

# The Bumper Book of *.cookies*

(The cGENIE.cookie user-manual and  
introduction to Earth system modelling)

Andy Ridgwell

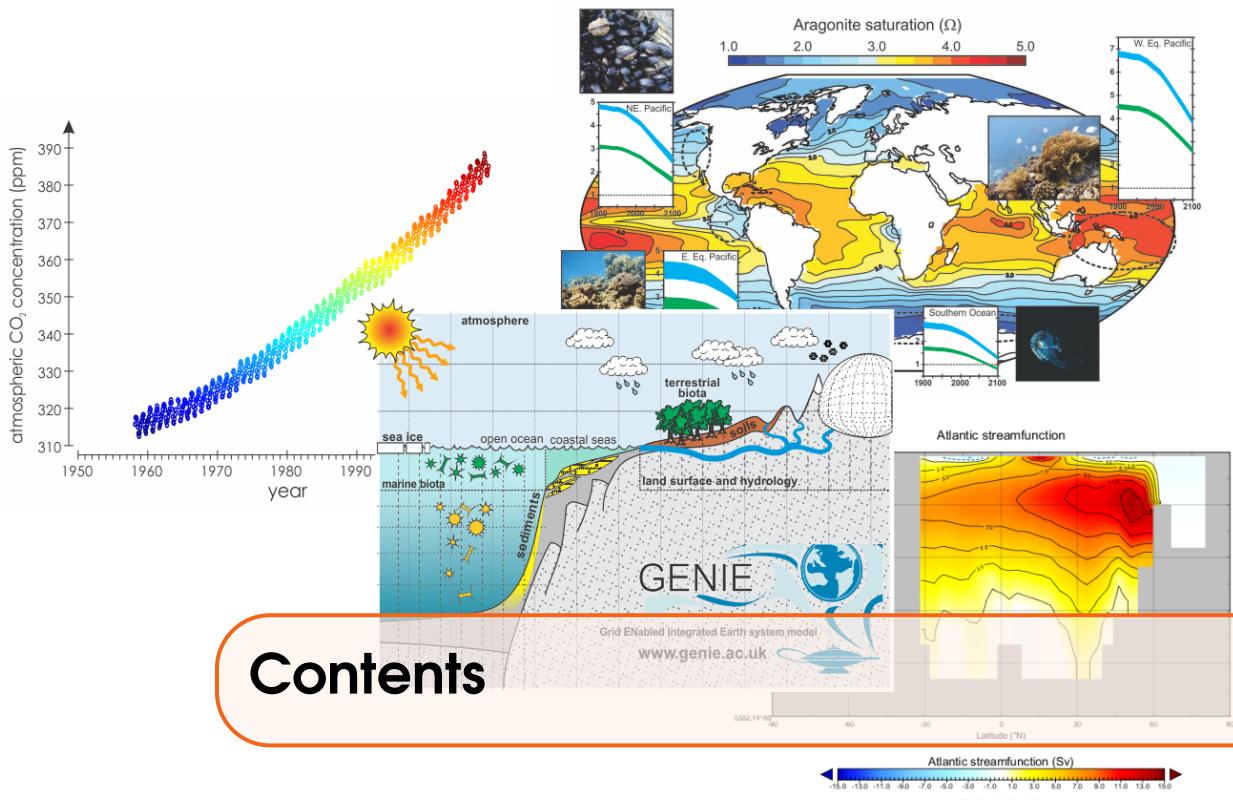
Copyright © 2017 Andy Ridgwell

PUBLISHED BY DERPY-COOKIES INC.

MYCOOKIE.SEAO2.ORG

Licensed under the Creative Commons Attribution-NonCommercial 3.0 Unported License (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <http://creativecommons.org/licenses/by-nc/3.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

*First printing, October 2017*



## Contents

<b>0</b>	<b>Installation, Configuration, Basic Usage</b>	<b>1</b>
<b>0.1</b>	<b>Before anything else ...</b>	<b>2</b>
<b>0.2</b>	<b>Starting (dozing?) off ...</b>	<b>6</b>
0.2.1	Logging in!	6
0.2.2	Downloading/installing the model code	7
0.2.3	Configuring the code	7
0.2.4	Testing the model code	9
<b>0.3</b>	<b>Running the model</b>	<b>12</b>
<b>0.4</b>	<b>Creating new experiments!</b>	<b>16</b>
<b>0.5</b>	<b>Model output</b>	<b>17</b>
0.5.1	Time-slice output	17
0.5.2	Time-series output	18
0.5.3	File naming convention	18
<b>0.6</b>	<b>Viewing model output</b>	<b>19</b>
0.6.1	Time-series output	19
0.6.2	2- and 3-D time-slice output	20
<b>0.7</b>	<b>Submitting experiment ‘jobs’</b>	<b>21</b>
<b>0.8</b>	<b>‘Restarts’</b>	<b>23</b>
<b>1</b>	<b>Climate dynamics &amp; experimental design</b>	<b>25</b>
1.1	Brrrrrrrrrrr – it’s chilly on ... snowball Earth!	27

<b>1.2</b>	<b>Further ideas</b>	<b>31</b>
1.2.1	Feedback loop analysis . . . . .	31
1.2.2	Continental configuration (1) . . . . .	31
1.2.3	Continental configuration (2) . . . . .	32
1.2.4	Geothermal heat input . . . . .	32
1.2.5	Seasonality . . . . .	33
<b>2</b>	<b>Ocean circulation I – modern</b>	<b>35</b>
2.1	<b>Simulating ocean circulation</b>	<b>37</b>
2.2	<b>Visualizing ocean circulation</b>	<b>39</b>
2.3	<b>Tracing ocean circulation</b>	<b>41</b>
2.4	<b>Tracing ocean ... ventilation</b>	<b>44</b>
2.5	<b>Poking the climate beast (1)</b>	<b>45</b>
2.6	<b>Poking the climate beast (2)</b>	<b>48</b>
2.7	<b>Further ideas</b>	<b>49</b>
2.7.1	Assessing the spatial sensitivity of deep-water formation to shutdown . . . . .	49
2.7.2	Forcing stronger deep-water formation ('anti-hosing' investigations) . . . . .	49
2.7.3	Regional hosing – Poking the Mediterranean . . . . .	49
<b>3</b>	<b>Ocean circulation II (fake worlds)</b>	<b>51</b>
3.1	<b>(none)</b>	<b>52</b>
<b>4</b>	<b>Atmospheric dynamics</b>	<b>53</b>
4.1	<b>(none)</b>	<b>54</b>
<b>5</b>	<b>Fossil fuel CO<sub>2</sub> and 'ocean acidification'</b>	<b>55</b>
5.1	<b>Exploring the consequences of fossil fuel CO<sub>2</sub> emissions</b>	<b>57</b>
5.1.1	Idealized emissions forcing . . . . .	58
5.1.2	Where has my carbon gone??? . . . . .	61
5.1.3	Historical (real-world!) emissions forcing . . . . .	62
5.1.4	Assessing future carbon emissions impacts . . . . .	64
5.2	<b>Further ideas</b>	<b>65</b>
5.2.1	Assessing the importance of emissions rate . . . . .	65
5.2.2	Determining thresholds of environmental impact . . . . .	65
5.2.3	Future atmospheric CO <sub>2</sub> concentration pathways ('RCPs') . . . . .	68
5.2.4	Isotopic tracing of fossil fuel CO <sub>2</sub> uptake . . . . .	69
5.3	<b>Appendix</b>	<b>70</b>

<b>6</b>	<b>Ocean biogeochemical cycles .....</b>	<b>73</b>
6.1	Basic controls on biological productivity	78
6.2	Comparing model vs. observations	83
6.3	Temperature and the biological pump	85
6.4	Iron co-limitation of biological productivity	87
6.5	Further ideas	90
6.5.1	Further analysis ideas .....	90
6.5.2	Further experiment ideas .....	90
6.5.3	Other thoughts and suggestions .....	91
6.5.4	Biological pump background information .....	92
6.5.5	Advanced experiment ideas .....	93
<b>7</b>	<b>Marine ecosystems and dynamics .....</b>	<b>95</b>
7.1	Getting going with ECOGEM	97
7.2	Ecosystem configuration	100
7.3	Increasing ecological complexity	103
7.4	Build it up, tear it down	106
7.5	Mixotrophy	110
7.6	Ecology in ... future oceans	111
7.7	Ecology in ... past oceans	113
7.8	Ecology in ... fake oceans!	121
7.9	EcoGENIE 1.1	123
7.10	Further ideas and investigations	124
<b>8</b>	<b>The geological cycle of carbon .....</b>	<b>125</b>
8.1	(none)	126
<b>9</b>	<b>The land surface and carbon-climate feedbacks .....</b>	<b>127</b>
9.1	(none)	128
<b>10</b>	<b>Climates of Past Worlds .....</b>	<b>129</b>
10.1	(none)	130
<b>11</b>	<b>Proxies and Reconstructing the Past .....</b>	<b>131</b>
11.1	(none)	132

<b>12</b>	<b>Synthesis: carbon in the Earth system .....</b>	<b>133</b>
12.1	(none)	134
<b>13</b>	<b>cookie model output .....</b>	<b>135</b>
13.1	(none)	136
<b>14</b>	<b>Introduction to model results analysis .....</b>	<b>137</b>
14.1	(none)	138
<b>15</b>	<b>muffinplot .....</b>	<b>139</b>
15.1	(none)	140
<b>16</b>	<b>muffindata .....</b>	<b>141</b>
16.1	(none)	142
<b>17</b>	<b>cookiegen .....</b>	<b>143</b>
17.1	<b>Introduction</b>	<b>144</b>
17.1.1	cookiegen .....	144
17.1.2	Installing cookiegen .....	144
17.2	<b>Quick Start Guide</b>	<b>145</b>
17.3	<b>Running cookiegen</b>	<b>148</b>
17.4	<b>Configuring cookiegen</b>	<b>149</b>
17.5	<b>Overview</b>	<b>149</b>
17.5.1	cookiegen configuration examples .....	152
17.6	<b>Creating base- and user-configs for cookie</b>	<b>153</b>
17.7	<b>cookiegen parameter settings – details</b>	<b>167</b>
17.8	<b>Appendix – pre-generated Worlds</b>	<b>170</b>
17.8.1	Conceptual worlds .....	170
17.8.2	Idealized-continent worlds .....	170
17.8.3	Modern worlds .....	172
<b>18</b>	<b>FAQ .....</b>	<b>173</b>
18.1	(none)	174
<b>19</b>	<b>HOW-TO (technical) .....</b>	<b>175</b>
19.1	(none)	176

<b>20</b>	<b>HOW-TO (experiment)</b>	<b>177</b>
20.1	(none)	178
<b>21</b>	<b>cookie Development</b>	<b>179</b>
21.1	(none)	180



Before anything else ...

Starting (dozing?) off ...

Logging in!

Downloading/installing the model code

Configuring the code

Testing the model code

**Running the model**

**Creating new experiments!**

**Model output**

Time-slice output

Time-series output

File naming convention

**Viewing model output**

Time-series output

2- and 3-D time-slice output

**Submitting experiment 'jobs'**

'Restarts'



## 0. Installation, Configuration, Basic Usage

Stuff to keep in mind:

- **(cGENIE.)cookie** is a model. Models ARE NOT the ‘real World’. (Don’t get confused!)
- The low resolution (for a 3-D ocean circulation model) of the **cookie** model limits its applicability for very short time-scale problems. In configurations not incorporating the **PLASIM** atmospheric GCM component, there are no atmospheric dynamics or inter-annual variability in the coupled ocean-atmosphere system.
- **cookie** is best thought of as a ‘discovery and exploring’ tool for learning how the Earth system has functioned in the past rather than a detailed ‘simulation’ tool for the future.
- It is possible to have fun.

## 0.1 Before anything else ...

### ReadMe

Some warnings and reminders in this manual are repeated over and over and over and ... over again. Some warnings and reminders are repeated over and over and over and ... over again. This is because you will forget immediately each time! ;)

<sup>!</sup><sup>1</sup>

### Software version naming conventions

You will be using the current version of the cGENIE Earth system model, code-named ‘**cookie**’ (if **Apple** can have ‘Leopard’, ‘Lion’ etc., I can have a baked goods version naming convention). The documentation may not be fully consistent in this respect ... and you may need to translate occurrences of e.g. a directory named ‘cgenie’ to ‘cgenie.cookie’. For brevity, the **cGENIE.cookie** model will be referenced by just ‘**cookie**’.

### linux ...

The (**cGENIE.**)**cookie** model currently naively compiles and runs under **linux**<sup>2</sup> (e.g. distributions such as **Ubuntu**) and unix-like operating systems (such as **macOS**). **cookie** (like most climate models) is configured and accessed (aka ‘run’) at the ‘command line’ of the **linux** (or **macOS** equivalent, which is unix-like) operating system. The command line is a place where you type text and when you press Return, something (hopefully, good!) happens. Typically the stuff you type started with a ‘command’ word, and often followed by one or more options and parameters. The command word and any options / parameters MUST be separated by SPACES.

The start of each line of the command line is indicated with something like: \$. The \$ is called the ‘prompt’ and is ... prompting you to type some input (commands, Tweets, swear words, etc.). See – the computer is just sat there waiting for you to command it to go do something (stupid?). Typically, you will also be informed (reminded) of the user-name, computer name, and current directory, e.g.:

```
[username@host ~]$
```

which is in this example is user ‘username’ (yours will be different!) on computer name ‘host’<sup>3</sup> and the current directory is the ‘home’ directory – represented by the ~ symbol. If, for example, you were instead currently in the **myfolder** directory, you would see at the command prompt something like:

```
[username@host myfolder]$
```

If you are not or not very familiar with the **linux/unix** command line, such as how to navigate up and down the directory tree and to display the contents of the current directory you are in – for a brief summary of some basic/useful **linux** commands and usage – see the **linux** HOW-TO Section towards the back of the book.

---

<sup>1</sup>Also read footnotes please.

<sup>2</sup>For some of you, the mechanics of running the model will be about as much fun as sticking your tongue in an electrical outlet (a popular hobby in England). (However, if you are an experienced **linux/unix/tongue-in-electrical-socket** user, you can skip onto the next Section and save yourself an entire 15 seconds of reading words.)

<sup>3</sup>Sprout the cat will eventually appear under ‘cat-of-the-day’ on my home-page if you press F5 enough times – all my computing clusters are named after my cats ...

NOTE: BE VERY CAREFUL that spaces are not missed out when typing out example lines. Also be careful not to confuse the number one (1) for the letter 'el' (l). Mis-spelling/typing is generally the most likely reason for things not working.

NOTE-the-second: If you find yourself terminally bored typing in long long instructions lines, you may be tempted to simply copy-paste from the cookie manual PDF to the command line, This can work, but firstly be aware that trying to copy-paste multiple lines at once is doomed to failure – copy-paste the first line and then following that add the second (or subsequent) line. Also note that the inverted comma symbol in the PDF is not the same inverted comma symbol that linux is expecting ... You should also make liberal use of the up arrow key that bring back the previously entered command (keep pressing the key for progressively older commands). For instance, a mistake in a command line can be corrected by bringing back the offending line and using the left/right arrows to navigate through the characters and correct any mistake.

NOTE-the-third: In places, instructions may be given for specific programs and computer platforms and hence may differ slightly from the software reality in front of you. Use your judgement in translating such instructions. Many other alternative software choices exist for editing files or viewing results, as are other ways of configuring software and file editing/transferring methodologies. Do what suits you best – you can view such instructions where they occur, as more representing an example methodology rather than a literal interpretation of the Constitution.

## Required computer hardware/software

For running **cupcake**, your options are:

### 1. Remotely

To do this, you will need an account on a linux-based server or cluster. You can use any platform (**linux**, **macOS**, **WindoZ**, as well as **iOS** and **Android** (which is based on **linux** in any case)) to connect to the cluster. You will need the following software on your local machine:

- (a) A terminal ('shell') window. This is no problem for **linux** and **Mac** users (you already have one built in). For **Windows**, either download a simple (and old) **SSH** client (ssh-client) from my website<sup>4</sup> or you can get hold of e.g. **PuTTY** (<http://www.putty.org/>).
- (b) A sftp (secure file transfer) client for convenience (i.e. dragging and dropping files between local and remote computers, and opening files directly on the remote computer cluster). If you have installed ssh-client (**Windows**, above) then a sftp client is already included as part of this software. If using **PuTTY** (**Windows**) you might try downloading **WinSCP** (<http://winscp.net/eng/index.php>). For **macOS**, you can connect to the server through the **Terminal**, but some sftp software for viewing/navigating server file structure include: **FileZilla** (recommended), **Cyberduck**, **TextWrangler**. For **linux**, maybe **FileZilla**.

### 2. Locally

It is possible to install and run the '**(cGENIE.)cookie**' Earth system model either on a linux box (e.g. **Ubuntu**) or on a **Mac**<sup>56</sup>. At a minimum, you will need:

---

<sup>4</sup><http://www.seao2.info/cgenie/software/ssh-client.exe>

<sup>5</sup>Sets of detailed installation instructions are available in the HOW-TO section of this manual.

<sup>6</sup>It is also possible to run **cookie** under **Windows**.

- (a) A FORTRAN (f90 and f77 combatable) compiler of some sort.  
This may come with the operating system as standard, possibly **gfortran**.
- (b) A **git** client.  
If not standard, this is relatively easy to add and install.
- (c) Compiled **netCDF** libraries (not so much fun ...).

To edit files and visualize results, you will need some specific software. The exact software will depend on your operating system, but essential are:

- (a) A viewer for netCDF format spatial data. A **Java** viewer called **Panoply** is provided by NCAR for all platforms – <http://www.giss.nasa.gov/tools/panoply/> (Note that you will need **Java** installed!) (Or alternatively: **MATLAB**, **python**, etc..)
- (b) A simple text editor, except not the rubbish default **Windows** one – you need one that can display **unix** ASCII text without screwing it up. Options for **Windows** users are: **notepad++** (<https://notepad-plus-plus.org/>) **SciTE** (<https://www.scintilla.org/SciTE.html>) (**linux** and **Mac** users need no special/different editor compared with your standard editor – everything will display just fine). You can also use **linux** command line based editors such as **vi**.

### **File editors ...**

You will need to edit text-based configuration files, possibly in installing and configuring **cookie**, but definitely in configuring model experiments. So now might be a good time to check that you can use the/an editor! (You will also be using the same editor to view some of the model output.)

You have two alternative options for editing and viewing text files, depending on whether you are a **unix** nerd with no life, or prefer anything to do with computers to be wrapped in cotton wool and covered with dollops of treacle. EITHER: Use the **linux vi** (**vim**) application (or similar e.g. **emacs**) if you are familiar with it. I think that this pretty much sucks as a text editor and life is far too short and brutal if you don't like this sort of thing ... OR ... use a suitable **linux**-friendly text editor (NOT **Micro\$oft Notepad**) in conjunction with the **Secure File Transfer Client**. For example: **SciTE** (<https://www.scintilla.org/SciTE.html>) is suitable, or **Notepad++**.

If you fiddle about with the settings under Options/Preferences in the **WinSCP** program and apply a little common sense, it should be possible to configure things so that you can simply double-click on a file in the remote (right-hand) window panel and it will open like magic (almost)! Saving the file after editing) should then result in the file being saved back to the cluster. Or you can select Edit With (and then **SciTE**) from right-mouse-button-clicking on the filename. Or ... a crude but workable approach is to use an sftp client to drag the file to your local machine (assuming **cookie** is installed remotely), edit it there, and then drag it back again.<sup>7</sup>

---

<sup>7</sup>Note that care still has to be taken to avoid certain **Microsoft** text editing programs under **Windoze**.)

## Model documentation in general

This (the **cookie** manual), and additional documentation (of varying degrees of up-to-date-ness) can be found:

1. On GitHub.

The **latex** source for the documentation lives here, allowing you to compile the most up-to-date PDF document. And ... make changes yourself and have them incorporated into the official documentation<sup>8</sup>.

However, note that you will have to compile the latex source yourself to create a PDF ...

2. On my website.

Here you can find compiled PDF versions of the documentation ... but it could be a little out of date (the up-to-date latex sources lives on **GitHub**).

## This document in particular

The instructions may not be entirely bug-free -- use your judgment.

## Go!

OK – now we are ready to start ...

---

<sup>8</sup>First clone the **git** repository. Make changes. Commit them locally. Make a 'pull request' ...

## 0.2 Starting (dozing?) off ...

You are going to be installing the model from scratch – why? Why not? It will be a happy character-building experience for you ... trust me ...

### 0.2.1 Logging in!

In running **cookie** locally – log into your **linux/macOS** box (and skip on to the next section)!<sup>9</sup>

Or ... and much more likely to be the case – if you are running **cookie** remotely (e.g. via a user account on a computing cluster or server), then log into the remote server or cluster account using a 'suitable terminal program',<sup>10</sup>:

- If logging in via a **linux/macOS** box, open the terminal/shell window and simply SSH in<sup>11</sup>, e.g.

```
$ ssh username@clusternname
```

where `clusternname`, the cluster (or remote server) name(!) might, for example, be `catname.ucr.edu`, and enter your account password (and tell it whether or not you want this password stored, if asked).

**IMPORTANT!** – When you type in a password in **linux**, NOTHING appears on the screen, not even `*****` as in common on **WindoZ**. As you type (the password), characters are being entered ... you just cannot see them. Don't panic – just type in the password (even if you cannot see characters appearing) and hit `Enter`.

- On a **WindoZ** machine – first start the **WinSCP** program (an sftp file transfer client). Under Host Name, enter the remote server or cluster name (e.g. `catname.ucr.edu`):

The Port number should be set to 22 (except for `sterling.ucr.edu`). Enter your computing cluster user-name on the line below this ('User Name') and then the Password. Click on Login. You will also need a terminal window. This can be opened by clicking on the 'Open session in PuTTY' icon on the top icon row, or pressing `Ctrl+P`.

You should now have TWO windows open – a 'shell' window (lines of text on an otherwise blank screen) and a file manager (transfer) window. Ensure that you have both these before moving on. It is recommended that you maximize both these windows to full screen. (But no-one will die horribly for not doing so. Probably ...).

You can also log in directly from the shell/terminal window e.g. **PuTTY** first (rather than first opening an sftp connection first). You will sill have to open an sftp connection for file transfer.

Note that the cluster you access may not use the standard 'port' number of 22. If your computer uses a different port number – in the login window of your sftp client, simply change the number in the Port box, or if you are using the command line, you will need to type:

```
$ ssh -p xxxx username@clusternname
```

where `xxxx` is the port number.

---

<sup>9</sup>If you fail at this step, you'll have to take up box-modelling instead.

<sup>10</sup>It very much depends on what software you are using. Provided are instructions for some examples, but only examples, and your reality may be rather different.

<sup>11</sup>If your current directory looks something like this: `[username@clustername ~]$` then you are probably already logged in! Otherwise, it will look like: `username@localcomputername:~$`

### 0.2.2 Downloading/installing the model code

The next step is to download/install a copy of the source code for **cookie**. The current release of the cGENIE model (**cookie**) lives here:

```
https://github.com/genie-model/cgenie.cookie
```

There are 2 options:

1. 'cloning'

The preferred/advised way is to *clone* the repository to where you intend to run **cookie**<sup>12</sup>. While you can also use a GUI based git client, easiest is at the command line (e.g. from your HOME directory), using the command `git clone`<sup>13</sup>:

```
$ git clone https://github.com/genie-model/cgenie.cookie.git
```

By doing this, you have created your own code repository (and an identical copy of the one hosted on GitHub). As part of the `git clone` command, you also automatically *check out* (from your very own personal repository) a copy of the code.<sup>14</sup>

2. Simple downloading.

Less good, but OK if you simply want a copy of the code to run an experiment just once, or simply only want to see the code.) By downloading an archive file, containing all the code etc. For this – click on the green Clone or Download button on **GitHub**, and select Download ZIP. You then unpack/unzip the files and directory structure where you want it. This [archive download] is a perfectly workable way to proceed ... as long as you neither want to update the code with whatever new developments or bug fixes occur in the future, nor want to have any code changes you might make, become part of the official **cookie** code (i.e. it becomes a one-off installation that has no connection to the **GitHub** repository.

