computer where your wget script can subsequently access it. Note that the certificate will expire after 12 hours, but it will be good for all wget scripts you run until that time.

b. If the *MyProxyLogon* cannot be launched as described above, you should try the following. First you should check that you have Java (version 1.5 or higher) correctly installed on your computer. Next you should check that your browser settings are set to launch Java *WebStart* applications. If you are unsuccessful in launching MyProxy from your browser, there are other alternatives:

i. The MyProxyLogon program can be downloaded to your computer and run as a command line program. Click on the link to download the program and follow the instructions to run it.

ii. Install MyProxyClient. MyProxyClient is a script written in the Python programming language. Follow the instructions to install this program and run it.

7. **Reporting bugs and seeking help.**

To report bugs, make suggestions, air complaints, or ask for help, please send an email to: esg-support@earthsystemgrid.org .

The following known bug is being worked on:

Sometimes (when you need to enter your OpendID), you'll arrive at a webpage that is blank except for the words "OpenID Login" (plus a header and footer). If you encounter this problem, you should clear all cookies in your browser and start over. Sometimes simply closing and reopening your browser   also   works.

Terms of use declared by each modeling center

<mark>Highlighted models</mark> are available for unrestricted use. The remaining models may only be used for non-commercial research and educational purposes. Gray shaded models are undeclared at this time. [See complete "Terms of Use": http://cmip-pcmdi.llnl.gov/cmip5/terms.html?submenuheader=2.]

name of modeling center (or group)	institute_id	model_id
The Centre for Australian Weather and Climate Research	CAWCR	ACCESS
Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1
College of Global Change and Earth System Science, Beijing Normal University	GCESS	BNU-ESM
Canadian Centre for Climate Modelling and Analysis	CCCMA	CanESM2 CanCM4 CanAM4
University of Miami - RSMAS	RSMAS	CCSM4(RSMAS)
Community Earth System Model	CESM1	CESM1(BGC) CESM1(CAM5) CESM1(CHEM,CAM5) CESM1(CHEM) CESM1(WACCM)
Centro Euro-Mediterraneo per I Cambiamenti Climatici	CMCC	CMCC-CESM CMCC-CM CMCC-CMS
Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	CNRM-CERFACS	CNRM-CM5
Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence	CSIRO/QCCCE	CSIRO-Mk3.6
EC-EARTH consortium	EC-EARTH	EC-EARTH
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences	undeclared	FGOALS-G2.0 FGOALS-gl FGOALS-S2.0
The First Institute of Oceanography, SOA, China	FIO	FIO-ESM

Global Modeling and Assimilation Office	NASA/GMAO	GEOS-5
Geophysical Fluid Dynamics Laboratory	GFDL	GFDL-CM2.1 GFDL-CM3 GFDL-ESM2G GFDL-ESM2M GFDL-HIRAM-C180 GFDL-HIRAM-C360
NASA Goddard Institute for Space Studies	GISS	GISS-E2-H GISS-E2-H-CC GISS-E2-R GISS-E2-R-CC GISS-E2CS-H GISS-E2CS-R
National Institute of Meteorological Research/Korea Meteorological Administration	NMR/KMA	HadGEM2-AO
Met Office Hadley Centre	MOHC	HadCM3 HadCM3Q HadGEM2-CC HadGEM2-ES
Natural and Environmental Research Council/Met Office Hadley Centre	undeclared	HiGEM1.2
Institute for Numerical Mathematics	INM	INM-CM4
Institut Pierre-Simon Laplace	IPSL	IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC	MIROC-ESM MIROC-ESM-CHEM
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC4h MIROC4m MIROC5
Max Planck Institute for Meteorology	MPI-M	MPI-ESM-HR MPI-ESM-LR
Meteorological Research Institute	MRI	MRI-AGCM3.2H MRI-AGCM3.2S MRI-CGCM3 MRI-ESM1
Norwegian Climate Centre	NCC	NorESM