### 0.2.3 Configuring the code

You may ... or may not, need to configure some local environment settings so that all the libraries etc. that are needed to compile **cookie** can be found. To a very limited extent, **cookie** will try and identify your computer and make the required setting automatically. The changes, if any, that you need to make will depend on the platform where you will be running **cookie** (i.e. the computer where you have just cloned (or downloaded) the code repository to). The currently recognized platforms and required actions are as follows:

The required configuration settings for **cookie** depend on the specific computing platform that you are running on. What follows is a list of the changes associated with some possible platforms that you might be using – find the list name (**bold**) that corresponds to the platform that you are using and ignore all the other options.<sup>15</sup>

---

<sup>12</sup>But see later for other/better ways of working.

<sup>13</sup>This: "... clones a repository into a newly created directory, creates remote-tracking branches for each branch in the cloned repository, and creates and checks out an initial branch that is forked from the cloned repository's currently active branch."

<sup>14</sup>Note that the major difference then with the **svn** system, is that previously, the GENIE code repository existed only on the University of Bristol server, and you *checked out* the code remotely from there.

<sup>15</sup>If you are using a platform not listed here, you may be able to simply adapt the instructions for one of the ones listed.

- **Ubuntu** – no changes are necessary (with caveats).

This is the default assumed platform. No changes are necessary IF the **netCDF** libraries are installed in their default locations and a relatively recent version of **netCDF** is used.<sup>16</sup>

- **sterling** (UCR cluster) – no changes are necessary (the cluster is automatically identified).

- **eevee** (UCR cluster) – no changes are necessary (the cluster is automatically identified).

- **macOS**

Please refer to the separate **macOS** instructions.

- **Windows**

Please refer to the separate **Windoz** instructions.

- **Otherwise ...**

... one or more edits will be required to the file user.mak, which lives in the genie-main directory.

1. At the end of the user.mak, the **netCDF** path needs to be changed.

First comment out<sup>17</sup>. i.e.,

```
### DEFAULT ###
NETCDF_DIR=/usr/local
```

to

```
### DEFAULT ###
#NETCDF_DIR=/usr/local
```

Then un-comment<sup>18</sup> the line: #NETCDF\_DIR=my\_path

AND change my\_path to be the path to your **netCDF** libraries ... the trick step :o)

2. For MacOS users, there are alternative machine type options depending on your CPU. You need to comment out MACHINE=LINUX in user.mak and then chose and un-comment one of the following:

```
#MACHINE=OSX # Intel processor
#MACHINE=OSX_M # Apple silicon (M1, M2 etc.)
```

If your computer is configured to use python v.2 by default and/or does not have python v.3, then you will need to make a few minor changes (see HOW-TO).

<sup>16</sup>By 'relatively recent' – the default settings assume that the **FORTRAN** and **C netCDF** libraries are separate.

<sup>17</sup>To 'comment out' a line – simply add a # symbol to the very start of the line. When **cookie** runs, this line will be ignored.

<sup>18</sup>To 'un-comment' a line – simply remove (delete) the # symbol from the beginning of the line.

## 0.2.4 Testing the model code

Finally, you need to test the code to ensure that all the files have been cloned/installed correctly.

First, change directory (see: Figure 1, and refer to the linux HOW-TO) to<sup>19</sup>:

```
cgenie.cookie/genie-main
```

If you are not ‘linux-friendly’ (see: Section ?? for **linux** basics) – maybe at first do this in steps – list the contents of the directory (`ls`) to check where you are (i.e. what directories are available to chance to), then change to `cgenie.cookie` (`cd cgenie.cookie`), then list again (`ls`) (and see what further directories are there), then change to `genie-main` (`cd genie-main`), and only then ... type<sup>20</sup>:

```
$ make testbiogem
```

This compiles a basic carbon cycle enabled configuration of **cookie** and runs a short test, comparing the results against those of a pre-run experiment (also downloaded alongside the model source code). It serves to check that you have the software environment correctly configured. There may be some ‘Warnings’ reported (== somepony’s sloppy programming) but these are not detrimental to the ultimate science results (we hope!). If you see the error message:

```
make: *** No rule to make target
```

when you try this, you are probably in the wrong directory (it should be `cgenie.cookie/genie-main`).

‘Success’ of this test is indicated by the message:

```
**TEST OK**
```

If you see this – you can then be certain that the model you have installed is producing identical (within tolerance) results to everyone else in the World who has ever installed **cookie**. Note that the model will pause for a l o o o n g time at the line:

```
./genie.job -t -k -f configs/eb_go_gs_ac_bg_test.xml -o /home/genie00/cgenie_output  
-c /home/genie00/cgenie -g ../../cgenie -m "" > testbiogem.out;
```

This is quite ‘normal’ – the model is thinking! Also – ignore the compiler warnings ...

This completes the basic code installation.

If the test doesn’t ‘work’,<sup>21</sup> – try issuing the command:

```
$ make cleanall
```

and then re-try the test.

Refer to the FAQ section at the end of this book for further clues as why the model installation appears not to be working.

<sup>19</sup>Note: the model is \*always\* run from `cgenie.cookie/genie-main`

<sup>20</sup>Remembering that the \$ is to indicate the command line, and you do not actually type it in.

<sup>21</sup>Note that if you have disabled the compilation of the C program that compares netCDF files, the test will run but never complete and you’ll not get a `**TEST OK**` reported, even if the test experiment ran correctly.

Note that **GitHub** does not host files larger than 100 MB, and the 'lookup table' for calculating opal dissolution in sediments (see e.g., *Ridgwell et al.* [2003]) is larger than this. It has hence been committed to **git** as an archived file. Only if a silica cycle is to be employed in your **cookie** experiments, does this file need to be unpacked. A script is provided for this – from `genie-main`, type:

```
$ ./installcookie.sh
```

and this will unpack the opal lookup table to `genie-sedgem/data/input` as well as unpacking a copy of the calcium carbonate lookup table<sup>22</sup>.

---

### Some brief notes on git

The simplest workable installation of **cookie**, as described above, is to use `git clone`. You end up with a carbon copy of the cookie code repo (I am too lazy to type out 'repository' again) which you have also automatically now 'checked out'. Changes and developments will occur to the code on the GitHub repo from time-to-time, and it may be that either you might benefit by using a more up-to-date code base, or specific changes may have been made that you absolutely need. To determine the status of your repo, type<sup>23</sup> (from `cgenie.cookie`):

```
$ git status -uno
```

However, git has not actually compared your repo with the one on GitHub, because this is apparently network expensive (as if you don't spend the rest of your life killing the internet by streaming Rick and Morty). So, use:

```
$ git fetch
```

which will 'download [new and modified] remote content but not update your local repo's working state'. If you now type `git status -uno`, **git** can tell you if there is newer content (e.g. 'Your branch is behind 'origin/master'<sup>24</sup>). To merge in the fetched content:

```
$ git merge
```

Both these commands are also combined in a single command<sup>25</sup>:

```
$ git pull
```

If you are not developing code, and hence not editing files in the repo (but e.g. only adding new model configuration files), your life should mostly be trouble-free with regards to updating your code via `pull`.<sup>26</sup>

---

<sup>22</sup>For now, the *CaCO<sub>3</sub>* lookup table also included in the git repo in its expended (unpacked) form.

<sup>23</sup>Here, the `-uno` ensures that files (e.g. experiment configuration files) you have created, but not added and committed, are not listed.

<sup>24</sup>`origin` is the origin of the repo – GitHub, and `master` is the name of the branch, in this case, the default branch name.

<sup>25</sup>If there are no changes (to `fetch` and `merge`), the you get the message: `Already up-to-date.`

<sup>26</sup>One exception is, if any of the model installation/configuration files that you might have edited, i.e. one or all of: `genie-main/user.mak`, `genie-main/makefile.arc`, and/or `genie-main/makefile`, have also changed on `origin/master`, then `pull` (or `merge`, after `fetch`) will fail (with the message: 'Please, commit your changes or stash them before you can merge. Aborting'). The root issue is a conflict between remote file changes, and local ones that you had not committed to your repo.

You probably want to keep your local configuration changes, otherwise you'll probably end up re-doing them all over again. One solution is to sneakily 'hide' (`stash`) them out of sight, by: `$ git stash` You can now `pull`, updating your repo with respect to `origin/master`. Then – you want your changes back, so apply the stashed changes: `$ git stash apply` If unlucky, there will be conflict as your stashed changes are merged back onto your local repo branch (master). If so, you'll need to edit the file, deciding which line(s) is correct, and delete the version you do not want, along with the ASCII tags/labels.

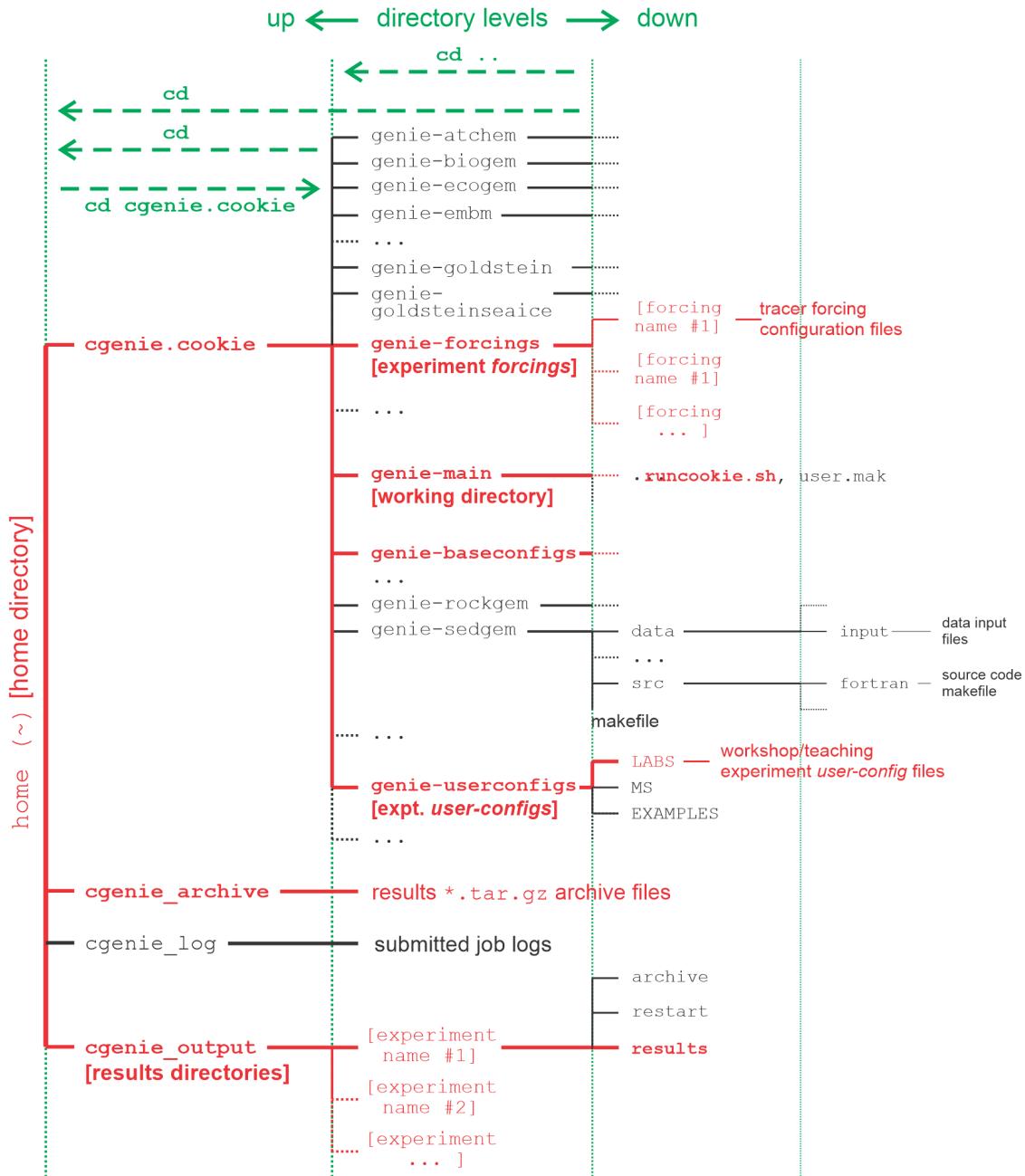


Figure 1: Directory structure of the **cookie** model. Highlighted in red are directories and sub-directories that you will need to access at some point. Vertical green lines designate directory levels, with example commands shown for moving between them.

**0.3****Running the model**

The overall sequence of configuring and running **cookie** (job submission to a cluster queue), is shown in Figure 2. Refer to this if in any doubt at any point.

At the command-line (\$) and in the genie-main directory (not your home directory), you will be entering in a command (`./runcookie.sh`) together with a list of parameters that will be passed to the model, and as if by magic the model will run (or sometimes not). The form of the command you are going to be issuing is:

```
$ ./runcookie.sh #1 #2 #3 #4 (#5)
```

(don't type it yet!) It requires that you must list at least 4 parameters after `./runcookie.sh`, separated by S P A C E S and on a single continuous line (even if it 'wraps' around across 2 lines of the screen). These parameters are:

- #1 ... is the name of the required base (or 'basic') configuration ('*base-config*') of the model.
- #2 ... is the name of the subdirectory (if any) containing the user configuration ('*user-config*') file (i.e., the file containing the specification of a particular experiment).
- #3 ... is the name of the experiment itself. There must exist a file in the directory specified by parameter #2 (LABS) with exactly the same name as you enter here for parameter #3 (i.e. parameter #3 points to a file in the directory given by parameter #2).
- #4 ... is the run length of the experiment in years – this must be entered as an integer.
- (#5) ... There is also one optional (5th) parameter (to be described later).

As an example of running the **cookie** Earth system model:

- #1 : The *base-config* is: `cookie.CB.p_worbe2.BASES`
- #2 : The *user-config* directory is: LABS
- #3 : The *user-config* file (the experiment name) is: LAB.0.3.EXAMPLE.
- #4 : Run the experiment for ten years: 10
- #5 : (There is no restart file, and so no 5th parameter needs to be passed ...)

The full command for your first example experiment, which you are going to issue from the `~/cgenie.cookie/genie-main` directory, then looks like:

```
$ ./runcookie.sh cookie.CB.p_worbe2.BASES LABS LAB.0.3.EXAMPLE 10
```

(you can try it now!)

**REMEMBER: This must be entered on a single CONTINUOUS LINE.**

**The (single) S P A C E S are vital.**

**ALSO – take care not to confuse an el ('l') with a one ('1') ... (it is a 'one' here).**

What should happen is: First, you will end up twiddling your thumbs a while, as all the components of **cookie** are compiled from the raw source code (**FORTRAN**). When it has finished doing this, the model will initialize and carry out some brief self-checking. Only then will it start actually ‘running’ and doing something ... in this particular example, the output should look something like this:

```
*****
*** Initialisation complete: simulation starting ...
*****
>   model year | pCO2(uatm)  SAT(°C)  AMOC(Sv)  ice(%)  SST(°C)  SSS(PSU)  | pH  OHMEGA  | [O2] (uM) fPOC(PgC/yr)
>   0.5 | 285.042 -3.667 | 12.715  0.274  1.408  34.902 | 8.048  2.537  | 181.360 12.327
>   1.5 | 295.295 -2.531 | 13.189  2.126  3.450  34.902 | 8.079  2.896  | 191.671 12.636
>   2.5 | 302.383 -1.270 | 12.266  4.282  5.004  34.904 | 8.088  3.131  | 196.879 12.394
>   3.5 | 307.657 -0.274 | 11.020  5.525  6.214  34.906 | 8.093  3.318  | 199.930 12.216
>   4.5 | 311.667  0.628 | 9.831   6.037  7.185  34.908 | 8.096  3.469  | 201.861 12.096
>   5.5 | 314.811  1.437 | 8.876   6.387  7.987  34.910 | 8.099  3.597  | 203.157 11.994
>   6.5 | 317.299  2.150 | 8.091   6.628  8.667  34.911 | 8.101  3.707  | 204.047 11.899
>   7.5 | 319.290  2.782 | 7.872   6.673  9.256  34.913 | 8.103  3.800  | 204.669 11.822
>   8.5 | 320.852  3.342 | 7.886   6.734  9.773  34.914 | 8.105  3.882  | 205.073 11.742
>>> SAVING BIGGEM TIME-SLICE AVERAGE CENTERED @ year : 9.500
>   9.5 | 322.066  3.848 | 7.935   6.684  10.232 34.914 | 8.106  3.954  | 205.320 11.670
*****
*** Simulation complete: shutdown starting ...
*****
```

(Note that you might want to increase the width of your terminal window in order to get the values reported for each year appearing on the same line and not wrapped-over to the next one.)

What do all these numbers mean? From left-to-right:

model year – ... guess!

Then:

$pCO_2$ (uatm) -- mean atmospheric  $CO_2$  concentration (in units of  $\mu\text{atm}$ )  
 $\langle SST \rangle$  – global mean surface air temperature (‘SAT’) °C

Note that if you are not running an experiment without a carbon cycle, there will be no values for  $pCO_2$ .

$AMO$ (Sv) – Atlantic meridional overturning circulation (AMOC) (Sv)  
 $ice(%)$  – global sea-ice fraction (%)  
 $\langle SST \rangle$  – global sea surface temperature (‘SST’) °C  
 $\langle SSS \rangle$  – global sea surface salinity ‘SSS’ (‰)

Note that if you are not running an experiment with a modern-like Atlantic basin, no value is given for the AMOC.

After that, a couple of carbonate chemistry indicators:

pH – Mean ocean surface pH.  
OHMEGA – Mean ocean surface OHMEGA ...

which will become apparent later.

Note that if you are not running an experiment without a carbon cycle, no values will be given.

Finally:

- 02 – Dissolved oxygen in the ocean(global mean). ( $\mu mol kg^{-1}$ )
- fPOC – Global total annual export production (in terms of particulate organic carbon ( $Pg Cyr^{-1}$ )).

Note that you need some sort of biological pump in the ocean for these to be meaningful.

---

The choice of what information to display on screen as the model is running is rather arbitrary, but the chosen metrics do tend to summarize some of the main properties of the climate system and carbon cycle – for my own personal convenience rather than reflecting any fundamental scientific truth ...

This information is reported at the same intervals as *time-series* data (see later and/or refer to the User Manual) is saved and is indicated by:

Interleaved between these lines are lines reporting the saving of *time-slice* data (the 2- and 3-D model states – more of which later as well as in the User Manual). These appear as:

>>> SAVING BIOGEM TIME-SLICE AVERAGE CENTERED @ year:

**You can stop the model at any point (all data up to that time will have been saved) by hitting: <Ctrl-C> (CONTROL key + ‘C’ key).**

---

Just from examining the screen output: how close to steady state does the system appear to have come after just 10 years? i.e., do SST and/or sea-ice extents appear to be converging towards stable (constant) values? This will be an important question to think about later on: ‘has the model reached steady-state (and does it matter)?’

---

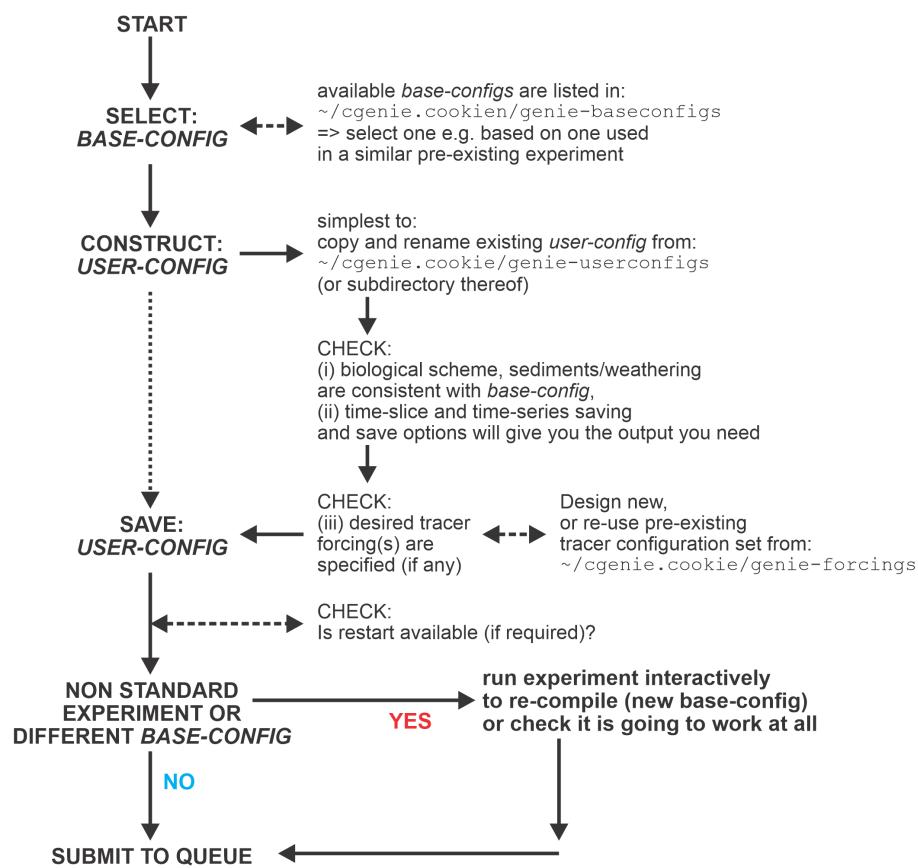


Figure 2: Schematic of the sequence-of-events in configuring and running an experiment.

**0.4****Creating new experiments!**

The key to creating new experiments is to **remember that the name of the *user-config* file, that contains the parameter settings that define that specific experiment, becomes the name of the experiment and hence name of the model results sub-directory in `cgenie_output`.** Changing the name of the *user-config* file, hence leads to a new experiment name and a new model results sub-directory. (Conversely, not changing the name of the *user-config* file and re-running it results in the results of any previous experiment run using that *user-config* file, being over-written.)

There are two obvious ways to create a new *user-config* file and hence new experiment:

1. Create a blank text file <sup>27</sup> and populate the contents with the parameter value assignments you need for your new experiment.

Inevitably, it is difficult to remember all the names and even values you want to specify, meaning that you'll end up looking at existing *user-config* files and copying and pasting lines (or even the entire contents) from the old file(s) into your new *user-config* file. So then you may was well ...

2. Copy and edit an existing file!

This is the most practical approach – pick an existing *user-config* file that is closest to the specific experiment that you want to run – copy and rename it, then edit the contents and save. For the purpose of trying this out, you can literally pick on any existing file in the `LABS` subdirectory (of `genie-userconfigs`).

You can do this by:

- Using your sftp client (program):

First – drag the existing *user-config* file to your local computer. On your local computer, rename the file as per how you would 'normally' rename any file.

Now you can edit the parameter values, re-save it, drag it back to the cluster (from your local computer) using the sftp client ... and finally run the experiment.

There are also ways of configuring **sftp** clients so that you can double-click on a file in the remove window, edit it, and save it back to the remote computer.<sup>28</sup>

- Or, if you are comfortable working at the linux command line; working from the same directory (e.g. `LABS`) that the file you want to copy lives in, you can copy a file (to a new filename) by:

```
$ cp oldfile newfile
```

(after which you can edit the new *user-config* file `newfile`, re-save, and then run the experiment.)

Remember that the new *user-config* file needs to be saved in (or copied to) the `genie-userconfigs/LABS` sub-directory (in the case of experiments carried out as part of the tutorials described in this manual), or `LABS` itself or any sub-directory (or sub-sub- etc directory) ... as long as you specify the correct path to the directory where you save the new file to<sup>29</sup>.

---

<sup>27</sup> \$ `touch file.txt` will achieve this at the linux command line.

<sup>28</sup> Actually what happens is that the file is transferred locally, opened, and you edit a local copy. When you save it (locally), it is automatically transferred back.

<sup>29</sup> See previous Section and the use of the 2nd parameter passed to `runcookie.sh`

## 0.5 Model output

Experiment results are saved in a single sub-directory of the experiment results directory (along with a sub-directory for *restart* files, and one for copies of your parameter choices). The directories containing the results of all your different experiments live in:

```
~/cgenie_output
```

and will be assigned a directory name something like:

```
LAB.0.3.EXAMPLE
```

(this being the results directory name for an experiment called LAB.0.3.EXAMPLE) and the actual results live in a sub-directory of this – results, hence:

```
~/cgenie_output/LAB.0.3.EXAMPLE/results
```

NOTE: If in an sftp client window you find that you cannot ‘see’ the cgenie\_output directory ... or cannot find any of the results sub-directories you are expecting, you will need to refresh the directory listing (e.g. for **WinSCP**, there is a double green-arrow Refresh icon button near the top right of the window that you can click on). sftp client programs generally do not automatically refresh directory listing on the remote computer.

**cGENIE.cookie** has a flexible and powerful facility of saving results by means of spatially explicit ‘time-slices’, and as a semi-continuous ‘time-series’ of a single global (or otherwise representative mean) variable, described as follows.

### 0.5.1 Time-slice output

One of the most informative data sets that can be saved is that of the spatial distribution of properties (such as tracers or physical ocean attributes). However, saving full spatial distributions (e.g., a  $36 \times 36 \times 8$  array) for any or all of the tracers each and every time-step is clearly not practical; not only in terms of data storage but also because of the detrimental effect that repeated file access has on model run-time. Instead, **BIOGEM** will save the full spatial distribution of tracer properties only at one or more predefined time points (in units of years). These are termed *time-slices*. At the specified time points, a set of spatially-explicit data fields are saved for all the key tracer, flux, and physical characteristics of the system. However, rather than taking an instantaneous snapshot, the time-slice is constructed as an average over a specified integration interval (the default is set to 1.0 years, i.e. an annual average). **BIOGEM** assumes that the specified time point represents the mid-point of the (annual) average with the results that output years end up being reported.

For example, to save regularly every 10 years you would set:

```
bg_par_data_save_slice_timeinterval=10.0
```

and which would give you save points at:

```
9.5  
19.5  
29.5  
39.5  
...
```

(the mid-points of averages made over the intervals: 9-10, 19-20, 29-30, 39-40 years, etc.). Otherwise, you can specify the name of a file containing times to save at, or simply only save at the end. (see later)

### 0.5.2 Time-series output

The second data format for model output is much more closely spaced in time. Model characteristics must then be reducible to a single meaningful variable for this to be practical (i.e., saving the time-varying nature of 3-D ocean tracer distributions is not). Suitable reduced indicators would be the total inventories in the ocean and/or atmosphere of various tracers (or equivalently, the mean global concentrations / partial pressures, respectively). Like the time-slices, the data values saved in the time-series files represent averages over a specified integration interval (the default is set to 1.0 years (annual average) but the results are reported with respect to the mid-point of the average which is where the ‘.5’ bits come in again).

The much smaller and simpler (text) file format now allow you to save much more frequently. For example, for experiments up to a few 100 years, you could save every single year, which you would set by:

```
bg_par_data_save_sig_timeinterval=1.0
```

Otherwise, you can specify the name of a file containing times to save at. (see later)

### 0.5.3 File naming convention

The results directory will contain files with names of the form:

- fields\_biogem\_2d.nc – 2-D fields of ocean and atmosphere properties, as **NetCDF**.
  - fields\_biogem\_3d.nc – 3-D fields of ocean properties, as **NetCDF**.
  - timeseries\_\*.txt – these are the time-series files (in ASCII / plain text format).
  - SUMMARY\_AT\_year\_\*\_diag\_GLOBAL.txt – these contain (global diagnostics) summary information and are saved at the same frequency as the time-slices (also as ASCII / plain text).
- 

#### Alternative file naming convention ...

Having the e.g., 2D and 3D **netCDF** files always called the same name (fields\_biogem\_2d.nc, fields\_biogem\_3d.nc) in each and every experiment, has the potential to get confusing if you have multiple experiments open simultaneously in a viewer such as **Panoply**.

You can specify that **netCDF** files are named following the name of your experiment, by adding the following line to your *user-config*:

```
bg_ctrl_ncout_expid_name=.true.
```

## 0.6 Viewing model output

### 0.6.1 Time-series output

A descriptive summary of all the time-series (biogem\_series\_\*.res) data files is given in the **cookie** User Manual if you are really that bored. The files of most immediate use/relevance are:

- timeseries\_atm\_humidity.res - mean atmospheric (surface) humidity
- timeseries\_atm\_temp.res - mean atmospheric (surface) air temperature
- timeseries\_misc\_opsi.res - overturning stream-function (e.g. AMOC) strength
- timeseries\_misc\_seaice.res - mean ocean sea-ice cover and thickness
- timeseries\_ocn\_sal.res - mean ocean surface and whole ocean salinity
- timeseries\_ocn\_temp.res - mean ocean surface and whole ocean temperature

---

One way of viewing the contents of files (in a shell window/terminal) is to change directory to the experiment results directory and opening the file in a file editor at the command line. But that is not so much fun.

Instead – change to the experiment results directory and then to the biogem sub-directory in the Secure File Transfer Client, and try double-clicking (if you have set up the **WinSCP** preferences correctly) or right-mouse-button-clicking (the then Edit with) on one of the .res files (listed above). For timeseries\_ocn\_temp.txt, you should see 5 columns:

```
% time (yr) / temperature (C) / _surT (ice-free) (C) / _benT (C) / _surT (C)
```

for: model time (years) (actually, the mid-point in time of an annual average), annual mean ocean temperature (averaging over the entire ocean volume), annual mean ocean surface temperature (excluding ice-cover areas), annual mean benthic (sea-floor) temperature, annual mean ocean surface temperature (now including ice-cover areas). Other results files may differ in the numbers of columns but all should be identifiable from the header (first line) information.

Remember: **WinSCP** does not automatically refresh the directory listing. If you cannot see the results sub-directory with the experiment name you have just run, 99 times out of 100, it is because the display of the **WinSCP** needs to be refreshed – there is an icon at the top of the program window or hit the ‘F5’ key.

---

For your information and edification (only): **Excel**, or **MUTLAB** if you prefer, can be used to graph the time-series results. Either way you will have to deal with the header line(s) that are present at the top of the file (and preceding the rows of data).

In **Excel**: Choose File then Open. You will want to select Files of Type ‘All Files (\*.\*)’. In the Text Import Wizard window you can request that **Excel** skips the first few lines to start the import on the 2nd or 3rd line of the text file. Alternatively: set an appropriate column width manually in **Excel** to ensure that the columns of data are correctly imported.

**MUTLAB** will ignore lines starting with a %, which the time-series starts with. However, it may be that the header line wraps-around and there is in effect a 2nd header line but without a %. In this case, extra care (or a quick edit of the header in the ASCII file) will be required to load the data into **MUTLAB**.

### 0.6.2 2- and 3-D time-slice output

For the time-slice **NetCDF** (\*.nc) files you will be using a program called **Panoply**. If you want your own (FREE!) copy of this utility, you can get it here (and is available for: **WindoZ**, **Mac**, and **linux** operating systems): <http://www.giss.nasa.gov/tools/panoply/>.

When you open the **NetCDF** file, you will be presented with a ‘Datasets and Variables’ window (on the left hand side of the application window). This contains a list of all the parameters available that you can display. You will find that the ‘Long Name’ description of the variable will be the most helpful to identify the one you want. Simply double-click on a variable to display.

For the 3-D fields you will be asked first whether you want a ‘Longitude-Latitude’ or ‘Latitude-Vertical’ plot (for the 2-D fields, the plot display will immediately open). For the ‘Longitude-Latitude’ plots – there are multiple levels (depth layers) in the ocean - these data that can be plotted from the surface to the abyssal ocean. For the ‘Latitude-Vertical’ plots – there are multiple possible longitudes at which to plot slices. The default is the global mean meridional distribution. There is also an option for ‘Longitude-Vertical’ plots (which we will not use).

There may be multiple time-slices (i.e., you can plot data saved from different years). By default, only the very first *time-slice* will be displayed.

You can choose interpolate the data or not (often you may find that it is clearer not to interpolate the data but to leave it as ‘blocky’ colors corresponding to the resolution of the model), change the scale and colors, overlay continental outline, change the projection, etc etc. Grey cells represent ‘dry’ grid points, i.e., continental or oceanic crust.

NOTE: The default settings in **Panoply** can mislead. Be aware of:

1. **Panoply** initially displays the very 1st time-slice (often year mid-point 0.5) time-slice rather than the experiment end. This can confuse and look like an experiment has not done anything!
2. By interpolating the data (not always misleading). To remove interpolation, un-tick: ‘Interpolate’ in the ‘Arrays’ tab.
3. By displaying a global zonal mean by default when selecting Latitude-Vertical plots. Then, to further confuse you, by plotting the output up-side-down (to invert: in the ‘Grid’ tab, hit ‘Swap B/T’ (for swap bottom/top)).
4. By listing all ‘Plottable variables’ (option at the bottom of the window), when what you *ideally* want are the shorter and less confusing list of ‘Georeferenced variables’.
5. In Longitude-Latitude plots, by overlaying the modern continental output. (**cookie** land is marked in grey.)
6. By fitting a scale to the plot when the display window is opened, but not changing the scale when e.g., time or depth is changed. (The point of confusion is that you can quickly move outside the scale and end up with all model points dark blue or red.) Re-fit the scale, or manually set limited, in the ‘Scale’ tab. So be careful when opening a new plot that you are looking at what you \*think\* you are looking at ... All the defaults can be changed via the ‘Edit’ drop-down menu and ‘Preferences’.

To save plots in **Panoply**:

File

Save Image As ...

Then select the location, filename, and graphics format.

## 0.7 Submitting experiment ‘jobs’

This bit is no particular fun at all, but it is a very handy ‘trick’ for running the model in the background, and maximizes drinking time in the bar vs. sat bored watching a computer screen :)

---

Running jobs interactively is all very well, but there are three important limitations:

1. The connection between your terminal and the server computer running the model must remain unbroken. Anything more than a fleeting loss of internet connectively may result in the experiment terminating.
  2. You can only run one experiment at a time ... unless you want to have thousands of separate terminal open ... ? I thought not ...
  3. Any cluster or computer you are likely to be accessing using a shell will not have many computing cores itself, either because it is a single machine with only one or two processors, or if a cluster, by using a terminal you are running on the ‘head node’, which will have similar computing core limitations to running on a single machine. The more experiments you run simultaneously, the slower they will all run ...
- 

The alternative is to submit your experiment as a ‘job’ to a queuing system which then manages what compute resources are used to run the model. Once you have submitted the experiment, that is it – you can go straight to the bar :)

For example – to run the same experiment as before (LAB\_0.EXAMPLE) for maybe 100 years (or even longer if you wish – I am just pulling factors of 10 out of thin air here) but now submit the experiment as a job to the cluster queue, type (again: SINGLE, CONTINUOUS LINE):

```
$ qsub -q QUEUE.q -j y -o cgenie_log -V -S /bin/bash
runcookie.sh cookie.CB.p_worbe2.BASES LABS LAB.0.3.EXAMPLE 100
```

Here, the queue name for this particular cluster is QUEUE.q. The cluster you are actually using will have a different queue name (e.g. refer to any cluster-specific information that you might have).

Note that now you should omit the ‘./’ bit before `runcookie.sh`. (If you are interested (I know that you are not): the options following `qsub` and before `runcookie.sh` do things like re-directing screen output and error messaging to a file and specify which linux ‘shell’ to assume. It is even possible to receive an email when the job is done :) ) The status of the cluster queue and how your experiment job is getting on (e.g., “Is it finished yet?”) can be checked by typing:

```
$ qstat -f
(qstat -f -u "*" will show all jobs on the cluster.)
```

After submitting an experiment, you receive a job number. This number appears in the first column in the queue status information when you issue a `qstat -f` command. You should see your job appear on one of a number of compute nodes, perhaps numbered 0-0 through 0-5), although it might briefly reside as a ‘PENDING JOB’. For each node, there are multiple processing cores (depending on the specific cluster and queue), meaning that multiple instances of `cookie` can run simultaneously on each node. For an 8-level ocean based configuration of `cookie`, being run for 100 years, the job should remain there in the queue for a few minutes before ‘disappearing’ (your clue that it has finished, or died<sup>30</sup> ...). If you periodically re-issue a `qstat -f` command you can follow your job’s progress.

---

<sup>30</sup>If your experiment appears on the queue but vanishes after a few seconds, it has most likely died :(

A rough rule of thumb is that 8-level ocean **cookie** @ a horizontal grid resolution of 36x36 will simulate about 1000 years per CPU hour. The 16-level version (which you will use later), runs at about 300-400 years per CPU hour.

**NOTE:** It may be that the **FORTRAN** compiler is not accessible by the computer nodes. The implication of this is that *the cookie executable must be already compiled BEFORE a job is submitted to the queue*. In other words; if you have just changed the model resolution or continental configuration, or number of tracers (i.e. changed the *base-config*) or issued a `make cleanall` command you MUST briefly run your desired experiment (or equivalent) interactively (i.e., in the shell window) to ensure that everything is correctly compiled. For instance, either run the experiment for a couple of years or start the experiment for the desired full duration, but 'kill it' (Ctrl-C) once the experiment is running successfully.

## 0.8 'Restarts'

Not much fun here either . . . but again... an important and time-saving (== increased drinking time!) modelling technique to learn to use.

By default, model experiments start from ‘cold’, i.e., the ocean is at rest and uniform in temperature and salinity while the atmosphere is uniform in temperature and humidity. All biogeochemical tracers in the ocean have uniform concentrations and/or are zero and there are no biogenic materials in deep-sea sediments. From this state it will take several thousand years (kyr) for the climate system to reach steady-state, and closer to 5 kyr (or more) for ocean biogeochemical cycles and atmosphere  $CO_2$  to reach steady-state, and exceeding 100 kyr for sediment composition to re-balance weathering ... Reaching this the equilibrium state is called the ‘*spin-up*’ phase of the model. There is evidently little point in repeating the *spin-up* for each and every model experiment that are similar except in a single detail (e.g., testing a variety of different  $CO_2$  emissions scenarios all starting from current year 2012 conditions). A facility is thus provided for requesting that a ‘*re-start*’ is used – starting a new experiment from the end of a previous one, usually a *spin-up* that has been run explicitly for the purpose of generating a starting point (*re-start*) of the system at steady-state (equilibrium) for subsequent experiments to continue on from. It is important to note that there is nothing special about a *re-start* – it is simply an experiment that you have already run. Equally, there is nothing special about the *re-starts* you will download next – these you could have generated yourself – it simply saves time to have them provided.

To experiment with using a *re-start*, you will first need to download a file that has been created (a pre-run 10,000 year spin-up). To fetch this: Change to the cgenie\_output directory (perhaps by going ‘home’ first (`cd <Enter>`), and then changing to cgenie\_output – refer to linux commands HOW-TO and Figure 1.1), and type:

```
$ wget --no-check-certificate http://www.seao2.info/cgenie_output/
cookie.CB.p_worbe2.BASES.ridgwelletal.SPIN.tar.gz
```

(all one line!)

This downloads an archived/compressed copy of the restart from a location on the interweb. Extract the contents of this archive by typing:

```
$ tar xfzv cookie.CB.p_worbe2.BASES.ridgwelletal.SPIN.tar.gz
```

Finally, change directory back to cgenie.cookie and then genie-main so that you are ready to run the model (the model is \*always\* run from the cgenie.cookie/genie-main directory).

A *re-start* can be requested for running on a new experiment from the end of a previous one, by providing a 5th (optional) parameter when entering in the `runcookie.sh` command. A spin-up of the modern World climate state is provided for you as a *restart* – `cookie.CB.p_worbe2.BASES.ridgwelletal.SPIN` – which you have just unpacked to the cgenie\_output results output directory.

To test out the use this *restart* – create a new (*user-config*) experiment configuration file in the directory:

```
~/cgenie.cookie/genie-userconfigs/LABS
```

taking the file `LAB_0.3.EXAMPLE` (which is provided) as a template (no parameter changes need to be made yet). As described earlier – copy this file and give it a new name – here, name it: `LAB.0.3.NEW`

You can then specify the use of the *restart* in your new LAB\_0.3.NEW experiment:

```
$ ./runcookie.sh cookie.CB.p_worbe2.BASES LABS LAB.0.3.NEW 100  
cookie.CB.p_worbe2.BASES.ridgwelltal.SPIN
```

The run-time output should now look noticeably different. There should be no (or perhaps just very little) drift in any of the various variable values outputted to the screen – this is because you have (re-)started from the end of a run that had already ready an equilibrium (or close).

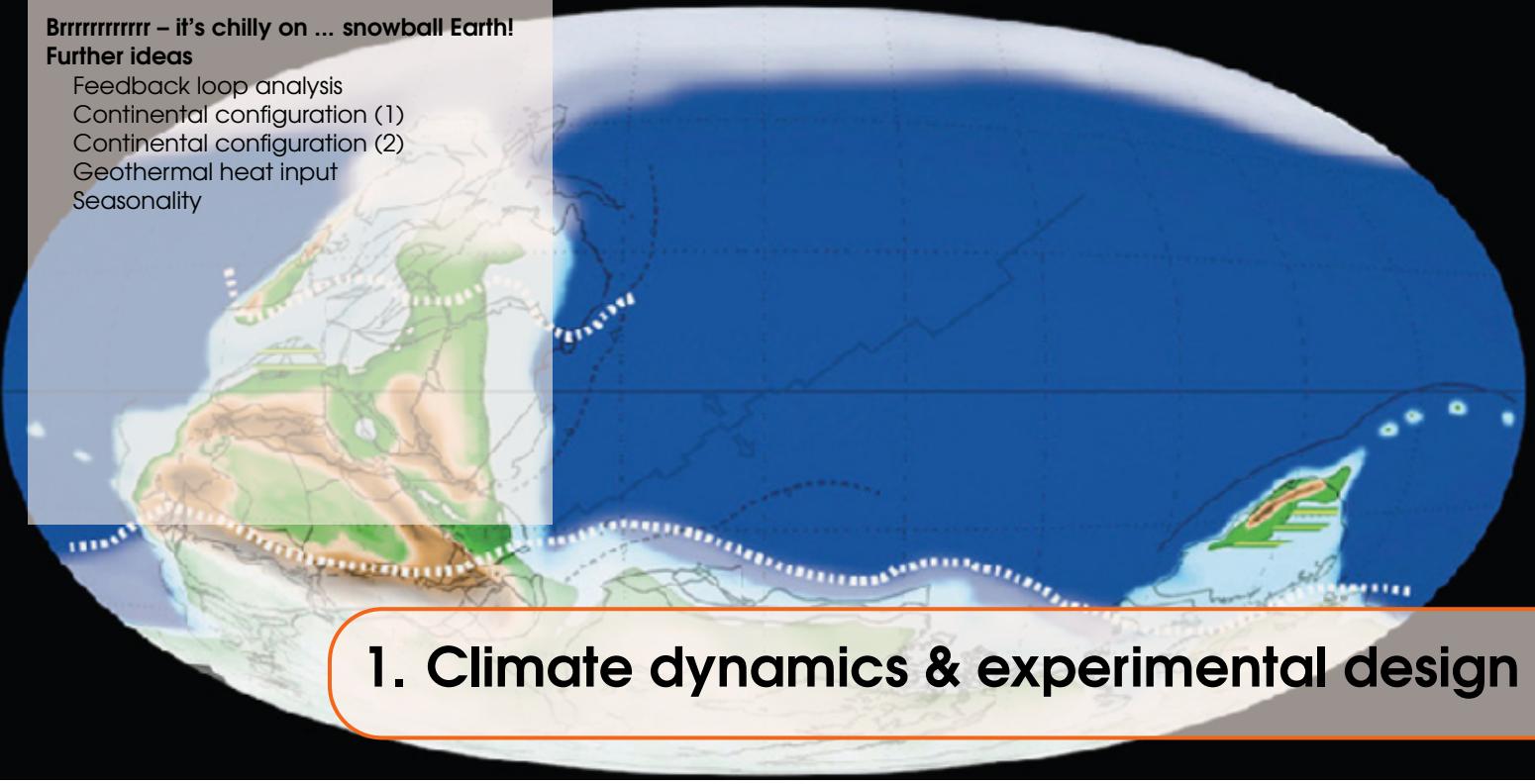
Because LAB.0.3.EXAMPLE (and hence your copy, LAB\_0.3.NEW) does not prescribe a value of  $pCO_2$  but instead allows it to 'float', your experiment is a good test of whether the carbon cycle in experiment

cookie.CB.p\_worbe2.BASES.ridgwelltal.SPIN was at steady state.

Brrrrrrrrrrr – it's chilly on ... snowball Earth!

#### Further ideas

- Feedback loop analysis
- Continental configuration (1)
- Continental configuration (2)
- Geothermal heat input
- Seasonality



## 1. Climate dynamics & experimental design

Stuff to keep in mind:

- Models ARE NOT the ‘real World’ (it is going to be pretty obvious this is the case here).
- Don’t believe what you read in *Nature* or *Science* ...

**Readme**

You will need to download a new *restart* file prior to embarking on the snowball Earth experiments. To fetch this: change directory to the cgenie\_output directory, and type (or copy and paste carefully from this PDF ...):

```
$ wget --no-check-certificate  
http://www.seao2.info/cgenie_output/cookie.C.p_woreq1.NONE.SPIN.tar.gz
```

(all one line)

This downloads an archived/compressed copy of the experiment cookie.C.p\_woreq1.NONE.SPIN – effectively, just an experiment (spin-up) that has already been run for 5,000 years for you. Extract the contents of this archive by typing (also from the cgenie\_output directory):

```
$ tar xfzv cookie.C.p_woreq1.NONE.SPIN.tar.gz
```

A new experiment results directly will then appear as if you had just run the entire 10,000 year experiment yourself (and you could in fact have done so). Remember to refresh the **WinSCP** directory view window if you are using this particular software, or it might appear that nothing has been extracted.

You'll then need to change directory back to genie-main to run the model.

If ... when you subsequently try and use this *restart* in an experiment, **cookie** stops and complains that it cannot find it, check:

1. That you in fact downloaded it to the correct directory, which should be: cgenie\_output, and not randomly to e.g. your home directory ...
2. That you unpacked it.

## 1.1 Brrrrrrrrrr – it's chilly on ... snowball Earth!

To illustrate how ‘easy’ it can be to configure an Earth system / climate model such as **cookie** and explore the behavior of the Earth system and its response to perturbation – you are going to induce an extreme cooling of the climate system and see what happens.

Solar output was weaker during the late Neoproterozoic, a time when the Earth experienced a series (2-ish) of extreme glaciations. Thus, having a mild climate state to start with must have been dependent on sufficient  $CO_2$  and/or  $CH_4$  in the atmosphere which presumably must have been highly elevated compared to the modern World ... sort of opposite to the problem we have today ...

You are going to be running experiments in a similar manner to before, and using the *re-start* experiment that you downloaded. You could then, for example, take the experiment configuration file LAB.1.1.EXAMPLE (which is provided for you), and run the experiment, which has a new *base-config*: cookie.C.p\_woreq1.NONE, by typing:

```
$ ./runcookie.sh cookie.C.p_woreq1.NONE LABS LAB.1.1.EXAMPLE 100
cookie.C.p_woreq1.NONE.SPIN
```

(All one line ... but remember there will need to be a space separating the experiment duration and the name of the *restart*).

However ... why not get into the habit of creating new and uniquely named experiments and their associated *user-config* files (no harder than copying it and renaming it – see earlier). If you keep using the same experiment name, the results will be over-written each time you re-run that experiment, whilst having 2 (or more) experiments running simultaneously with exactly the same name causes havoc as they try and over-write each others results files in a somewhat entertaining but ultimately useless way.

So – copy and rename the file LAB.1.1.EXAMPLE to ... LAB.1.1.EXPT or whatever (in the subdirectory LABS), and then run your newly created experiment:

```
$ ./runcookie.sh cookie.C.p_woreq1.NONE LABS LAB.1.1.EXPT 100
cookie.C.p_woreq1.NONE.SPIN
```

Remember that because you have changed *base-config* since the exercises in the previous chapter, running the model first interactively (at the command line) is essential in order for the model code to be re-compiled consistent with the new configuration. i.e. you cannot directly start submitting experiments using the *base-config* cookie.C.p\_woreq1.NONE, straight-away after having previously run the model using the cookie.CB.p\_worbe2.BASES *base-config*. Once you have re-compiled the model with the new *base-config* and starting running (any experiment), you need not re-compile again and can submit as many jobs to the cluster queue as you like, up until you change *base-config* once again.

Overall: your task in this exercise will be to determine the radiative forcing (or rather, p $CO_2$  equivalent) threshold required to drive the climate system into a full ice-covered ocean (snowball Earth) state. (Make sure you have read the Hyde *et al.* [2000] paper.)

Useful 2-D (netCDF — **Panoply**) variables to view include: surface air temperature and sea-ice extent (and/or thickness). Ocean surface temperature and salinity can be viewed in the 3-D netCDF results file but are likely to be of less interest.

Time-series (ASCII .res files) are useful for providing simple mean indicators of global climate such as global ocean fractional sea-ice covered.

Note that the model configuration of an idealized super-continent you are using and as defined by the `cookie.C.p_woreq1.NONE base-config` file, positioned symmetrically about the Equator, is pretty unrealistic. But the further you go back in time, the more uncertain it becomes as to exactly where and in what orientation the continents were. Sometimes modelers have to resort to somewhat idealized experiments if the uncertainties are too great. In addition, one can conduct sensitivity experiments to test whether the continental configuration is important to the results. For instance, Hoffman and Schrag [2002] discuss the potential importance of continental configuration, while the hypothesis of Donnadieu *et al.* [2004] rests on specific details of the continental configuration.

For this configuration, the solar constant is set weaker than modern to reflect the fact that the Sun's output has increased with time and during the Neoproterozoic the solar constant would have been ca. 5% weaker. This is set in the *user-config* file by the model *parameter*:

```
ma_genie_solar_constant= 1285.92
```

which is set at the top of the provided *user-config* file `LAB.1.1.EXAMPLE`. (For reference, the modern value is  $1368\text{Wm}^{-2}$ .)

To search for the atmospheric  $\text{CO}_2$  concentration (or rather, radiative forcing equivalent) that would lead to a ‘snowball Earth’ state in the Neoproterozoic and answer the question: ‘How low does  $\text{CO}_2$  have to be to trigger a ‘snowball’?’ you are going to edit the experiment file that controls the specific details of the experiment – the *user-config* file. From the `genie-userconfigs/LABS` directory, open one of the snowball experiments in your preferred text editor.

Near the top of the file you should see something like:

```
#  
# --- CLIMATE -----  
#  
...  
# scaling for atmospheric CO2 radiative forcing, relative to 278 ppm  
ea_radfor_scl_co2=20.0  
...
```

Each line that is not commented out (i.e., no #) contains a *parameter* name and assigned value pair, with the format:

`PARAMETER=VALUE`

The value of each parameter can be edited to change the experiment. (Additional parameter value specifications can also be added, or existing ones deleted.) In this example, the line:

`ea_radfor_scl_co2=20.0`

specifies a radiative forcing of climate by  $\text{CO}_2$  equivalent to x20 modern ( $20 \times 278 = 2560$  ppm). If you instead wrote:

`ea_radfor_scl_co2=1.0`

this would give you a modern<sup>1</sup> (x1, or  $1 \times 278 = 278$  ppm) radiative forcing. Note:  $\text{CO}_2$  is not being explicitly modeled in this experiment<sup>2</sup>, but the long-wave radiative forcing associated with a specified concentration of  $\text{CO}_2$  (in ratio to modern) is being set instead.

<sup>1</sup>Technically: the pre-industrial  $\text{CO}_2$  value rather than ‘modern’ *per se*.

<sup>2</sup>And hence the parameter:

```
# set no CO2 climate feedback  
ea_36=n
```

which tell cookie to ignore the concentration of  $\text{CO}_2$  in the atmosphere, in case, as here, there isn't one defined ...

Edit the value of `ea_radfor_scl_co2` (lower or higher – your choice) and save the file (with a different filename!). Re-run the experiment until sea-ice extent starts to approach a new steady state. You may want to try longer simulations than suggested (>100 years) if it becomes clear that the model is still far from steady-state by the end of the experiment. You can judge how close to equilibrium things have got by following (and/or plotting) the evolution of e.g., global surface air temperature or sea-ice extent (both time-series files).

---

In terms of methodology – or; 'how am I going to answer the question' – you will need to run multiple different model experiments, each with a different value for the radiative forcing parameter, and find out for what value of the parameter, the Earth becomes completely ice-covered. It is your choice whether you change the radiative forcing parameter value, run the experiment, but wait for it to finish before deciding the radiative forcing parameter value to try for the next experiment – i.e. a sequential approach, or try running multiple different 'guesses' simultaneously by submitting multiple different experiments to the queue. Ideally, each experiment will have a different name.

In each experiment, you want to be assessing how far towards the Equator the sea-ice limit encroaches by viewing some of the *time-series* and *time-slice* files or even the on-screen summary feedback (assuming that you are running interactively rather than via a job submission to the cluster queue). Informative *time-series* variables include (but are not necessarily be limited to): atmospheric temperature and sea-ice cover. (Sea-ice thickness, on account of the simple physics in the model, low resolution and long time-step, can fluctuate a little. This is also true for sea-ice area. Sea-ice volume is then the most reliable data (column) to keep track of in the *time-series* output.)

For the *time-slice* data: atmospheric and ocean surface temperature and particularly sea-ice extent (fractional cover), (2-D biogeochemical **NetCDF** file) may be informative.

**HINT:** Be careful with the 'Fit to data' scaling feature in **Panoply** – at near complete sea-ice cover, you may find Panoply scaling min and max sea-ice between 99.1 and 99.9% or something. Specific fixed scale limits (e.g. 0 and 100) can be set instead.

In answering the question ('How low does  $CO_2$  have to be to trigger a 'snowball'?'), think about what an appropriate degree of precision (rather than accuracy) might be for your experiments. Just because computer models generally calculate to around 16 significant places of precision, does not mean you have 16 significant figures of realism. For instance – how many significant figures is the solar constant quoted to and what do you think is the uncertainty in this? Harder to judge is how the assumed (incorrect) continental configuration creates additional uncertainty, or the simple physics assumed in the ocean or sea-ice, or lack of snow on land ...

Other questions to think about with regards to numerical modeling are:

- (Is the model configuration and experimental design 'realistic' ... ?)
- What is 'missing' in the model and what might the implications for your predictions and conclusions be? For example, there is no land-surface scheme (and hence no concept of 'snow') in this particular configuration.
- Are the simulations being run for sufficiently long? Why not if not (i.e., justify your choices of parameter values and experimental assumptions)? How might the results and conclusions be biased (if at all)?
- How would you test model predictions and your overall conclusions?
- How could the experimental design be improved?

Associated with the question of 'precision' is: '*How long do I have to run the experiment for?*'

This also has a vague-to-no answer. It depends on the science question and what you might judge to be appropriate precision in the context of the various uncertainties ... In other words – if you can define a sufficient or minimum precision of some property of the system, you only need to run the model as long as needed to achieve this. However, to determine the precision of the model at any point in time, you need to know the final or equilibrium value. So ideally you would run the model once, for much longer than you think you need to, and then determine how as a function of time, precision increases and error between current state and final steady state decreases.

Here I am crudely equating the concept precision with how far the model is from steady-state. You could think about it in terms of accuracy if you assumed that the steady-state of the model was 'perfect' and deviations from that steady-state (i.e. incomplete spin-up) equates to model 'error'.

Once you are happy with determining snowball threshold, try and answer the associated question:

*How high does the (CO<sub>2</sub>) radiative forcing have to be in order to escape from a snowball?*

Having determined the appropriate radiative forcing value required to create a snowball state, you can use that particular experiment as a *re-start*, and hence carry out a series of experiments with increasing radiative forcing, all starting from the same within-snowball climate state you have just created. Defining the radiative forcing / climate path going out of a snowball would complete the hysteresis loop of *Hyde et al.* [2000].

Note that a good *re-start* is one for which the experiment did not sit too long in the snowball state before finishing (the more sea-ice thickness you create in the first experiment, the more you are going to have to melt in the next and hence the longer this new series of experiments will take to ruin...). To achieve this, you can fine-tune the number of years the experiment is run, i.e. having determined the appropriate radiative forcing value required to create a snowball state, find out when (in terms of number of years) in the experiment the snowball state first occurred, and then run a new experiment that finishes only a decade or so after the snowball state was initiated.<sup>3</sup>

HINT: If you are having trouble deciding whether or not the snowball is heading in the right direction (i.e. towards an exit!), e.g. because sea-ice is always reported at 100% (or close to), you can keep track of whether there is net melting or net freezing by following mean sea-ice thickness (m) as reported in the biogem\_series\_misc\_seaice.res time-series output file. Your indication of a melting snowball state is a progressive decline in mean thickness (a proxy for global ice volume). Note that you can open and review the results of *time-series* files at \*any\* time during the experiment as the lines are written while the data for each time point is generated.

Overall: think critically about the model configuration, the experimental design, and the nature of the scientific question (based on your background reading of snowball Earth). Some of the exploration/testing suggestions (above) may not necessarily give substantially different results. Such a finding would be as valid and interesting as determining an important dependence of a certain assumption, and would for instance indicate that the associated paleo uncertainties are not critical to model assessment of the question.

Always be prepared to justify all your choices for experimental design and model settings, e.g., range of radiative forcing assessed, continental configuration(s), solar forcing, use of re-starts (if any), run duration, etc. etc. etc.

<sup>3</sup>You cannot select when the *re-start* is saved – it is always saved at the end of an experiment.

## 1.2 Further ideas

### 1.2.1 Feedback loop analysis

To quantify the snowball Earth hysteresis loop in **cookie** as per Figure 2 in *Hyde et al.* [2000] you will need to extract from the model ‘meaningful’ measures of climate (e.g., global surface air temperature, fractional sea-ice coverage) as a function of  $CO_2$  multiples,  $CO_2$  concentration, or (better) radiative forcing. For the latter, in **cookie**, the radiative forcing for a doubling of  $CO_2$  is set at:  $5.77Wm^{-2}$ . See: *Myhre et al.* [1998] (*Geophys. Res. Lett.* 25, 2715–2718) and/or *IPCC* [2007] for more on what radiative forcing is and how it is related to a relative change in  $CO_2$  concentration. Also, for making a comparison with *Hyde et al.* [2000] – for going into the snowball, note that they plot the change in radiative with a ‘cooling’ as positive (a bit daft). Their baseline radiative forcing state (an anomaly of  $0Wm^{-2}$ ) you might assume is equivalent to 278 ppm and hence  $\sim 130$  ppm is an approximately halving of  $CO_2$  and hence creates  $\sim 5Wm^{-2}$  of cooling. (You might prefer to plot the radiative forcing change as warming being positive, which makes rather more sense ...)

For coming out of a snowball, because the  $CO_2$  and hence radiative forcing threshold is so high as compared to going in, you may want to be creative in the plotting (assuming attempting to combine both thresholds into a single plot) and, for instance, one might break the scale between the low radiative forcing interval spanning going in and the high one spanning coming out.

Another example is as per Figures 3 and 4 in *Stone and Yao* [2004] (*Clim. Dyn.* 22, 815–822) (although here it is the solar constant rather than long-wave radiation forcing that is being varied). So in fact, you could try varying the solar constant as an alternative to radiative forcing and hence be able to come up with a plot directly comparable to *Stone and Yao* [2004].

### 1.2.2 Continental configuration (1)

It was mentioned earlier that the position of the continents is an area of modeling uncertainty and might be important. You can test for this. Four alternative *base-configs* are provided, each defining a different continental configuration (all with an ocean resolution of 36x36 with 8 vertical levels):

1. `cookie.C.p_worpl1.NONE`  
– a single polar super-continent. (Note potential ‘l’ and ‘l’ confusion in ‘p\_worpl1’.)
2. `cookie.C.p_worpl2.NONE`  
– one continent at each pole.
3. `cookie.C.p_woreq1.NONE`  
– a single Equatorially-centred super-continent. [this is the current configuration (that you used previously)]
4. `cookie.C.p_woreq2.NONE`  
– two continents straddling the Equator.

You can use the given *user-config* file (LAB.1.1.EXAMPLE) again as an experiment template file, and any of the alternative configurations can be run very similarly to as per before, i.e.:

```
$ ./runcookie.sh cookie.C.p_xxxxxx.NONE LABS LAB.1.1.EXAMPLE 100
```

Note that you are using a different *base-config* file name: `cookie.C.p_xxxxxx.NONE` where `xxxxx` is one of: `worpl1`, `worpl2`, `woreq1`, or `woreq2`.

Also note that no re-starts are provided for any of these configurations. You may (or may not) want to create some (you will need to judge for yourselves how long to run the *re-start* experiments to achieve as close to steady-state as you think is ‘sufficient’). Recall again, that *re-starts* are just ‘normal’ experiments that have already been run. Be careful that when changing from one *base-config*

to another, the model re-compiles. Simply running the new configuration briefly (for even just a single year) is sufficient to ensure this. Experiments can then be safely submitted to a cluster queue, i.e. do not try and submit an experiment using a different *base-config* straight to the cluster queue without having run it (or a short version of the experiment you want) interactively first (to ensure the model is re-compiled). This is also good practice – checking that a new sort of experiment and/or model configuration works as you intend and without hiccups.

### 1.2.3 Continental configuration (2)

Although much useful can be learned from conceptual configurations and Worlds regarding climate dynamics, it is invariably aesthetically more ‘pleasing’ to also test ideas in a more paleogeographically realistic configuration. Provided is a set of *base-config* and *user-config* files (plus associated boundary conditions) for the position of the continents and climate 635 millions years ago (635 Ma). For this:

The *base-config* is named: cookie.C.fm0635cb.NONE

The *user-config* is: cookie.C.fm0635cb.NONE.SPIN

NOTE that the *user-config* lives in the directory: ~\cgenie.cookie\user-configs\EXAMPLES so that you either need to specify this different (EXAMPLES) directory when running **cookie**, or copy the *user-config* file into LABS.

A *re-start* experiment is provided called cookie.CB.fm0635cb.NONE.SPIN and which can be downloaded as per before:

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/cookie.C.fm0635cb.NONE.SPIN.tar.gz
```

and unpack by:

```
$ tar xfzv cookie.C.fm0635cb.NONE.SPIN.tar.gz
```

(Remember: you should be in the cgenie\_output directory when you do this downloading and unpacking.)

To run (e.g. for 100 years), following on from its *re-start* (and copying (/renaming) the *user-config* to the LABS directory):

```
$ ./runcookie.sh cookie.C.fm0635cb.NONE LABS cookie.C.fm0635cb.NONE.SPIN 100
cookie.C.fm0635cb.NONE.SPIN
```

(all on one line)<sup>4</sup>

### 1.2.4 Geothermal heat input

Finally, **cookie** will fairly happily build up sea-ice, apparently without limit (with the remaining wet ocean becoming progressively colder and more saline). In the real world, one might expect some sort of limit to the maximum thickness achieved as the heat diffusion across a progressively greater thickness of sea-ice approaches the heat input at the bottom of the ocean from geothermal energy. Different modes of ocean circulation are also possible if one considers heating from the bottom as well as cooling (and brine rejection) from the top and which might affect the entry into or exit from a snowball state.

---

<sup>4</sup>NOTE that the *base-config* and *user-config* filenames start cookie.C. whereas the *re-start* filename starts cookie.CB. ... just to try and trip you up ...

In the experimental setup you have been given, a geothermal heat input is specified in the ocean circulation module via the following:

```
bg_par_Fgeothermal=100.0E-3
```

The parameter `bg_par_Fgeothermal` sets the geothermal flux in units of  $\text{W m}^{-2}$ . (Note that in the Neoproterozoic, the geothermal heat flux could have been somewhat higher than modern. How much higher? A question for **Google** ... ?)

An appropriate research question might be to determine in radiative forcing *vs.* geothermal space (and requiring a 2D grid of parameter combinations to be created and submitted to the cluster), the equilibrium sea-ice thickness and region in which a snowball solution is not possible. However, more simply and suitable to a short exercise: How much of a difference, to the estimated snowball entry and exit thresholds of radiative forcing, does the inclusion of a geothermal input make? E.g., what happens if you set it to zero? What about 10 times modern (or more, although \*extreme\* seafloor heating can cause numerical instability and the model to crash)?

### 1.2.5 Seasonality

By default, the idealized model configurations are non-seasonally forced (by solar insolation). You can switch to a seasonally-forced to model by adding the following lines to the *user-config* file:

```
ea_dosc=.true.  
go_dosc=.true.  
gs_dosc=.true.
```

The scientific question here in trying this would be whether or not taking into account a seasonally-varying climate substantially affects the entry (and/or exit) thresholds for a snowball climate state. (At least, whether it is important in the context of the resolution and physics of the model you are using.)

You can also save the data seasonally if you like – see Section 12.2.3 in the **cookie** User Manual (this document!).<sup>5</sup>

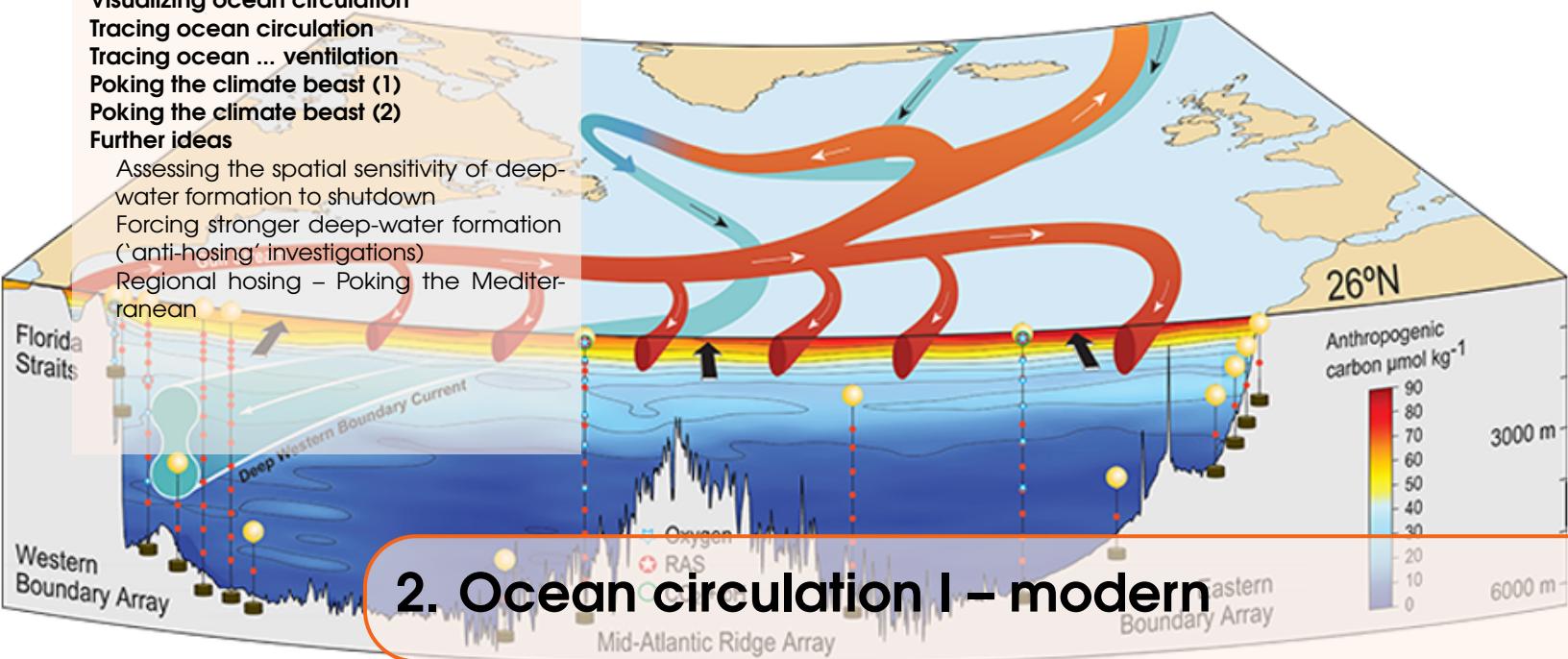
---

<sup>5</sup>For reference, your configuration has 24 time-steps per year set for the **BIOGEM** module.



Simulating ocean circulation  
Visualizing ocean circulation  
Tracing ocean circulation  
Tracing ocean ... ventilation  
Poking the climate beast (1)  
Poking the climate beast (2)  
Further ideas

Assessing the spatial sensitivity of deep-water formation to shutdown  
Forcing stronger deep-water formation  
(‘anti-hosing’ investigations)  
Regional hosing – Poking the Mediterranean



## 2. Ocean circulation I – modern

Stuff to keep in mind:

- Nothing at all – keep your mind completely empty and let the wonderful truths of **muffin** permeate your entire being.

Background reading (Atlantic circulation and stability in **muffin**):

- Hargreaves et al. [2004] (Climate Dynamics 23, 2004, Pages 745 – 760)  
→Simple assessment of the likelihood of AMOC collapse.
- Marsh et al. [2004] (Climate Dynamics, 23 2004, Pages 761 – 777)  
→Characterization of thresholds of AMOC collapse.
- Singaray et al. [2008] (GRL 35, doi:10.1029/2008GL034074)  
→Role of changing ocean circulation in atmospheric radiocarbon variability during the Younger Dryas.

Background reading (Miscellaneous (model) Atlantic circulation and stability):

- IPCC [2007] (e.g., Section 10.3.4)  
→Future predictions of AMOC strength.
- Schmittner [2005] (Nature 434, 628– 633)  
→Impacts on marine ecosystems and carbon cycling.
- Obata [2007] (J. Clim. 20, 5962–5976)  
→Climate-carbon cycle model response to freshwater discharge.

---

**READ.ME**

You will need to download new *re-start* files prior to embarking on the experiments with modern ocean circulation. To fetch them: change to the cgenie\_output directory, and type (or copy and paste carefully from the PDF ...):

```
$ wget --no-check-certificate http://www.seao2.info/cgenie_output/
  cookie.C.p_worjh2.r.SPIN.tar.gz
$ wget --no-check-certificate http://www.seao2.info/cgenie_output/
  cookie.C.p_worjh2.rb.SPIN.tar.gz
```

(single line ... no space between line fragments ...) This downloads archived/compressed copies of the 10,000 year *spin-up* experiments: cookie.C.p\_worjh2.r.SPIN and cookie.C.p\_worjh2.rb.SPIN. Extract the contents by typing:

```
$ tar xfzv cookie.C.p_worjh2.r.SPIN.tar.gz
$ tar xfzv cookie.C.p_worjh2.rb.SPIN.tar.gz
```

You will then need to change directory back to genie-main to run the model.

## 2.1 Simulating ocean circulation

### Important.

1. Note that the configurations provided for this exercise do not have a carbon cycle, but do include a numerical 'age' tracer to help diagnose ocean circulation and rate of ventilation of the deep ocean.
  2. Make use of *control* experiments in order to create anomaly maps (see later). An ideal control experiment is one that is run for as long as your 'real' (perturbation) experiments, but has no changes made to the *user-config* (compared to the *re-start*).
- 

To start off – run a *control* experiment using the *re-start* experiment that you have just downloaded as the initial state of your experiment. A *control* experiment is basically one in which you do not change any parameter settings and is used top quantify any residual drift (that you might otherwise interpret as the result of a parameter change in your actual experiment). The template *user-config* file you are provided for, for this, is: LABS.2.1.EXAMPLE

Run the model for ... whatever, 10 years will do for now::

```
$ ./runcookie.sh cookie.C.p_worjh2.r LABS LABS.2.1.EXAMPLE 10  
cookie.C.p_worjh2.r.SPIN
```

where cookie.C.p\_worjh2.r.SPIN is the *re-start* experiment name.

You will analyze the results of this in the next couple of Sections.

NOTE: For the initial steps, you will be using a seasonally-forced modern continental configuration of cGENIE with 16 vertical levels in the ocean. This is indicated by the code 'p\_worjh2' in the *base-config* file-name. The grid and definition of the vertical levels are illustrated in Figure 2.1) and (Figure 2.2), respectively.

---

36	10	09	08	09	10	11	11	10	08	07	07	07	07	07	07	07	11	13	14	14	12	12	93	93	91	91	09	05	05	08	10	11	11	11	12	12	11	11
35	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94
34	94	94	94	94	94	92	92	08	07	10	15	12	09	93	93	94	94	94	91	91	91	06	05	05	06	09	12	93	93	94	94	94	94	94	94	94	94	
33	94	91	91	91	16	09	04	03	04	05	03	03	06	93	93	91	91	91	91	91	09	05	03	04	06	93	93	93	94	94	94	94	94	94	94	94		
32	94	91	91	91	10	04	01	01	01	01	01	01	04	11	93	91	91	91	91	91	12	05	03	03	03	11	93	93	94	94	94	94	94	94	94	94		
31	94	91	91	11	05	01	01	01	01	01	01	01	03	08	93	93	91	91	91	91	12	06	03	03	02	09	93	93	94	94	94	94	94	94	94	94		
30	91	91	91	07	02	01	01	01	01	01	01	01	02	05	93	91	91	91	91	91	09	03	01	01	04	02	93	10	94	94	94	94	94	94	94	94		
29	91	91	91	08	02	01	01	02	01	01	01	01	01	03	93	93	92	92	92	92	06	01	01	02	04	03	02	12	10	08	08	11	94	94	94	94	94	94
28	91	91	16	05	01	01	01	01	01	01	01	01	01	10	93	92	92	92	92	92	05	01	01	03	01	03	01	93	93	12	11	94	94	94	94	92	92	
27	91	91	09	03	01	01	01	01	01	01	01	01	02	02	07	91	13	09	04	01	01	02	01	01	05	93	93	94	94	94	94	92	92	92	92			
26	91	91	05	02	02	01	01	01	02	01	01	01	02	05	91	08	07	04	01	01	02	01	01	09	93	93	94	94	94	91	16	09	92	92	92			
25	91	08	02	01	02	01	01	02	02	01	01	01	02	04	91	09	08	05	02	01	02	01	12	93	93	94	94	94	91	14	06	92	92	92				
24	93	06	02	01	02	01	01	02	01	01	01	01	02	03	06	91	10	06	03	01	03	01	02	12	93	93	94	94	91	09	03	02	08	11				
23	91	05	03	01	02	01	01	01	01	01	01	01	01	02	02	04	06	91	05	04	01	02	01	10	93	93	94	94	94	91	05	02	11	05	09			
22	93	06	04	01	02	01	01	01	01	01	01	01	01	02	02	03	04	07	93	08	03	01	01	07	93	93	94	94	94	91	03	02	09	04	08			
21	91	07	05	02	03	03	02	01	01	01	01	01	02	03	03	03	07	91	91	07	02	03	02	05	93	93	93	93	91	91	03	02	06	03	08			
20	93	09	04	02	02	03	02	01	01	02	02	02	02	03	03	03	05	91	91	11	04	02	02	03	05	08	93	93	91	11	02	02	04	03	07			
19	91	91	04	03	03	03	02	01	01	01	02	02	02	02	04	04	05	91	91	91	05	03	02	02	04	04	93	93	91	07	02	02	03	02	04			
18	93	91	07	08	07	05	03	01	01	01	02	02	02	02	03	04	05	91	91	91	08	04	01	02	03	03	93	93	91	04	02	02	02	01	02			
17	09	91	07	10	10	07	04	01	01	01	01	02	02	02	03	03	03	91	91	91	91	05	01	02	02	02	93	93	91	03	03	04	02	01	01			
16	04	93	07	14	10	05	03	01	01	02	01	01	02	02	03	03	02	02	91	91	91	91	05	01	03	02	01	93	91	91	03	03	04	01	01	01		
15	01	03	92	94	10	04	03	03	02	01	02	02	03	03	02	02	91	91	91	91	04	01	03	02	01	93	91	91	04	03	04	01	02	01				
14	01	02	93	91	09	04	04	04	04	01	02	03	03	03	03	02	02	91	91	91	91	03	01	03	01	01	93	91	91	07	03	03	02	02	01			
13	01	04	93	91	13	06	04	04	04	01	02	03	03	03	02	02	04	93	91	91	03	01	03	01	01	93	93	91	09	03	03	01	01	01				
12	01	08	93	91	91	08	06	04	03	01	02	02	03	04	02	02	02	93	91	91	03	01	03	01	02	93	93	10	08	02	03	02	01	01				
11	01	10	93	93	91	07	05	03	02	01	02	02	03	04	03	03	03	93	91	91	09	02	01	02	10	93	07	05	01	02	02	01	02	01				
10	01	09	93	92	91	05	06	03	01	01	02	02	03	04	03	03	03	93	91	06	02	01	03	02	02	07	93	05	04	01	02	03	03	02				
09	02	08	92	91	04	06	04	01	01	01	02	02	03	03	03	03	03	93	91	04	02	03	02	02	06	93	04	03	02	02	03	04	03	02				
08	01	03	03	05	92	03	05	07	01	01	01	02	03	03	03	03	04	93	09	02	02	03	02	01	03	05	02	03	02	03	02	03	02					
07	07	02	01	01	02	07	02	04	10	02	01	01	01	02	03	03	03	06	93	05	01	01	02	04	02	01	01	02	02	04	02	03	03	03				
06	03	02	01	01	04	02	04	12	03	02	01	01	01	02	04	03	02	03	06	93	03	01	01	02	03	03	02	01	01	03	05	02	02	04	03	03		
05	03	03	02	02	04	02	05	07	02	01	01	02	03	04	02	02	08	93	04	01	01	03	04	03	02	01	02	04	03	04	06	03	03					
04	02	02	02	03	03	04	03	01	01	02	03	03	03	02	01	02	05	05	04	03	02	02	04	04	03	02	02	01	03	05	03	02						
03	02	01	01	02	03	03	03	01	01	02	03	03	02	02	02	01	01	02	04	04	03	02	02	01	01	01	01	01	01	02	04	03	03					
02	94	10	09	09	08	07	06	03	04	03	03	02	02	02	02	03	05	09	07	04	02	01	01	02	03	03	04	06	06	06	08	08	04					
01	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94				

Figure 2.1: The cookie grid for a modern 36 × 36 ‘p\_worjh2’ configuration. Light blue numbers are the ‘i’ co-ordinates. Green numbers are the ‘j’ co-ordinates. The depth of the ocean at any location is indicated by its ‘k’ value – a number between 1 and 16, with 16 being the surface layer of the ocean, and 1 the maximum possible depth anywhere. Numbers > 90 (i.e., 91, 92, 93, 94) and shaded grey are land (the numbers specify the compass direction of rainfall run-off). The longitude of the western edge of this particular modern ocean grid is at 260°W, and the increments are 10°.

k	mid depth (m)	base of layer (m)
16.00	38.91	60.84
15.00	126.04	174.75
14.00	227.26	283.85
13.00	344.84	410.58
12.00	481.44	557.80
11.00	640.12	728.83
10.00	824.45	927.51
9.00	1038.59	1158.31
8.00	1287.35	1426.43
7.00	1576.33	1737.90
6.00	1912.04	2099.73
5.00	2302.02	2520.05
4.00	2755.05	3008.34
3.00	3281.33	3575.57
2.00	3892.71	4234.52
1.00	4602.92	5000.00

Figure 2.2: The cookie ocean vertical level definitions for a modern 16-level ocean grid.

## 2.2 Visualizing ocean circulation

Visualizing the 3D flow (/transport) of the ocean, much less the rate at which this occurs, is no trivial matter. Even with the aid of a model. (Or rather, the problem then becomes: how to visualize ocean circulation in a model.) We'll consider two different ways of analysing model velocity fields first – simply utilizing the *restart* that you have downloaded and unpacked, and then later take a more pro-active approach in subsequent Sections with some new experiments.

The first approach we can take is to simply visualize the raw velocity fields, but plotted as ocean currents.

1. In the 3D **netCDF** file, the three components of ocean velocity are represented by the variables: ocean velocity – u (Eastwards), ocean velocity – v (Northwards), and ocean velocity – w (upwards). 2. Open up velocity – u. Choose ‘lon-lat’.
2. Select/highlight velocity – v. and click on the ‘Combine Plot’ icon (as per before).
3. Rather than a difference map, which is what you get by default, i.e., ‘Array 1 – Array 2’ – from the drop-down menu (next to the ‘Interpolate’ button) select ‘Vector Magnitude’.
4. You should have a color contoured (or not if you prefer plotting without contouring on) map of ocean current speed, with velocity vectors (direction and magnitude) overlain. You'll need to re-scale the velocity vectors to properly see them – from the ‘Contours and Vectors’ tab – change the ‘Scale Length’ to e.g., 0.1. (On a Mac, look under the ‘Vectors’ tab for a ‘Reference Value’ to something like 0.1.) When fresh-water hosing – look out for impacts on the N. Atlantic current system associated with the AMOC.
5. You can repeat this for deeper depth levels in the ocean – e.g., between about 1500 and 2000 m is a good place to go looking for the Western boundary current (and AMOC return) in the model (such as it exists at this low resolution) but you'll need to re-scale the velocity vectors again (e.g., to 0.01 to less).

An example plot (using **Panoply** for visualizing surface ocean current fields) is given in Figure 2.3.

The second approach is to visualize the large-scale ocean transport in terms of the meridional overturning circulation (‘stream-function’) (e.g. see background literature).

Two example plots (using **Panoply**) are shown in Figure 2.4 for the Atlantic basin, and Figure 2.5 for the Pacific. In the In the 2D **netCDF** file, relevant fields (netCDF variables)<sup>1</sup> are:

```
phys_opsia == global overturning stream-function
phys_opsip == overturning in the Atlantic
```

<sup>1</sup>Note that these fields are only meaningful for the modern arrangement of the continents and a continuous separation of the Atlantic from the Pacific from high northern latitudes down to the tip of South America. Different (e.g. paleo) arrangements of the continents may not have recognisable (or definable) Atlantic and/or Pacific basins and it may only be possible to define and visualize the global meridional overturning circulation – variable: phys\_opsi.

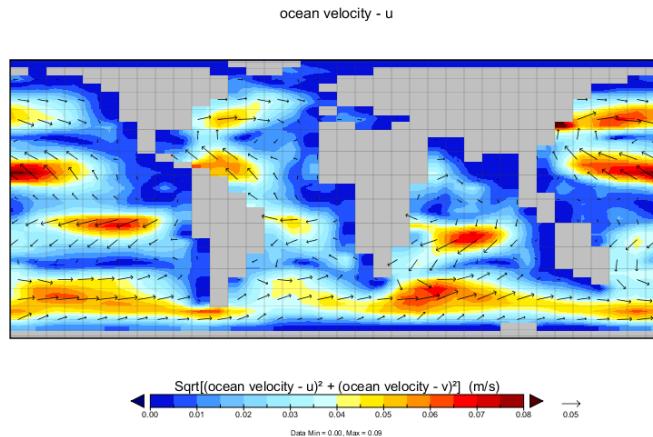


Figure 2.3: Example plot of (normal/default modern) ocean current fields (3D netCDF file). Again scaling has been set manually to create an easy-to-interpret axis scale. On the left is the surface field, and on the right an intermediate depth (illustrating what approximates the Deep Western Boundary current in the model in the Atlantic).

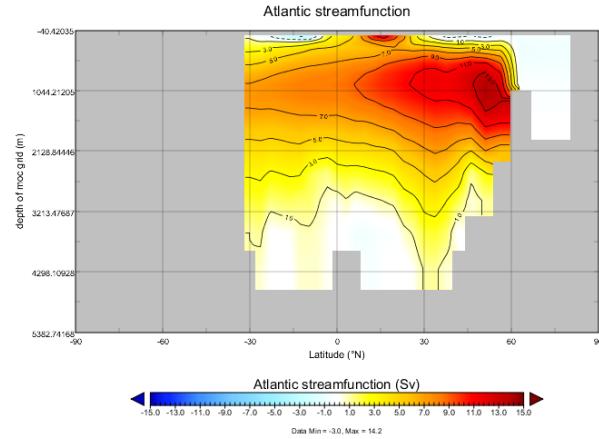


Figure 2.4: Example plot of (normal/default modern) overturning stream-function (2D netCDF file). (e.g., for Atlantic: **netCDF** parameter name: `phys_opsia`, long-name: Atlantic stream-function). Note that auto-scaling has been turned off and the min and max plotting limits set manually. By convention, stream-functions are plotted with their scale symmetrical around zero, giving red and ‘warm’ colors for positive value and clockwise overturning, and blues and ‘cold’ colors for negative values and anti-clockwise overturning. (The plot has been tart-ed up by overlaying solid contours plus contour labels.) It may be necessary in **Panoply** to re-orient (invert) the vertical grid.

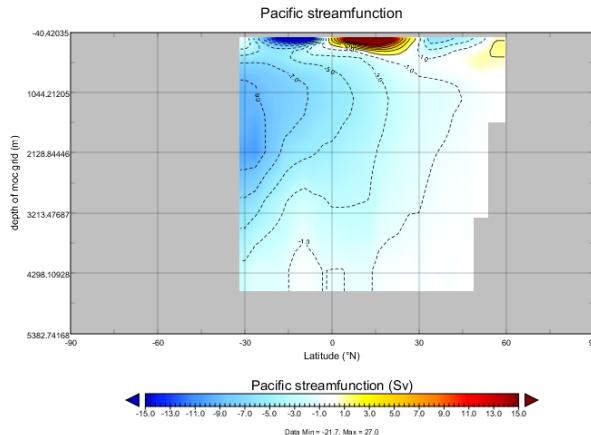


Figure 2.5: Pacific meridional overturning circulation (PMOC).

## 2.3 Tracing ocean circulation

The ocean biogeochemistry module (**BIOGEM**) in **cookie** provides a framework for applying time- and spatially-variable ‘forcings’ of the Earth system<sup>2</sup> – fluxes or ‘restored-to’ boundary conditions that can be prescribed for any gas, dissolved substance (including temperature and salinity), or particulate matter. Examples include freshwater input (== a negative salinity flux forcing) of the North Atlantic to alter ocean circulation, fossil fuel  $CO_2$  emissions to the atmosphere (== a  $CO_2$  gas flux forcing), or aeolian iron supply to the surface ocean (a 2-D dust flux forcing).

For example: view the *user-config* file: LAB.2.3.colortracer – you will see the following lines (under the heading: '# -- FORCINGS --')

```
bg_par_forcing_name="pyyyyz_Fblue"
bg_par_force_point_i=22
bg_par_force_point_j=33
bg_par_force_point_k=8
bg_par_ocn_force_scale_val_49=1.0E12
```

The first line points **cookie** to a directory located in cgenie.cookie/genie-forcings that contains a set of files that define what geochemical property is going to be altered plus information about how the magnitude of the forcing changes with time.

There are then three lines (bg\_par\_force\_point\_i=20, ...) that specify the location in the ocean of the geochemical forcing that is going to be applied. The point sources are specified in (i,j,k) coordinates, which in this case is (22,33,08). For the ocean model resolution we are using, the grid is 36x36x16, and in which: longitude (i) is counted from left-to-right (1 to 36); latitude (j) is counted from bottom-to-top (1 to 36); level depth (k) is counted from downwards top-to-bottom (16 down to 1). Thus, (22,33,08) is a release of tracer in the North Atlantic, a little south of Greenland, and intermediate depth (level = 8 out of 16). Refer to the Figures for how the horizontal (Figure 2.1) and vertical (Figure 2.2) grid is specified.

Finally, there is a scaling parameter (bg\_par\_ocn\_force\_scale\_val\_49) which modifies the magnitude of the flux to be applied<sup>3</sup> – the default value in the forcing definition itself is zero  $0.0\text{mol}\text{yr}^{-1}$  (so you will need to set something here!).

<sup>2</sup>Refer to the ‘force the system’ HOW-TO in the cookie manual for further details on *forcings*.

<sup>3</sup>Flux *forcings* in **cookie** are in units of  $\text{mol}\text{yr}^{-1}$  per model grid point.

You are going to run a brief experiment in which you will be injecting a conservative ‘dye’ tracer into the ocean. The **BIOGEM** module has two tracers that can be defined for this purpose – ‘blue’ and ‘red’ – you will be using the blue one here. You can control the flux of blue dye by opening the *user-config* file: LAB.2.2.colortracer and editing the flux scaling parameter:

```
bg_par_ocn_force_scale_val_49=1.0E12
```

The *base-config* you will be using is different from previously: cookie.C.p\_worjh2.rb – this specifies a 16 vertical levels ocean and also includes seasonality of solar insolation.<sup>4</sup>

Run the model for ... whatever, 20 years will do. Use the *re-start* experiment that you have just downloaded to start from:

```
$ ./runcookie.sh cookie.C.p_worjh2.rb LABS LAB.2.3.colortracer 20  
cookie.C.p_worjh2.rb.SPIN
```

View the results – for instance how the blue tracer distribution evolves with time – in the *time-slice* files (full ocean/atmosphere) properties saved in the **netCDF** format (.nc) files). You can follow the progress of the dye (and hence diagnose the properties of ocean circulation in the model) by plotting vertical and/or horizontal slices that go through (or near) the cell location in which you inject the dye tracer in the 3D **netCDF** file. Note that **Panoply** appears to ‘count’ the ocean layers in the opposite direction to the way in which the ocean model is actually counting them – the correct definition is with ‘1’ being very deepest level possible (and as displayed in the figure) and ‘16’ is the surface.

You can also view the tracer distributions in terms of a water-column integrated tracer inventory (**netCDF** variable name: ocn\_int\_colb; long name: colb water-column integrated tracer inventory) in the 2D **netCDF** output. (See: Sabine *et al.* [2004] for the use of water column integrals in the context of the distribution of anthropogenic CO<sub>2</sub> uptake and storage.) Changes in tracer inventory with time can be tracked in the time-series file: biogem\_series\_ocn\_colb.res

Spend a little while altering the flux (bg\_par\_ocn\_force\_scale\_val\_49) and/or location (bg\_par\_force\_point\_i, bg\_par\_force\_point\_j, bg\_par\_force\_point\_k) of tracer input. Overall – note how you can use numerical ‘tracers’ to help diagnose (and better understand) the circulation of the ocean.

Ignore the ‘red’ tracer for now ... we’ll look at that shortly.

<sup>4</sup>Note that because the *base-config* is different from that used in the previous chapter, you need to force a re-compile of the model code before any experiments can be submitted as cluster jobs. The easiest way to do this is to run an experiment interactively at the command line.

When you add the numerical dye, particularly early on in the experiment, you may see a 'front' of negative tracer concentrations leading the (positive) tracer as it spreads. DON'T PANIC!

The model is ... a model (of the numerical flavor) and not an exact analytical solution. So errors in how it solves ocean transport are to be expected.

Moreover, by default the ocean circulation model employs an isopycnal/diapycnal mixing scheme. This can lead to unwanted negative tracer values when sharp horizontal (or vertical) transitions in concentration occur. (In this example, e.g. by injection dye at a point location into a surrounding ocean of initially zero concentration.)

You can change to a simpler horizontal/vertical (.false.) mixing scheme by adding to the *user-config* file:

```
# turn OFF (=false.) isopycnal/diapycnal mixing
go_diso=.false.
```

If you try this, you should (hopefully) find much less (or no) negative tracer concentrations occur. However, also note that by changing the physics of ocean mixing, you also slightly alter the large-scale circulation of the ocean (and e.g. the AMOC might change slightly in strength).<sup>5</sup>

Lastly, an interesting (honest!) and illustrative exercise is to use the dye tracer to pick out the path taken by Mediterranean Intermediate Water. Despite the low resolution of the **cookie** ocean circulation model component and the highly restricted representation of the Mediterranean, the model does project a salty Mediterranean as a consequence of P-E across this basin (and its catchments) being negative and this higher density water makes its way out in the subsurface into the Atlantic.

To do this – simply specify a dye injection somewhere in the Mediterranean (be careful with the restricted depth of the Mediterranean – if you inject too deeply (into the crust!) then you will not see anything (refer to the figure for the depth level (k) number of the maximum depth of the water column in each location), and it is better to inject it relatively close to the opening of the gateway (try some different locations and see which ones produce a reasonably instructive tracing of Mediterranean outflow). Run for e.g., 20 or 50 years (from the provided spin-up). Then:

1. View the dye-tagged plume of Mediterranean Intermediate Water by plotting a lat-lon slice (from the 3D **netCDF** file). This will give you the depth of the plume. How does this compare with salinity observations (salinity observations and appropriate global data-sets can be found on the web with a little patience)? You can also view the water-column integrated distribution (2D **netCDF**).
2. Try viewing the plume via a lat-depth slice. Refer to the figure to determine the 'i' value up the Atlantic that will just graze the edge of what passes for Spain at this low model resolution. Which direction does it head after exiting the Mediterranean? Is this 'realistic'?

<sup>5</sup>You might contrast the overturning stream-functions for experiments run both with and without horizontal/vertical mixing.

## 2.4 Tracing ocean ... ventilation

Yet another way to think about global ocean circulation is through considering the connection (rate of mass exchange) between surface and deep ocean – ‘ventilation’.

**cookie** has the capability to employ/simulate a ventilation age tracer<sup>6</sup> – a numerical tracer carried in the ocean circulation model that tracks the time since a parcel of water last ‘saw’ the surface. The older the ‘age’ of the parcel of water, the longer the time since it last saw the surface.

We can use the second numerical tracer (red’) to keep track of age, but rather than apply a flux forcing to the surface, we let the model automatically restore the tracer value at the surface to zero and everywhere else (in the ocean interior) increase the age each time-step (by the duration of the time-step) such that a parcel of water away from the surface ages by 1 year, each year.

The *base-config*, and *restart*, provided for the (‘blue’) circulation tracing, already has the (‘red’) age tracer included and activated. As a result, all of your experiments on ocean circulation which you have conducted so far already have simulated ocean ventilation age and you do not need to run a new experiment (unless you want to!).

In the 3D netCDF output file – `misc_col_Dage` is the output variable that is the calculated ventilation age.

Explore the distribution of water mass age and think about the physical ocean circulations reasons for this. Are all the modelled distributions reasonable? Are there indicators of facets of the simulated circulation that are not particularly realistic? Try plotting both lon-lat and lat-depth slices through the ocean. How does the distribution of water mass age relate to the overturning stream-function for Atlantic or Pacific basins (more on this in the next Section)?

For the remainder of the sections in this chapter, we’ll ditch the ‘blue’ color tracer and use a configuration (and different *base-config*) with just the ‘red’ one: `cookie.C.p_worjh2.r`

<sup>6</sup>Under the ‘screw with and/or diagnose the climate system’ HOW-DO – see ‘Add a water mass age tracer’ (and the ‘easy’/automatic method described towards the end of that section).

## 2.5 Poking the climate beast (1)

The ocean biogeochemistry module (**BIOGEM**) in **cookie** provides a framework for applying time- and spatially-variable ‘forcings’ of the Earth system<sup>7</sup> – fluxes or ‘restored-to’ boundary conditions that can be prescribed for any gas, dissolved substance (including temperature and salinity), or particulate matter. Examples include freshwater input (== a negative salinity flux forcing) of the North Atlantic to alter ocean circulation, fossil fuel  $CO_2$  emissions to the atmosphere (== a  $CO_2$  gas flux forcing), or aeolian iron supply to the surface ocean (a 2-D dust flux forcing). In this exercise you are going to add fresh water to the ocean surface to assess the sensitivity of the Atlantic Meridional Overturning Circulation (AMOC) to collapse, in a classic ‘hosing’ experiment.

The *user-config* file for this is called: LAB.2.5.hosing. This includes a prescribed transport of salt to the North Atlantic (north of ca. 30°N) from the result of the global surface ocean (for mass balance). This is specified by the line:

```
bg_par_forcing_name="p_worjh2.Fsal_SUR_NAtl"
```

This is paired with a scaling parameter for ease of changing the magnitude of the flux forcing:

```
bg_par_ocn_force_scale_val_2=0.0
```

which by default is zero (i.e., no flux forcing).

To orientate you in freshwater forcing space: `bg_par_ocn_force_scale_val_2=-1.1E18` – equivalent to 1Sv – should be sufficient to make ‘stuff happen’ and quickly. BUT, this is a pretty extreme flux (see overleaf for a rough conversion between salinity forcing units ( $mol\ yr^{-1}$ ) and fresh water flux (in  $m^3\ s^{-1}$  or Sv). Much more than this and the model may crash or at the very least, you might be left with a large freshwater pond in the North Atlantic. Don’t forget the minus sign if you want a net freshwater flux to the N Atlantic rather than a salinity one ...

To run the model for 10 years using the same *re-start* as previously:

```
$ ./runcookie.sh cookie.C.p_worjh2.r LABS LAB.2.5.hosing 10
cookie.C.p_worjh2.r.SPIN
```

10 years should be long enough to see a collapse start to occur, but you might want to run the model for longer (and it can be submitted as a job, of course). Running for longer will also allow you to have a smaller, less extreme (and maybe more realistic) freshwater input flux.

Make sure that you run a control experiment – an experiment of the same duration of your hosing experiments, but with a zero freshwater flux. The impact of freshwater input, is the difference, at the same model year, between the perturbation experiment and the control. (You only ever need to run one control, regardless of how many different freshwater flux perturbation experiments you run.)

Note that as the model is running rather slow than in the snowball configuration, you might want to think carefully of making use of cluster queuing possibilities (i.e., running multiple experiments at once in the background).

---

**What to look at?** The most obvious property of the Earth system to follow is the Atlantic overturning strength (`biogem_series_misc_opsi.res`). The AMOC stream-function (in `fields_biogem_2d.nc` 2-D time-slice **netCDF** results file, field: `phys_opsi`) is also illustrative. You can also try and identify the salinity anomaly (see below) due to freshwater input in the 3D salinity tracer field.

---

<sup>7</sup>Refer to the ‘force the system’ HOW-TO in the cookie manual for further details on *forcings*.

There are also important (but not necessarily painfully exciting) impacts on surface air temperatures and maybe sea-ice extent (in `fields.biogem_2d.nc`) (but see below for a better way to visualize these changes). Note the importance (sort of) of the AMOC in transporting heat to the N Atlantic region (the film the Day After Tomorrow was not entirely inaccurate in this particular respect). Be aware of the possibility of climate impacts far from the location of fresh water forcing. Look out for any significant-looking impacts on sea-ice extent, etc.

You might plot current velocity fields and visualize how these change in response to the fresh water forcing, and if you ran much longer than 10 years, see impacts on ocean ventilation ages.

---

To more easily assess some of these impacts (and for other sorts of analysis) it is possible to create an anomaly (difference) map in **Panoply**:

1. First open a dataset, e.g., `atm_temp` (surface air temperature) in the 2D **netCDF** file. Double-click the variable name, or, with the variable name highlighted, click the ‘Create Plot’ icon.
2. Now, with the `atm_temp` still selected (and the first plot window still open), click on the ‘Combine Plot’ icon. A dialogue box will appear and ask you to select a plot to combine the new one with. Make sure the name of your first plot window is selected/highlighted. Click ‘Combine’. OR, simply drag a second dataset into the plot window of the first dataset.
3. You now have a plot window that by default it is showing you the difference between two identical (in time) slices. The two slices are labeled Array 1 (LH side) and Array 2 (RH side).

Keep one array (Array 1) fixed to the initial (year 1 (centered on 0.5)) and vary the year in the second array (Array 2). Note that you can select in Panoply whether  $\text{Array 1} - \text{Array 2}$  is plotted, or  $\text{Array 2} - \text{Array 1}$ , or various proportional or relative differences.

Note that you can switch off the auto-scaling feature (Always fit to data) and center the scale so that no change is white, with positive deviations = red and negative = blue by clicking on Center on 0 (an often used convention in climate field plotting).

---

Try and answer the question: **How much fresh water flux (in Sv) is required to collapse the AMOC?** and how non-linear is the response of overturning to fresh water flux?

You will need to submit multiple experiment jobs to the cluster queue, with each experiment differing in the assumed radiative forcing parameter value. Plot the strength of the AMOC that is reached at the end of each experiment (from the time-series file: `biogem_series_misc_opsi.res`) vs. radiative forcing.

---

A brief note on units ... the freshwater forcing is implemented as negative salinity, just to really screw with your mind. The generic internal **cookie** model units for the forcing end up as  $PSU kg^{-1} yr^{-1}$ . Which sort of does not make much sense ...

Start, by thinking of a value of `bg_par_ocn_force_scale_val_2` of  $-34.9$  as equivalent to taking all the salt out of  $1kg$  of freshwater (since the mean global salinity is  $34.9PSU$ ). Or equivalently, since the ocean volume is fixed, an applied forcing value of  $-34.9$  is equivalent to adding  $1kg$  of freshwater to a (surface) box. So, a value of `bg_par_ocn_force_scale_val_2` of  $-3.49 \times 10^4$  ( $-3.49E04$ ) would be a flux of  $1m^3 yr^{-1}$  ( $1000kg m^{-3}$ ) of freshwater.

So, in the example earlier (`bg_par_ocn_force_scale_val_2=-1.0E18`), the freshwater flux is  $1.0 \times 10^{18} / 3.49 \times 10^4 = 2.8653 \times 10^{13} m^3 yr^{-1}$ .

The literature invariably gives freshwater fluxes in units of  $Sv$  ( $10^6 m^3 s^{-1}$ ). So in the example, the freshwater flux is:  $9.0797 \times 10^5 m^3 s^{-1}$  ( $365.25 \times 24 \times 3600 = 31557600 syr^{-1}$ ). Or  $0.9Sv$ .

Or or ...  $1.0Sv$  is equivalent to a model (`bg_par_ocn_force_scale_val_2`) total forcing flux of  $-1.0E18 / 0.90797 = -1.1E18$

Read the literature ... but generally, fluxes of ca.  $0.05Sv$  and larger (and to quite specific places) are applied in models to induce an AMOC collapse.

## 2.6 Poking the climate beast (2)

A current concern regarding anthropogenic climate change is the ocean circulation (and marine ecological and biogeochemical) response to a strong warming of the surface, as rapid surface warming is assumed (and demonstrated in models) to result in surface stratification of the ocean, likely restricting the nutrient supply to phytoplankton and reducing ventilation of the ocean interior with dissolved oxygen.

You can explore the transient response of ocean circulation to warming by simply adjusting the radiative forcing parameter as per used in the snowball Earth experiments: `ea_radfor_scl_co2`. By default in the modern continental configuration, this has a value of 1.0, corresponding to 278 ppm atmospheric  $CO_2$ . A value of 2.0 would reflect warming equivalent to 556 ppm  $CO_2$ . And 3.0 would be more like an end-of-the-century warming. Note that you are applying the warming instantaneously by manipulating the climate system in this way and hence the changes will be more extreme than those occurring over the time-scale of this century. Also note that a cooling (a radiative scaling smaller than 1.0) could be applied instead.

The same parameter as before appears at the start of the *user-config* provided for this Section (LAB.2.6.EXAMPLE) for you to make this radiative forcing adjustment:

```
# scaling for atmospheric CO2 radiative forcing, relative to 278 ppm
ea_radfor_scl_co2=1.0
```

Remember that it is best practice to copy-rename-edit LAB.2.6.warming to create a new and unique experiment name for a new experiment.

You are running your new experiment from the same *re-start* state as before, e.g.:

```
$ ./runcookie.sh cookie.C.p_worjh2.r LABS LAB.2.6.warming 10
cookie.C.p_worjh2.r.SPIN
```

Potentially interesting properties of the Earth system to look at include sea-ice extent and AMOC strength (in the ASCII time-series files), and the overturning stream-function (2D with **netCDF** variable `phys_opsia`) and sea-ice extent in the 2-D **netCDF** output. With the age tracer included you can also make projections of how patterns of ventilation age change with transient warming.

Finally – is the spatial pattern of warming even? You will need to create a difference map of surface temperature – variable `atm_temp` in `fields_biogem_2d.nc` – to be able to see this well.

Try and answer the question: **How much radiative forcing is required to collapse the AMOC?** (and how much atmospheric  $CO_2$  does this approximately correspond to?) How non-linear is the response of overturning to fresh water flux?

You will need to submit multiple experiment jobs to the cluster queue, with each experiment differing in the assumed radiative forcing parameter value. Plot the strength of the AMOC that is reached at the end of each experiment (from the time-series file: `biogem_series_misc_opsi.res`) vs. radiative forcing.

## 2.7 Further ideas

### 2.7.1 Assessing the spatial sensitivity of deep-water formation to shutdown

What is the largest freshwater flux that can be sustained without ‘collapsing’ the AMOC? Is there a ‘threshold’ (‘tipping point’) of freshwater input, beyond which the AMOC rapidly decreases in strength? For this – you would run (submit to the cluster) a series of experiments, each with different (increasing) values for the fresh water flux. Remember to include an experiment with zero freshwater flux to act as a ‘control’.

Is the precise location of the freshwater input important (i.e., try tipping it in somewhere else)? For this – you could piece-meal run an experiment, analyse the results, then choose a new location to input the freshwater, or, come up with a systematic search pattern of freshwater input patterns.

Outside of the North Atlantic, are any other regions (of deep water formation) sensitive to freshwater perturbation and what are the consequences (could it happen in the future)?

### 2.7.2 Forcing stronger deep-water formation (‘anti-hosing’ investigations)

There are questions concerning past changes in the AMOC as to whether it is ‘pushed’ or ‘pulled’. i.e., if the AMOC shoals in depth and/or weakens, is it because its production has weakened, or as Antarctic Bottom water (AABW) strengthened and ‘pushed’ it out of the way (to shallower depths)?

What you might try then is to inject salt in the Southern Ocean as opposed to fresh water in the North Atlantic. All you need do is pick an appropriate grid point (this is worth thinking about carefully and maybe testing different locations) and rather than giving the parameter `bg_par_ocn_force_scale_val_2` a negative value, you give it a positive one. (Start by trying similar magnitudes of value as before and see what happens.)

**Is the AMOC (for the same magnitude of forcing) more sensitive to being ‘pushed’ or ‘pulled’?** (Obviously the answer will depend on where the perturbations are being applied.)

### 2.7.3 Regional hosing – Poking the Mediterranean

It is also possible to apply freshwater fluxes in a specific pattern/region, rather than just at a single location.

The experiment configuration (*user-config*) – LAB.2.7.saltymed – is almost identical to all the other ones in this Chapter, except that it prescribes a net salt transport into the Eastern Mediterranean (5 cells on the model grid) with a balancing freshwater flux distributed across the remainder of the global ocean surface. This is implemented by the forcing:

```
bg_par_forcing_name="p_worxx2.Fsal_SUR_EMed"
```

This is paired with a scaling parameter for changing the magnitude of the flux forcing:

```
bg_par_ocn_force_scale_val_2=0.0
```

which by default is zero (i.e., no flux forcing). Setting the value of this parameter to:

`bg_par_ocn_force_scale_val_2=1.1E18` would create a net freshwater removal from the Eastern Mediterranean of 1Sv (probably the largest you want to try). Note that you need to apply a value that is positive so that we get a salt addition rather than freshwater addition in the Eastern Mediterranean.

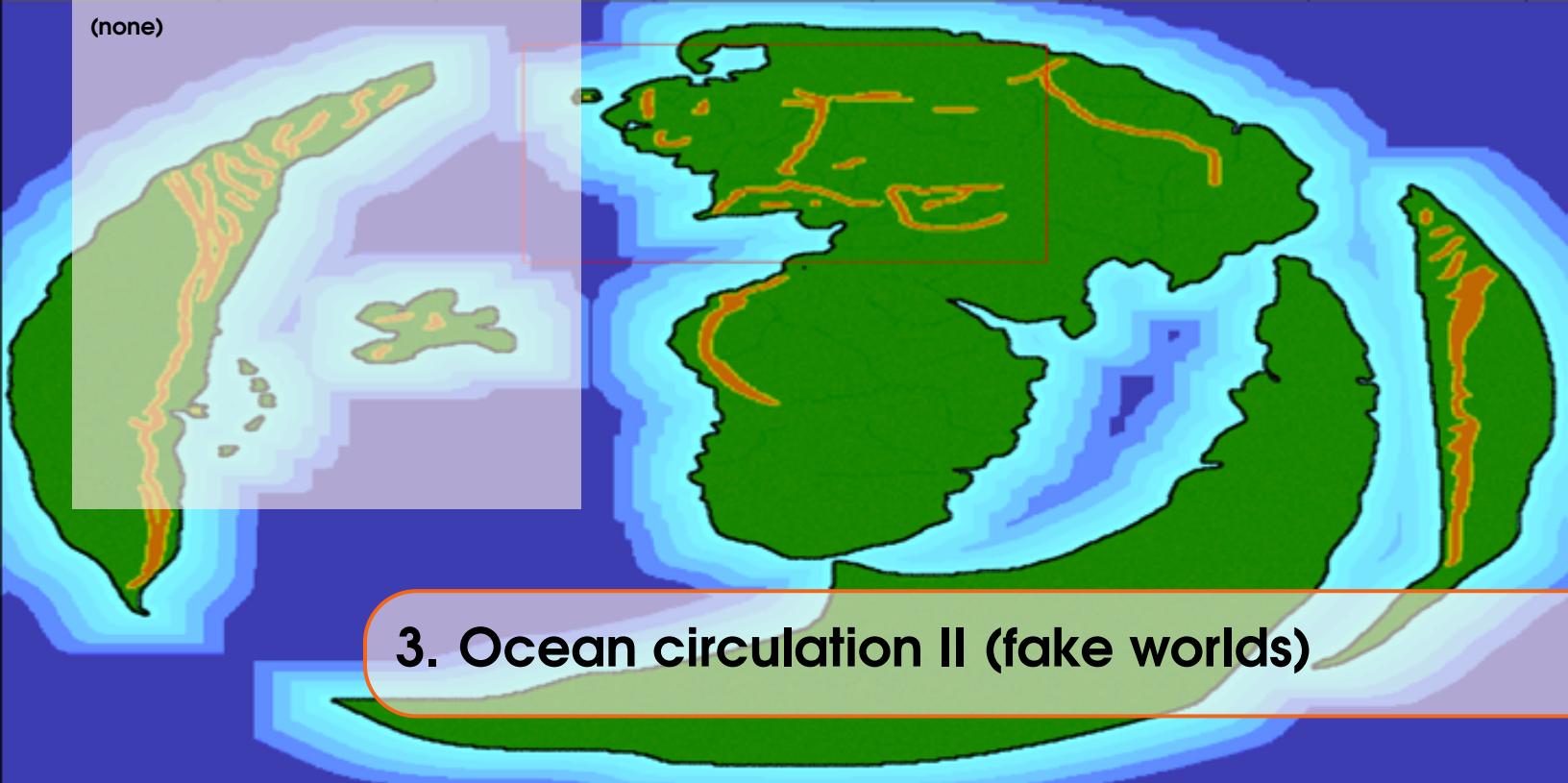
Try running the experiment for 20 years to start with. i.e.,

```
$ ./runcookie.sh cookie.C.p_worjh2.r LABS LAB.2.7.saltymed 20  
cookie.C.p_worjh2.r.SPIN
```

(again running on from a *re-start*).

You should find a rich response to adding salt to the Eastern Med. Initially, the AMOC weakens because you are taking salt away from everywhere including the N. Atlantic to feed into the Med. After year ca. 15, something interesting happens ... I am not going to spoil the surprise – you need to run the experiment and then work out 'why' the model does what it does (by repeating out some of all of the analyses you have carried out previously).

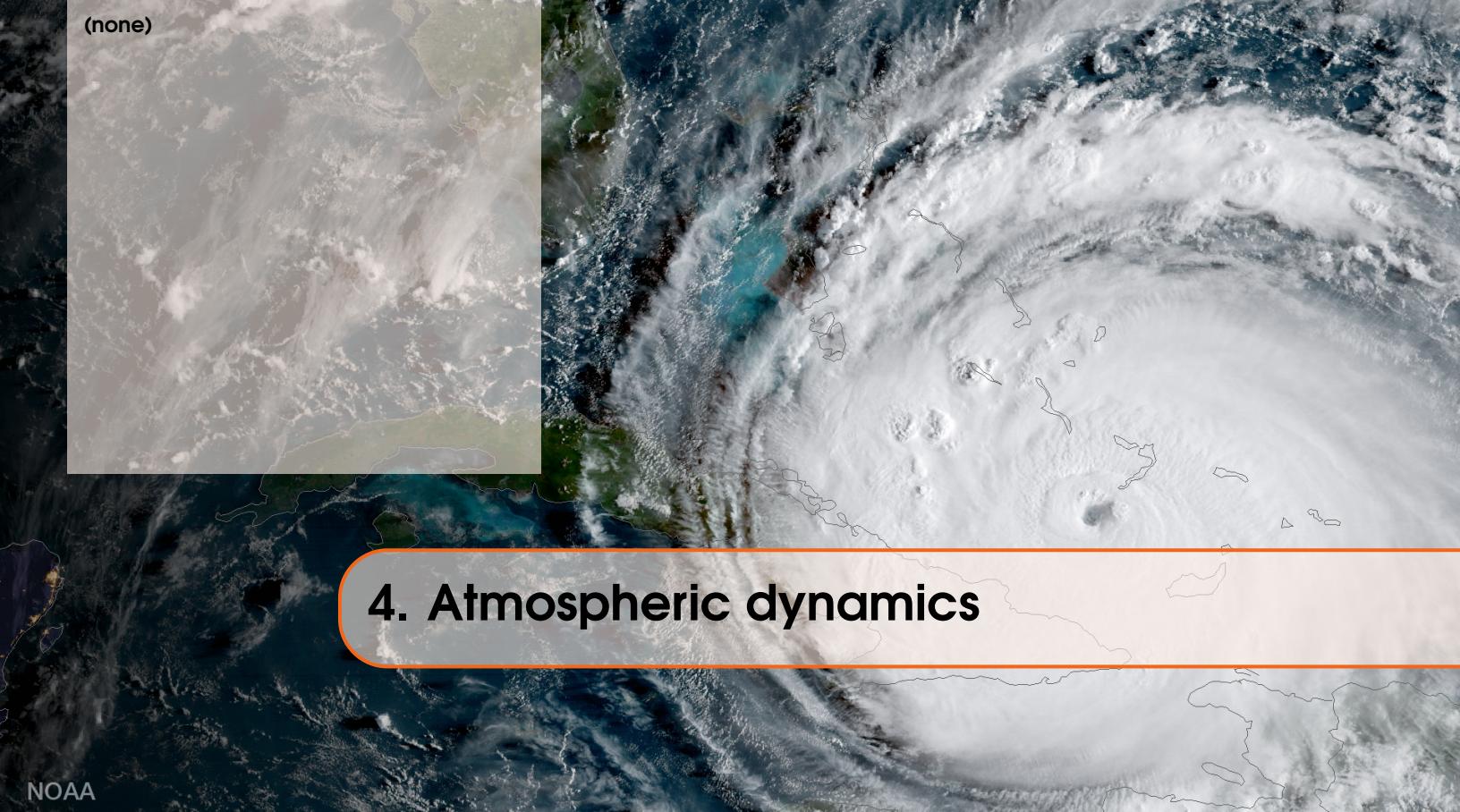
(none)



### 3. Ocean circulation II (fake worlds)

**3.1 (none)**

(none)



## 4. Atmospheric dynamics

NOAA

**4.1** (none)

## Exploring the consequences of fossil fuel

### $\text{CO}_2$ emissions

Idealized emissions forcing

Where has my carbon gone???

Historical (real-world!) emissions forcing

Assessing future carbon emissions impacts

### Further ideas

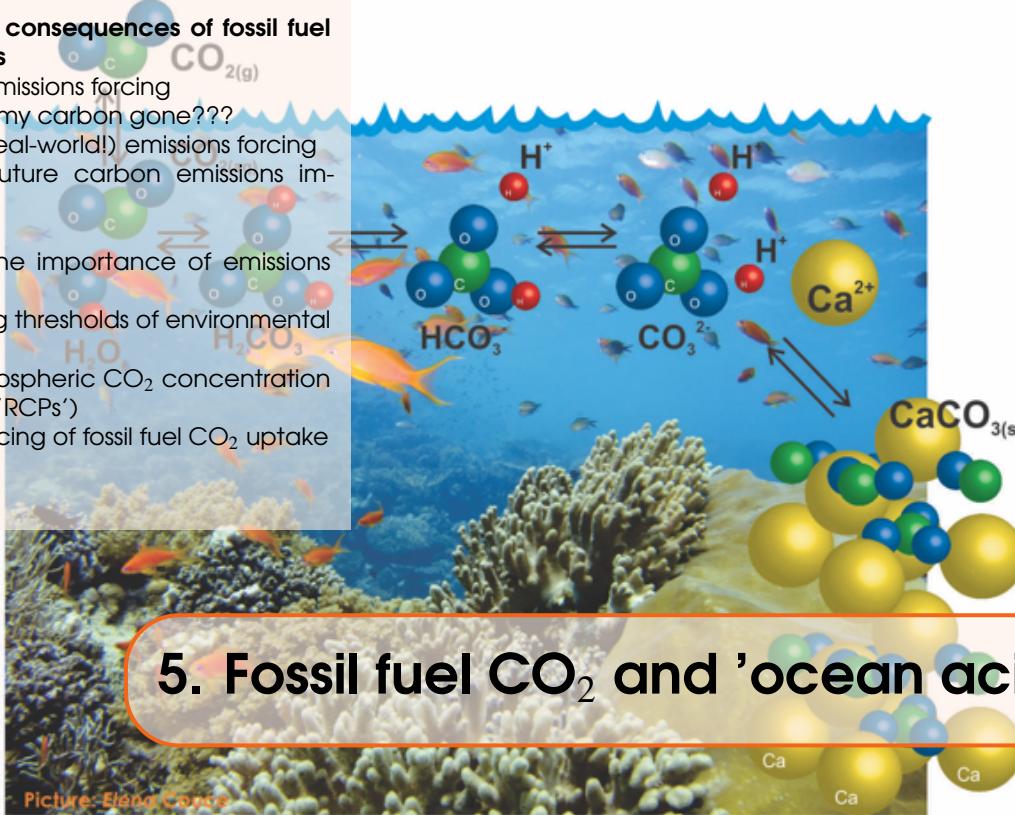
Assessing the importance of emissions rate

Determining thresholds of environmental impact

Future atmospheric  $\text{CO}_2$  concentration pathways ('RCPs')

Isotopic tracing of fossil fuel  $\text{CO}_2$  uptake

### Appendix



## 5. Fossil fuel $\text{CO}_2$ and 'ocean acidification'

**Readme**

You will need to download a new *restart* file prior to embarking on the experiments. This pre-industrial spin-up includes a basic ocean (-atmosphere) carbon cycle. It follows the physical climate configuration of *Cao et al.* [2009] but to simplify your introduction of carbonate chemistry, does not include biology (or associated e.g., nutrient tracer).

To fetch this: change to the cgenie\_output directory, and type (or copy-paste with care<sup>1</sup>):

```
$ wget --no-check-certificate  
http://www.seao2.info/cgenie_output/cookie.CB.p_worjh2.rCARB.SPIN.tar.gz
```

Extract the contents of this archive by typing:

```
$ tar xfzv cookie.CB.p_worjh2.rCARB.SPIN.tar.gz
```

(and then change directory back to genie-main to run the model)

---

<sup>1</sup>ALL one line with a space before the URL.

## 5.1 Exploring the consequences of fossil fuel CO<sub>2</sub> emissions

For the next experiment(s) you can chuck CO<sub>2</sub> into the atmosphere, just for the hell of it. As much as you want! Apparently, humans are actually doing this now. Imagine that!

A new *user-config* is provided – LAB.5.1.EXAMPLE – which unlike experimental setups in previous Chapters, is configured with climate being responsive to any changes in atmospheric CO<sub>2</sub> (i.e., it takes account of CO<sub>2</sub>-climate feedback). (We can make this choice because now we are going to calculate atmospheric CO<sub>2</sub>.) The setting (near the top of the *user-config* file) that does this is:

```
# set CO2-climate feedback  
ea_36=y
```

(Previously it was disabled (ea\_36=n) and climate was fixed at a given radiative forcing equivalent, set by parameter ea\_radfor\_scl\_co2.)

Additional (netCDF) output has also been prescribed, via the *user-config* parameter setting: bg\_par\_data\_save\_level=10 so that more data relevant to assessing ocean acidification is saved.

---

First, you should start with a control experiment to provide a baseline response (or hopefully, lack of response) to which you can contrast your perturbation experiments. It is also a good experiment to interrogate and explore new output and concepts (i.e., carbonate chemistry). For this, copy and rename<sup>2</sup> the example *user-config* for this section – LAB.5.1.EXAMPLE.

```
$ ./runcookie.sh cookie.CB.p_worjh2.rCARB LABS LAB.5.1.control 10  
cookie.CB.p_worjh2.rCARB.SPIN
```

Now, run an actual experiment. The *user-config* LAB.5.1.emissions is almost identical to your control experiment, except it specifies a release of CO<sub>2</sub> to the atmosphere, which by default is set to a value of just 1 PgC (cf. current emissions are ca. 10 PgCyr<sup>-1</sup>) and over an interval of just a single year. (Releasing CO<sub>2</sub> just over a single year is obviously rather unrealistic and many impacts will decay away rapidly, but represents a useful idealized experiment for assessing the time-scale(s) of fossil fuel CO<sub>2</sub> uptake by the ocean.)

Run for e.g., 10 years (or for somewhat longer if you like but make sure that you have a *control* experiment of the same length) and also starting from the preindustrial *re-start* experiment cookie.CB.p\_worjh2.rCARB.SPIN, i.e.:

```
$ ./runcookie.sh cookie.CB.p_worjh2.rCARB LABS LAB.5.1.emissions 10  
cookie.CB.p_worjh2.rCARB.SPIN
```

**A summary of relevant *time-slices* and *time-slice* output is given in the Appendix to this Chapter.**

What specific model results variables to consider ... ? Obviously atmospheric CO<sup>2</sup> to start with. Think about the climate change and ocean acidification literature and which environmental (physical and geochemical) properties are considered either critical for ecosystems or are simply helpful and/or illustrative. In the 3-D netCDF *time-slice* file remember, for instance, that ocean surface waters in which aragonite becomes under-saturated (OHMEGA < 1.0) is regarded as a critical threshold for organisms making aragonite shells and skeletons and spells TROUBLE for some poor calcifying marine organism somewhere. For climate change ... the variables of particular interest should be obvious. Remember that there are both *time-series* outputs, as well as 2D and 3D fields, any or all of which might be helpful for elucidating impacts.

---

<sup>2</sup>e.g., LAB.5.1.control

### 5.1.1 Idealized emissions forcing

The default settings in the experiment that you have just run was for  $1 \text{ PgCyr}^{-1}$  over a single year (i.e., an idealized pulse of unit size and duration). You can easily modify the experimental design to release more/less CO<sub>2</sub> very much as you did for the red dye tracer. In the *user-config* file, the line:

```
bg_par_atm_force_scale_val_3=8.3333e+013
```

scales the time-history of the CO<sub>2</sub> flux, given in the forcing file:

```
biogem_force_flux_atm_pCO2_sig.dat
```

... which can be found in the directory:

```
cgenie.cookie/genie-forcings/pyyyz.FpCO2_Fp13CO2
```

The format of this file is:

```
-START-OF-DATA-
 0.0  1.0
 1.0  1.0
 1.0  0.0
999999.9 0.0
-END-OF-DATA-
```

and defines an emission of  $1 \text{ molC}$  (carbon) per year over the first year (year 1.0) of the model experiment (between year 0.0 and 1.0), but which in the example *user-config* is then scaled by a value of  $8.333 \times 10^{13}$  (by the parameter *bg\_par\_atm\_force\_scale\_val\_3*) to give a total of  $1 \text{ PgCyr}^{-1}$ . (Year 999999.9 has no special meaning and is simply just waaaaa into the future ...)

Pause ... and note briefly how the final CO<sub>2</sub> flux is arrived at. **cookie** calculates it by multiplying the value in the forcing file (1.0) by a modifying parameter in the *user-config* file (8.3333e+13). The total flux is hence:  $1.0 \times 8.333 \times 10^{13} = 8.333 \times 10^{13} \text{ molCO}_2\text{yr}^{-1}$ . If you set both values as 1.0, you'd get very little carbon released (a single mol!). If you screw up and multiply 8.3333e+013 and 8.3333e+013 together as the total flux ... you'll soon know it as you cook the Earth ... But it does not matter which parameter has value 1.0 and which scales the units (8.3333e+013). For now, it is simply more convenient to be able to edit the *forcing* file with 'simple' numbers (and leave the large numbers and units conversion in the *user-config* file).

Together, the scaling and forcing value gives a CO<sub>2</sub> release of  $1 \text{ PgCyr}^{-1}$  for just a single year compared to current emissions are about  $10 \text{ PgCyr}^{-1}$ . So, do not expect anything exciting if all you emit to the atmosphere is a single measly  $1 \text{ PgC}$  (over 1 year). (The parameter: *bg\_par\_atm\_force\_scale\_val\_4=-27.0* specifies the carbon isotopic composition of fossil fuel carbon and can be ignored for now.)

If you want modify files in a forcing directory, it is good practice to copy the entire directory (in the example here: pyyyz.FpCO2\_Fp13CO2 and then copy and rename it (within the *forcings* folder –cgenie.cookie/genie-forcings). And then edit the file(s) you want modified. The intention is that you always retain a copy of the original, unmodified *forcing*.

Having created a forcing folder with a different name, you point to it by setting the parameter:

```
bg_par_forcing_name="MYFORCING"
```

where MYFORCING is whatever you called the forcing directory  
(it was originally: *bg\_par\_forcing\_name="pyyyz.FpCO2\_Fp13CO2"*).

The importance of the control experiment which you have already run, is that ‘accidents can happen’ and the global environmental changes induced by the massive fossil fuel CO<sub>2</sub> release can obscure mistakes made in the experiment configuration (parameter values) and/or the *re-start* used. The template experiment provided that you used as a control – LAB.5.1.EXAMPLE is designed to be identical to that of the actual experiment itself (LAB.5.1.emissions) with the exception of the scaling of the CO<sub>2</sub> emissions, which is set to zero.<sup>3</sup> i.e.:

```
bg_par_atm_force_scale_val_3=0.0
```

If everything is OK with the control experiment, atmospheric CO<sub>2</sub> (and climate) following on from the *re-start* should be stable and there should be little (or no) drift in any of the output variables (because the *spin-up* you are re-starting from should have been run to an equilibrium state and you have not changed anything in the control experiment, right?).

It is good practice (i.e., always do it!) to always run a control experiment for each different type of experiment – e.g., you only need to run one control experiment for a set CO<sub>2</sub> emissions experiments differing only in total carbon release or the time-history of that release. When you have run both the real and control experiment, compare the results. View (or plot) both relevant *time-series* output, and create anomaly maps of key *time-slice* variables in **Panoply** or **MATLAB**, using a corresponding *time-slice* from the control experiment to create the experiment anomaly with.

OK. You might want to run something a little more exciting now. For instance, rather than

```
-START-OF-DATA-
 0.0  1.0
 1.0  1.0
 1.0  0.0
999999.9  0.0
-END-OF-DATA-
```

you might have:

```
-START-OF-DATA-
 0.0  1000.0
 1.0  1000.0
 1.0      0.0
999999.9      0.0
-END-OF-DATA-
```

for a total of 1000PgC is released over a single year. Now you should see some policy-relevant impacts occur :o)

You can control the shape of the emissions profile as well as its magnitude. Between the start and end ‘tags’ in the text *forcing* file, the data is arranged into 2 columns: the first contains a series of tie-points for defining the timing of changes in emissions, and the 2nd column contains flux information (units of PgCyr<sup>-1</sup> when scaled by the parameter parameter bg\_par\_atm\_force\_scale\_val\_3 in the *user-config*). At each time-step of the model, the CO<sub>2</sub> flux to be applied to the atmosphere is interpolated between these time points.

<sup>3</sup>For completeness, the isotopic value of emitted CO<sub>2</sub> is also scaled to zero, but because there are no emissions in the control experiment, this does not actually achieve anything in practice.

For instance, in the *forcing* (directory) file biogem\_force\_flux\_atm\_pCO2\_sig.dat, the purpose of:

```
0.0 1.0
1.0 1.0
1.0 0.0
999999.9 0.0
```

is to specify a uniform flux of 1.0 (scaled to  $PgCyr^{-1}$ ) over the first full year of the model run, followed by a sharp turn-off to zero flux at the end of first year (and remaining zero thereafter). To extend the period of emissions – for example:

```
0.0 1.0
10.0 1.0
10.0 0.0
999999.9 0.0
```

would result in a uniform flux lasting 10 years with a sudden cut-off and zero thereafter (i.e., once scaled by the parameter in the *user-config* –  $1PgCyr^{-1}$  over 10 years –  $10PgC$  total emissions).

In contrast:

```
0.0 0.0
10.0 1.0
10.0 0.0
999999.9 0.0
```

would result in a linear ramp, starting from zero at the start of year 0.0, to  $1.0PgCyr^{-1}$  at year 10.0 and then suddenly ceasing and remaining at zero for the remainder of the experiment (a total  $CO_2$  emission of  $1 \times 1.0 \times 0.5 = 5PgC$  over 10 years).

To ramp up (over 10 years), and then down again (over 10 years), you would specify:

```
0.0 0.0
10.0 1.0
20.0 0.0
999999.9 0.0
```

Try making up a few 'shapes' (and hence different experiments), maybe for the same integrated/total emissions, explore the effect of different rates of rise/fall in the release rate. And/or for the same release rate and/or duration, explore the impact of different total emissions of  $CO_2$ . Try and think in terms of hypotheses and formulate questions to guide your experimental design/configuration. (An alternative approach is to create random scenarios and in the analysis, fish for interesting patterns that could lead to knowledge and/or specific questions to be tested further, but it is better to start off with a hypothesis in mind.)

Note that you can either edit and re-use the same *forcing* directory and name, modifying the file biogem\_force\_flux\_atm\_pCO2\_sig.dat each time but then losing an explicit record of how you might have set the emissions profile previously, or you can copy and rename the entire *forcing* directory (and then edit biogem\_force\_flux\_atm\_pCO2\_sig.dat). If you copy and rename the entire *forcing* directory, in the *user-config*, you then need to specify this new forcing (directory) name, e.g.:

```
# specify forcings
bg_par_forcing_name="pyyyzz.FpCO2_Fp13CO2.NEW"
```

if you, for instance, called your new *forcing* directory (in genie-forcings): pyyyzz.FpCO2\_Fp13CO2.NEW.<sup>4</sup>

---

<sup>4</sup>Refer to the directory map in Figure 1.1. if in doubt here.

### 5.1.2 Where has my carbon gone???

In any of the emissions experiments you have tried out, in the time-series of atmospheric  $pCO_2$  (file: biogem\_series\_atm\_pCO2.res) you'll undoubtedly see atmospheric  $pCO_2$  initially rise, but then once the emissions cease, start decaying back down again. Where is it 'going'?

Well obviously the ocean. D'uh! (At least, this is true in this particularly configuration of **cookie** without a terrestrial biosphere.) A better question would be: 'where in the ocean has it gone?', and even better: 'why there?'.

In the 3D netCDF, the variable ocn\_DIC is the total dissolved carbon inorganic concentration (*DIC*). Open this up ... and by slicing horizontally (e.g. start at the surface, and then slice downwards), or vertically (up through the middle of the Atlantic would be a good latitude-vertical section to create), can you 'see' where the carbon (as *DIC*) is going? If not ... why not? Try making the same data sections from the same year of the control experiment. *DIC* is everywhere in the ocean in the control, with a highly spatially variable distribution. It could be then that the carbon taken up from the atmosphere in your experiment, simply overprints too small a pattern of (fossil fuel *DIC*) to tell background + fossil fuel from just background. (A similar situation arose when looking for the surface temperature impact of a weakening AMOC.)

To resolve this, create difference maps of a slice (lon-lat, or lat-depth) from your experiment at time  $t$  minus the control, also at time  $t$ . Now re-evaluate whether you can tell where the fossil fuel  $CO_2$  is going. Why (there)? (We'll also consider this question in the next exercise.)

---

What about the other carbonate system parameters? Can you track identify patterns of uptake and ocean circulation transport. For instance  $CO_{2(aq)}$ ? What about  $CO_3^{2-}$  (before looking, and remembering your basic carbonate chemistry, what would you expect)? ( $HCO_3^{2-}$  should primarily track *DIC*.)

Are there any other carbonate system impacts you can discern? Some fields to consider might include:

- $pH$  – variable misc\_pH

(There is also a field for the hydrogen ion concentration if you really want to see it.)

- Calcite saturation state ( $\Omega_{(cal)}$ ) – variable carb\_ohm\_cal

- Aragonite saturation state ( $\Omega_{(arg)}$ ) – variable carb\_ohm\_arg

---

In looking at the different fields and based on your reading of the literature (you did read the background papers, right ... ?), think about what organisms live where and what environmental (carbonate system) variables might affect them (or their prey).

---

Also see the subsection – "Isotopic tracing of fossil fuel  $CO_2$  uptake".

**5.1.3****Historical (real-world!) emissions forcing**

Historical and future (e.g. IPCC 'SRES') emissions scenarios can be prescribed explicitly and simply in **cookie**. An example is given as *user-config* file: LAB.5.1.historical. In this, a historical emissions forcing (technically: a prescribed concentration profile of *pCO<sub>2</sub>* is specified by the *forcing*):

```
bg_par_forcing_name='pyyyzz.RpCO2_Rp13CO2.historical2010'
```

In contrast to before, no additional scaling is needed because the forcing specification directly follows the observed change in atmospheric concentration with time (in units of atm *CO<sub>2</sub>*).

Note that an additional line appears in the *user-config*. This is because the historical *pCO<sub>2</sub>* transient starts in the 1700s (for which a nominal date of 1765 is often used) rather than year zero. To start **cookie** counting from year 1765 rather than year zero, a start year parameter value is specified:

```
bg_par_misc_t_start=1765.0
```

It is also convenient to specify a set of points in time at which data is saved that are consistent with the historical period. In the example *user-config*, the addition of the parameter settings:

```
bg_par_infile_slice_name='save_timeslice_historicalfuture.dat'  
bg_par_infile_sig_name='save_timeseries_historicalfuture.dat'
```

specifies a series of time points at which data is saved that aligns with historically relevant years. Try viewing the contents of these (text) files:

```
save_timeslice_historicalfuture.dat  
save_timeseries_historicalfuture.dat
```

which can be found in the directory: `genie-biogem/data/input`, to get a sense of how frequently the data will be saved, and how this differs from the default settings, which are defined in the files:

```
save_timeseries.dat  
save_timeslice.dat
```

Running a transient historically-forced experiment looks like this (all one line!):

```
$ ./runcookie.sh cookie.CB.p_worjh2.rCARB LABS LAB.5.1.historical 245  
cookie.CB.p_worjh2.rCARB.SPIN
```

The 245 parameter value for experiment duration (years) arises on the basis of the start year being 1765 (as specified by `bg_par_misc_t_start=1765.0`) and to give an experiment end at year 2100.

Note that from year 1765 onward, changes in atmospheric *CO<sub>2</sub>* only rise very slowly initially. Don't expect to see anything happen in 10 seconds flat because relatively few people and countries in the 1800s could be bothered to burn much more than a little local coal. You could potentially start your experiment at year 1850, changing the value of `bg_par_misc_t_start` and specifying shorter experiment duration (150 years) if you are desperate for the End of the World to come.<sup>5</sup>

Given that there is observationally-based information on the distribution of anthropogenic *CO<sub>2</sub>* taken up by the ocean (e.g., *Sabine et al.* [2004]) ... and you are running a historical transient experiment with the model driven by observed increases in atmospheric *pCO<sub>2</sub>* ... you are in a position to critically evaluate the models ability (or lack of) to represent the future-critical process of oceanic fossil fuel *CO<sub>2</sub>* uptake and transport by large scale ocean circulation.

<sup>5</sup>Don't forget: you could submit this experiment to the cluster and do more (idealized emissions) 'playing' which it runs.

In the 2D **netCDF** output, there is a variable for the water column integrated inventory of DIC – equivalent to the Sabine map except you will need to subtract the preindustrial background of DIC first, i.e., to create a DIC anomaly map representing only the added fossil fuel CO<sub>2</sub> component of ocean DIC. The data in the Sabine paper clusters around 1994. A *time-slice* centered on this year (1994.5) has been configured in the model exactly for this purpose. Your baseline state can either be from prior to CO<sub>2</sub> emissions commencing at any significant rate (e.g., 1750.5) or (better), from a control experiment. Note that similar comparisons could be (and are regularly) made with other tracers such as CFCs, which provide additional insights into the patterns and time-scales of trace gas update and ocean circulation. (See: *Cao et al.* [2009])

Observational data, re-gridded to the **cookie** grid and in netCDF format can be downloaded from: <http://www.seao2.org/mymuffin.html> (and under the ‘got data?’ box on the left). You could for instance, compare horizontal or vertical slices (3D netCDF) and create difference (anomaly) maps. Somewhat more representative of the entire ocean is to compare (or calculate difference maps) of zonal average profiles. Unfortunately, the observations are not in the form of water column integrals and hence you cannot create difference maps of model as per the *Sabine* paper ... unless you use the 3D **BIOGEM MATLAB** plotting scripts. Examples of **MATLAB** plotting of the model *vs.* observed anthropogenic anomaly are shown in Figure 5.1.

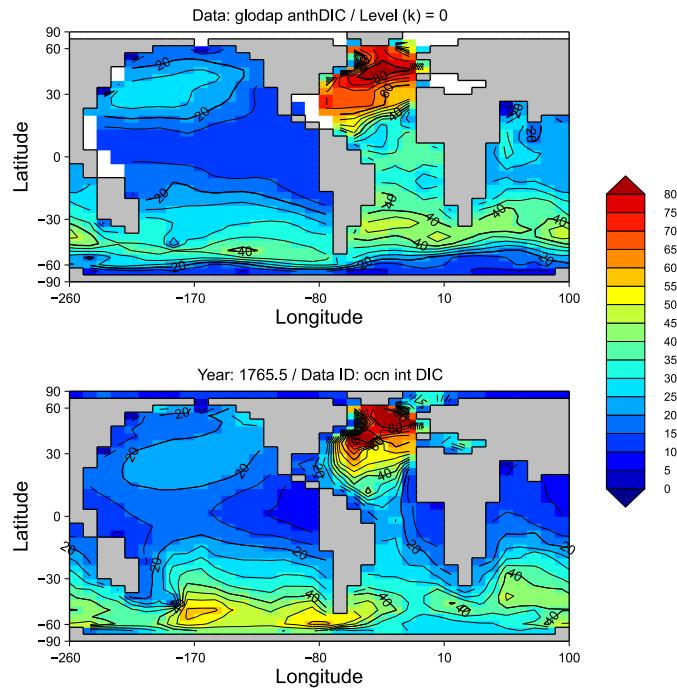


Figure 5.1: Observed (top) *vs.* Model (bottom) anthropogenic CO<sub>2</sub> inventories. Data and model water column integrals in units of mol CO<sub>2</sub> m<sup>-2</sup> and are nominally with respect to year 1994.

Note that model and data are not strictly comparable in the particular configuration used here, as we do not have any biology or a ‘biological pump’ (more on this later) in the ocean. (The example in the figure uses the with-biology biogeochemical configuration described by *Cao et al.* [2009].)

**5.1.4****Assessing future carbon emissions impacts**

Finally, and the closest to being slightly interesting: rather than applying highly idealized pulses of CO<sub>2</sub> emissions, the IPCC 'SRES' emissions scenarios<sup>6</sup> can be used to make future projections with. An example forcing of this sort is provided and can be selected by changing the name of the forcing selection parameter (`bg_par_forcing_name`) to any one of the following:

- pyyyyz.FpCO2\_Fp13CO2.A1\_AIM
- pyyyyz.FpCO2\_Fp13CO2.A1G\_MINICAM
- pyyyyz.FpCO2\_Fp13CO2.A1T\_MESSAGE
- pyyyyz.FpCO2\_Fp13CO2.A2 ASF
- pyyyyz.FpCO2\_Fp13CO2.B1 IMAGE
- pyyyyz.FpCO2\_Fp13CO2.B2\_MESSAGE

These are 'future' emissions scenarios, which all start at year 2010, and end at year 2100. They are derived from different socio-economic and future technological assumptions in making their future emissions projections.

Again, as these forcings have units of  $PgCyr^{-1}$  in the time-series files, you will need to ensure that the scaling parameter in your *user-config* file is set to turn units of  $PgCyr^{-1}$  into  $molCyr^{-1}$ .<sup>7</sup>

```
bg_par_atm_force_scale_val_3=8.3333e+013
```

You will want to run your experiment starting from the end of the historical transient experiment you have just run rather than the original steady-state *re-start*:<sup>8</sup>

```
$ ./runcookie.sh cookie.CB.p_worjh2.rCARB LABS LAB.5.1.future 90 LAB.5.1.historical
```

and to run for 90 years from 20120 to 2100, set the start year to the year that the previous historical transient finished on (2010):

```
bg_par_misc_t_start=2010.0
```

You can also easily replace the details of the emissions with other SRES scenarios – simply find the year *vs.* emissions rate information from the interweb<sup>9</sup> and edit or copy-and-paste the flux values for each decade into the file `biogem_force_flux_atm_pCO2_sig.dat` in the forcing directory. `cookie` will then automatically interpolate between the decadal tie-points to give a continuous change in emissions. Now you are able to make a rather more realistic/plausible assessment of when and where potential ecological impacts (via assumed ocean chemistry criteria) might occur.

Try running (e.g. as jobs submitted to the cluster queue) some other actual or made up CO<sub>2</sub> emissions scenarios.

<sup>6</sup>Note that the past few IPCC assessment reports have switched to 'using 'RCP's – Representative Emissions Pathways, rather than the SRES emissions scenarios – see later Section.

<sup>7</sup>For completeness also add in the specification of the isotopic composition of the carbon emissions – refer back to the idealized experiments.

<sup>8</sup>Note that the *user-config* `LAB.5.1.future` is not provided for you – you will need to create this (or a file named whatever you like) by copying e.g., `LAB.5.1.emissions` and making the parameter changes described above (forcing specification parameter, emissions scaling parameter, and start year parameter).

<sup>9</sup>e.g., [http://sres.ciesin.columbia.edu/final\\_data.html](http://sres.ciesin.columbia.edu/final_data.html)

## 5.2 Further ideas

Some further possibilities for investigations that build on the basic previous ones.

### 5.2.1 Assessing the importance of emissions rate

By editing the flux magnitude and/or timing (i.e. the years that are assigned to the different forcing time-points) information of the idealized emissions *forcings*, you can control the  $CO_2$  emissions trajectory as well as the total of fossil fuel carbon burned. Explore some different assumptions about  $CO_2$  release rate but for the same total carbon emitted, and note their differing impact on climate and ocean (carbonate) geochemistry. For example rather than  $1000 PgC$  over a single year (which you tested earlier), in terms of emissions rate, more realistic would be e.g., 10 or  $20 PgCyr^{-1}$  but spread over a longer interval (order of 100 years), for the same total carbon release. For instance, one might try and address the question: “For a given total release of fossil fuel  $CO_2$ , is it safer to burn it slower?” The answer is maybe not completely obvious, as burning carbon resources slower will result in a small global impact, but perhaps one that persists for longer(??). You could conceive of an ensemble (related set) of model experiments, maybe one of  $100 PgCyr^{-1}$  for 1 yr, one of  $10 PgCyr^{-1}$  for 10 years, and one of  $1 PgCyr^{-1}$  for 100 years.<sup>10</sup>

Because the experiments are getting longer to run in real time . . . remember to make appropriate use of the cluster queuing facility – i.e., think about whether you want to sit around starting at the screen for 15 minutes waiting for a new line of numbers appear – if not: submit to the cluster queue. (Don’t forget the control experiment! (configured the same, except with  $0 PgCyr^{-1}$  of emissions for 100 years).)

Note note that ideally you would create a new *forcing* based on the original if you are editing the same original *forcing* and expecting to run different ones at the same time. Really, this is little more than you did in copying and renaming *user-config* files in order to create new experiments ... except that now it involves copying and renaming entire directories in *genie-forcings*. Remember that the *forcing* is specified by the directory name assigned to *bg\_par\_forcing\_name* (enclosed in '') and you will need to change this to match the name of your new *forcing* directory.

### 5.2.2 Determining thresholds of environmental impact

There are various concerns about the impacts of continuing fossil fuel  $CO_2$  emissions and a number of proposed climatic (e.g., the  $1.5^\circ C$  paris protocol global warming limit often mentioned in policy documents) and ecological ‘tipping points’. You could assess the maximum allowable  $CO_2$  emissions to remain within particular global environmental limits in the model. For example:

- What is the maximum total  $CO_2$  release that can be made without inducing aragonite undersaturation at the ocean surface anywhere (or any season – see Section 5.2.3 in the User Manual for seasonal time-slice data saving)? How important is the time-scale of emissions in determining this? For total emissions above this: where in the ocean does the surface first become under-saturated and what sort of (calcifying) organisms might be impacted there? How large would the emissions have to be in order to start to induce under-saturation with respect to aragonite at the surface in the tropics (home to socio-economically important reef systems)? These are questions that can be addressed with simple  $CO_2$  release experiments in

<sup>10</sup>These all represent rather unrealistically small total  $CO_2$  releases and you may want to consider a total more like  $1000 PgC$  or rather more. You may also want to think about more realistic shapes rather than pulses, such a some sort of ramp up and then down in the emissions rate (but for the same total emissions).

ocean carbon cycle models and everyone seems to get a GRL paper out of it each and every time!

- How important are CO<sub>2</sub>-climate feedback in amplifying or diminishing future climate and ocean carbonate chemistry changes – e.g., is the same atmospheric pCO<sub>2</sub> value reached with and without climate feedback (and surface warming) – if not, why?

Hint: the solubility of CO<sub>2</sub> in sea-water is a function of temperature and at higher temperatures, CO<sub>2</sub> is less soluble. This means that with climate warming, CO<sub>2</sub> solubility declines, less is taken up by the ocean and more is left in the atmosphere – driving further heating in a positive feedback.

All your CO<sub>2</sub> emissions experiments to date have had this feedback enabled – specified at the top of the *user-config* file by:

```
# set climate feedback
ea_36=y
```

To quantify the role of the carbon-climate feedback, you need to run an identical emissions experiment, but with the feedback disabled:

```
# set climate feedback
ea_36=n
```

Note that you also to need to run a historical transient experiments with no carbon-climate feedback, if you are starting emissions experiments from the year 2010. i.e. you will have a set of future emissions experiments including the carbon-climate feedback that are run from a historical transient experiments that also includes carbon-climate feedback, vs. a set of future emissions experiments without the carbon-climate feedback that are run from a historical transient experiments that also does not include carbon-climate feedback.

The importance of the feedback is simply the difference between the 2 sets of experiments, at the same year.

- Also: How large a CO<sub>2</sub> emission does it take to significantly ‘collapse’ the AMOC and over what time-scale? (Or alternatively: what is the atmospheric pCO<sub>2</sub> threshold for AMOC collapse in **cookie**?)

If the AMOC weakens or collapses ... why in the absence of a prescribed freshwater perturbation does this happen? What physical process are at play in response to rapid CO<sub>2</sub> release too the atmosphere that may act to reduce or shutdown deep-water formation in the ocean model? (Plotting appropriate ocean property anomalies between the CO<sub>2</sub> release experiment and a control experiment might help.)

Related to a previous possible investigation – does the rate of CO<sub>2</sub> increase (for the same total release) matter? If so, why? What is happening in general to the structure and dynamics of the upper ocean when surface warming is very rapid?

Experiments could be hypothetical and consisting of CO<sub>2</sub> pulses or ramps (or exponential) and run on directly from a pre-industrial spin-up, or more ‘realistic’ and run on from the end of a historical transient experiment (e.g., starting in year 2010).

---

Also:

- How much carbon can be burned (and how quickly) such that atmospheric  $pCO_2$  does not increase any further and remains at some specific value (you might pick  $\times 2$  or  $\times 4$  preindustrial  $pCO_2$  (ca. 278 ppm))? This is difficult to determine well because to hold atmospheric  $pCO_2$  constant, continuing emissions are required (as  $CO_2$  continues to be taken up by the ocean).

One possible approach to tackling this, might be:

1. Run one or more (SRES) emissions scenarios and find the year at which your  $pCO_2$  threshold value is crossed.
2. Re-run the same emissions experiment, but only for as many years as the threshold-crossing year you had identified, is reached. This will become your new *re-start*.
3. Create a new experiment, with zero emissions, and run on from your new re-start. You should see atmospheric  $pCO_2$  start close to the value you reached at the end of the last experiment, but then decay away as the ocean continues to absorb carbon from the atmosphere. The decline in atmospheric  $pCO_2$  each year tells you something about the yearly emissions needed to keep atmospheric  $pCO_2$  constant.  
An approximate rule-of-thumb, is that 1 ppm in atmospheric  $pCO_2$  is equivalent to 2 PgC (actually, a more exact conversion is 1 ppm = 2.123 PgC). So you could calculate a yearly time-series of how much (in ppm) atmospheric  $pCO_2$  falls by each year ... convert this to PgC, and then create a new *forcing*, with this as the annual emissions.
4. Create and run a new experiment, from the same *re-start* experiment you created, and apply the new emissions forcing you created. See if atmospheric  $pCO_2$  is approximately maintained constant(?)
5. If not sufficient constant to your satisfaction, you could also carry out a second iteration – calculating from your latest experiment, the ppm change in  $pCO_2$  each year, converting to an emissions rate, and running a further experiment ...  
Note that if you 'overheat' and  $pCO_2$  rises above the threshold value rather than falls below, you are allowed to have a negative carbon emission to the atmosphere in your forcing.
6. The sum of the carbon emissions up to the threshold being reached, is your answer to how much more carbon can be burned and not cross that threshold ... but you will see that after that, further emissions are allowed without  $pCO_2$  exceeding the threshold. So the answer e.g. for year 2100 (or 2200) will be larger.

There are automated ways provided in the model framework of achieving this, and you can for instance simply tell **cookie** to maintain a specific (or changing) value of atmospheric  $pCO_2$ , and from this diagnose the emissions rate that was required. For instance, you'll see in the original *re-start user-config* for this chapter – cookie.CB.p\_worjh2.rCARB.SPIN – the *forcing* used is pyyyyz.RpCO2\_Rp13CO2. If you go look in that forcing directory at:

biogem\_force\_restore\_atm\_pCO2\_sig.dat

you will see a unit forcing. This is then scaled by the parameter bg\_par\_atm\_force\_scale\_val\_3 to achieve a constant atmospheric  $pCO_2$  value of 278 ppm –  $0.278 \times 10^{-6} atm$ . To specify a different  $pCO_2$  value, simple change the scaling factor (as per e.g., for the flux forcing).

The equivalent carbon emissions required to do this, are diagnosed and provided as a *time-series* output (in units of  $molCyr^{-1}$ ) in file:

biogem\_series\_diag\_misc\_specified\_forcing\_pCO2.res

- Similarly – how much more carbon can be burned but still keep global mean surface air temperature from rising beyond the 'Paris' limit of 2°C (as compared to preindustrial)? This is also difficult to determine well because there are significant climate lags in the system with warming continuing even if atmospheric pCO<sub>2</sub> was held constant.
- Trial-and error would be one approach ...

### 5.2.3 Future atmospheric CO<sub>2</sub> concentration pathways ('RCPs')

The more recent/current incarnation of IPCC future scenarios revolves not around making projections of future greenhouse gas emissions rates, but rather future greenhouse gas concentration pathways. (The reasoning is partly to cut out the differences in carbon cycle feedback between models, where for the same unit CO<sub>2</sub> release, different climate/Earth system models might project a different atmospheric CO<sub>2</sub> concentration and hence climate change (in addition to differences between models in climate sensitivity).) In **cookie**, these work pretty well much like in the historical forcing scenario, where the observed change in atmospheric CO<sub>2</sub> concentration in the atmosphere with time, is prescribed and the model 'forced' to conform to this trajectory.

A series of RCP scenarios for how atmospheric CO<sub>2</sub> concentrations may evolve with time are provided. Each actually starts at year 1765 and hence incorporates the historical transient. They can hence be used with an experiment starting at 1765 and *re-starting* from a steady-state preindustrial *spin-up*. Or they can be jumped into at any point, and e.g. experiments started at year 2010, when only the year 2010 onwards part of the CO<sub>2</sub> restoring *forcing* is utilized. The RCP *forcings* currently provided are:

- pyyyyz.RpCO2\_Rp13CO2.RCP3PD
- pyyyyz.RpCO2\_Rp13CO2.RCP4p5
- pyyyyz.RpCO2\_Rp13CO2.RCP6p0
- pyyyyz.RpCO2\_Rp13CO2.RCP8p5

and can be selected simply by changing the name of the forcing in the *user-config*, e.g.:

```
bg_par_forcing_name='pyyyzz.RpCO2_Rp13CO2.RCP8p5'
```

which would select the RCP8.5 scenario.

NOTE: ensure that the scaling parameters, if set, are either commented out, e.g.:

```
#bg_par_atm_force_scale_val_3=8.3333e+013
#bg_par_atm_force_scale_val_4=-27.0
```

or simply deleted in their entirety (because the RCP forcing contains the actual values/correct units and you do not need to modify them any further).

To start at year 2010:

```
# change start year
bg_par_misc_t_start=2010.0
```

Note that the equivalent carbon emissions required to follow these concentration pathways are diagnosed by **cookie** and provided as a *time-series* output (in units of molCyr<sup>-1</sup>) in file:

```
biogem_series_diag_misc_specified_forcing_pCO2.res
```

### 5.2.4 Isotopic tracing of fossil fuel CO<sub>2</sub> uptake

The experiments on fossil fuel CO<sub>2</sub> emissions to the atmosphere include an assumed isotopic composition of the emitted carbon. For any of the carbon emissions experiments you have run (including the historical transient) – explore in the 3D output how the isotopic composition of fossil fuel carbon is propagated from the atmosphere into the ocean and through the ocean via its large-scale circulation. The variable you want to plot is: ocn\_DIC\_13C<sup>11</sup>. In this, it is helpful to also take a control experiment (you did run one, right ... ?) and create a difference map to better visualize how the  $\delta^{13}\text{C}$  patterns in the ocean evolve through time.

---

By default, fossil fuel carbon is tagged with a mean fossil fuel isotopic signature of  $-27\text{\textperthousand}$ . This is set by the parameter:

```
bg_par_atm_force_scale_val_4=-27.0
```

which scales the value of 1.0 specified in the file biogem\_force\_flux\_atm\_pCO2\_13C\_sig.dat in the *forcing* directory, which in the previous fossil fuel release experiments was the genie-forcing sub-directory: pyyyyz.FpCO2\_Fp13CO2

To create a more pronounced 'tag' (or tracer) of fossil fuel carbon, you could, for instance, make the assumed value of the CO<sub>2</sub> more negative, e.g.  $-60\text{\textperthousand}$  would be the signature of methane (natural gas). You could also push the value even more negative and consider it an idealized numerical tracer of fossil fuel carbon (note that the lowest value you are allowed to set is  $-999\text{\textperthousand}$ ).

---

<sup>11</sup>Also calculated and saved are the isotopic compositions of the 3 different aqueous carbonate chemistry components – carb\_d13C\_CO2, carb\_d13C\_CO32, and carb\_d13C\_HCO3.

### 5.3 Appendix

#### Relevant CO<sub>2</sub> and carbonate chemistry time-series data

Filename	Data	Application
biogem_series_*.res		
atm_pCO <sub>2</sub>	Global inventory ( <i>mol</i> ), mean concentration ( <i>atm</i> ) of atmospheric CO <sub>2</sub> .	Drivers of and feedbacks with climate. Diagnostic of response to carbon emissions (and removal).
atm_pCO <sub>2</sub> _13C	<sup>13</sup> C inventory ( <i>mol</i> ) and δ <sup>13</sup> C of atmospheric CO <sub>2</sub> .	Diagnostic of carbon emissions (and removal). Comparison with (terrestrial) proxy δ <sup>13</sup> C data.
carb_sur_conc_*	Carbonate chemistry components (mean surface) (mol kg <sup>-1</sup> ).	Not generally useful.
carb_sur_H	Surface ocean mean [H <sup>+</sup> ] (mol kg <sup>-1</sup> ).	More useful is pH – reported under misc (see below).
carb_sur_ohm_arg	Mean surface aragonite saturation.	Ocean acidification impacts of CO <sub>2</sub> release. Weathering impacts. Relates to carbonate production by (modern) corals, pteropods.
carb_sur_ohm_arg	Mean surface calcite saturation.	Ocean acidification impacts of CO <sub>2</sub> release. Weathering impacts. Carbonate production by foraminifera and coccolithophorids.
diag_misc_specified_forcing_*	Applied flux <i>forcings</i> (mol yr <sup>-1</sup> ).	Whenever a <i>restoring</i> , or <i>flux forcing</i> is specified, the actual flux employed, is saved here. Useful for diagnosing the flux associated with a restoring forcing (e.g. allowing emissions flux associated with RCP ( <i>restoring forcing</i> ) scenario to be diagnosed.)
misc_surpH	Mean surface <i>pH</i> .	Ocean acidification.
ocn_DIC	Global inventory ( <i>mol</i> ), mean global, surface, benthic concentrations (mol kg <sup>-1</sup> ) of DIC.	Carbon release and removal.
ocn_DIC_13C	Global inventory ( <i>mol</i> ), mean global, surface, and benthic δ <sup>13</sup> C.	Carbon release and removal.
ocn_ALK	Global inventory ( <i>mol</i> ), mean global, surface, benthic concentrations (mol kg <sup>-1</sup> ) of ALK.	'Ocean Alkalinity Enhancement' CDR.

Table 5.1: Summary of the main (useful, plus notes on a few less used) *time-series* output for (bio)geochemistry (non ecological) investigations.

**Relevant CO<sub>2</sub> and carbonate chemistry time-slice data**

<b>variable</b>	<b>variable (long name)</b>	<b>Description</b>	<b>Application</b>
carb_ben_ohm_arg		Benthic aragonite and calcite saturation.	Impacts of ocean acidification of distribution of benthic organisms. Indicator of sediment preservation.
carb_sur_ohm_cal			
carb_sur_ohm_arg		Ocean surface aragonite and calcite saturation.	Impacts of ocean acidification of distribution of planktic organisms.
carb_sur_ohm_cal			
fseaair_pCO <sub>2</sub>	pCO <sub>2</sub> : net sea->air gas exchange flux density	Air-sea CO <sub>2</sub> gas exchange.	Indicator of air-sea gas disequilibrium, regions of out-gassing/in-gassing.
misc_pH	ocean pH	Ocean surface pH.	Ocean acidification.
ocn_int_DIC	DIC water-column integrated tracer inventory	Pattern of water column integrated ocean DIC (i.e. dissolved carbon storage) ( $mol m^{-2}$ ).	Indicator of CO <sub>2</sub> emissions storage and transport when used in difference/anomaly maps and calculations.

Table 5.2: Summary of the main (mostly useful) 2D time-slice output for (bio)geochemistry.



Basic controls on biological productivity  
Comparing model vs. observations  
Temperature and the biological pump  
Iron co-limitation of biological productivity

Further ideas

Further analysis ideas  
Further experiment ideas  
Other thoughts and suggestions  
Biological pump background information  
Advanced experiment ideas ...

## 6. Ocean biogeochemical cycles

**READ.ME**

You will need the following *re-starts* files prior to embarking on the experiments in any of the listed Sections in this Chapter.

6.1 This is the basic single ( $PO_4$ ) nutrient biological export scheme configuration for playing with.

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/cookie.CB.p_worbe2.BASES.ridgwelletal.SPIN.tar.gz
```

6.2 A single ( $PO_4$ ) nutrient biological export scheme but in a higher (vertical) resolution ocean.

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/cookie.CB.p_worjh2.BASES.caoetal.SPIN.tar.gz
```

World Ocean Atlas modern ( $PO_4$ ) climatology (re-gridded/interpolated observations), re-gridded to the same ocean grid as EXAMPLE.worjh2.Caoetal2009.SPIN

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/worjh2.p_an.200709.nc
```

World Ocean Atlas modern ( $O_2$ ) climatology (re-gridded/interpolated observations), re-gridded to the same ocean grid as EXAMPLE.worjh2.Caoetal2009.SPIN

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/worjh2.o_an.200709.nc
```

6.3 This is a paired T-dependent and non-T-dependent export and remineralizaton configuration.

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/cookie.CB.p_worjh2.BASES.crichtonetal.STND.SPIN.tar.gz
```

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/cookie.CB.p_worjh2.BASES.crichtonetal.TDEP.SPIN.tar.gz
```

6.4 This is a scheme with T-dependent biological export (but with a non T-dependent, fixed remineralization profile) and with Fe co-limitation alongside P.

```
$ wget --no-check-certificate
http://www.seao2.info/cgenie_output/cookie.CB.p_worjh2.BASESFe.FeMIP.SPIN.tar.gz
```

6.5 (same *restart* as per for 6.4)

Extract the results in the usual way and in the usual place ... and return to genie-main in the usual way ... all ... as usual.

## The ocean's biological pump

In this Chapter we'll step through some of the facets of the cycle of carbon (and nutrients) in the ocean – the 'biological pump'. And then in a later section we'll look at the same processes but in a different way – through the lens of 'geoengineering', and hopefully learn something further about how everything 'works' regardless of the desirability and effectiveness (or not) of geoengineering.

**cookie** incorporates a variety of options for simulating biological export production. To date, these have all been rooted in what is effectively a nutrient mass balance approach to estimating export production, nicely encapsulated by Ernst Maier-Reimer as: "*conceptually not a model of biology in the ocean but rather a model of biogenically induced chemical fluxes [from the surface ocean]*" [Maier-Reimer, 1993]. Hence in schemes of this nature – which we will term 'biogenic flux' schemes – there is no attempt to explicitly account for changes in cell numbers/biomass and hence, nor zooplankton, which will impact a bias particularly in the seasonal time-dependent response of export. Previous biogenic flux schemes utilized in **cookie** considered a single nutrient (P limitation) only and took either restoring to observations [Cameron et al., 2005] or explicit P-limitation [Ridgwell et al., 2007a,b] approaches. More recently, additional limitation by iron (P+Fe) has been implemented, while nitrogen cycling (including N fixation and denitrification) and hence P+N limitation, has been implemented and explored in the context of extreme nutrient and oxygen cycle perturbation associated with the Cretaceous Oceanic Anoxic Events [Monteiro et al., 2012; Naafs et al., 2029]. Work currently in development includes the additional consideration of Si limitation (P+Si limitation) control on production of diatoms vs. non-diatoms following Ridgwell et al. [2002].

Overall, biological production and net export of particulate (POM) and dissolved (DOM) organic matter plus calcium carbonate ( $CaCO_3$ ) are directly determined by the availability of nutrients (phosphate and/or total dissolved iron and/or  $NO_3^{2-}$  (and/or  $NH_4^+$ ) and/or dissolved silica) together with the degree to which physical conditions, particularly light and temperature, are conducive to growth. DOM is remineralized (transformed back to inorganic dissolved constituents) relatively rapidly (a ca. annual time-scale) and hence mostly at or close to the ocean surface, while POM is assumed to sink down into the ocean interior. See: *Hülse et al. [2017]*<sup>1</sup> for a comprehensive review (esp. of implementation in models).

Note that a full description of the various biological and ocean interior remineralization schemes (together constituting the biological pump) is not given here as background. Rather, the relevant provided references should be read (rather than e.g. downloaded and immediately forgotten ...).

---

<sup>1</sup>Hülse, D., S. Arndt, J.D. Wilson, G. Munhoven, and A. Ridgwell, Understanding the causes and consequences of past marine carbon cycling variability through models, *Earth-Science Reviews* 171, dx.doi.org/10.1016/j.earscirev.2017.06.004 (2017).

What should you be thinking of looking at in terms of model output to understanding something about the role of the oceans biological pump in the Earth system (and climate dynamics)? Each subsequent exercise will direct you to some of the outputs to analyze/visualize that are directly relevant to the question (model experiment). In addition to that, what follows is a brief general over-view of relevant fields.

- Firstly, it is worth noting, if you have not already, the summary text file:  
biogem\_year\_yyyy\_yyy\_diag\_GLOBAL\_AVERAGE.res where yyyy\_yyy is the mid-point of the annual average year saved. Hence for the provided 10,000 year spin-up cookie.CB.p\_worbe2.BASES.ridgwelletal.SPIN, the file is:  
biogem\_year\_09999\_500\_diag\_GLOBAL\_AVERAGE.res.

Each of these files contains summary information associated with each *time-slice*. Model properties that you might pay particular attention to includes:

– ATMOSPHERIC PROPERTIES

Atmospheric pCO <sub>2</sub>	:	278.000 uatm
Atmospheric pCO <sub>2</sub> _13C	:	-6.500 ‰

is a useful place to find e.g. the final (annual average) value of atmospheric *pCO<sub>2</sub>* (rather than searching to the end of the relevant *time-series* file).

– BULK OCEAN PROPERTIES

Ocean DIC	..... :	2211.435 umol kg <sup>-1</sup> <-> 0.2977295E+19 mol
-----------	---------	--

reflecting how much carbon is stored in the ocean,

Ocean P04	..... :	2.154 umol kg <sup>-1</sup> <-> 0.2899897E+16 mol
-----------	---------	---

the ocean nutrient (here: phosphorous) inventory,

Ocean O2	..... :	221.288 umol kg <sup>-1</sup> <-> 0.2979245E+18 mol
----------	---------	---

is the average concentration of dissolved oxygen in the ocean (and its inventory).

– SURFACE EXPORT PRODUCTION

Export flux POC	:	204.253 umol cm <sup>-2</sup> yr <sup>-1</sup> <-> 0.7505843E+15 mol yr <sup>-1</sup>
Export flux CaCO <sub>3</sub>	:	28.624 umol cm <sup>-2</sup> yr <sup>-1</sup> <-> 0.1051875E+15 mol yr <sup>-1</sup>

fluxes reflecting the global biological export of particulate organic matter and calcium carbonate, and then

SURFACE EXPORT PRODUCTION

Export flux POC	:	204.253 umol cm <sup>-2</sup> yr <sup>-1</sup> <-> 0.7505843E+15 mol yr <sup>-1</sup>
Export flux CaCO <sub>3</sub>	:	28.624 umol cm <sup>-2</sup> yr <sup>-1</sup> <-> 0.1051875E+15 mol yr <sup>-1</sup>

is the residual flux that actually reaches the sea-floor.

Finally, there is a summary of the summary, which is often the most useful of all:

– SURFACE EXPORT & SEDIMENT DEPOSITION (RAIN) FLUX SUMMARY

Total POC export	:	0.7505843E+15 mol yr <sup>-1</sup> = 9.007 PgC yr <sup>-1</sup>
------------------	---	---

Total CaCO <sub>3</sub> export	:	0.1051875E+15 mol yr <sup>-1</sup> = 1.262 PgC yr <sup>-1</sup>
--------------------------------	---	---

Total POC rain	:	0.8654062E+14 mol yr <sup>-1</sup> = 1.039 PgC yr <sup>-1</sup>
----------------	---	---

Total CaCO <sub>3</sub> rain	:	0.5191799E+14 mol yr <sup>-1</sup> = 0.623 PgC yr <sup>-1</sup>
------------------------------	---	---

where carbon fluxes are also given in helpful and literature-friendly units of PgCyr<sup>-1</sup>.

- Secondly, there are a number of biological pump related *time-series* outputs, e.g.:

biogem\_series\_fexport\_POC.res  
biogem\_series\_fexport\_POPres  
biogem\_series\_fexport\_CaCO3.res

are the *time-series* of the carbon and phosphorous of particulate organic matter export, and that for  $CaCO_3$ , and

biogem\_series\_ocnsed\_POC.res  
biogem\_series\_ocnsed\_POPres  
biogem\_series\_ocnsed\_CaCO3.res

are the corresponding fluxes at the sea-floor. The *time-series* output:

biogem\_series\_ocn\_PO4.res

will tell you how what the global annual average nutrient concentration is (left) at the surface, and

biogem\_series\_ocn\_O2.res

the average concentration dissolved oxygen in the ocean and at the sea-floor.

- Thirdly, as always – **netCDF** outputs.

- In the 2D **netCDF**:

bio\_export\_POC  
bio\_export\_POP  
bio\_export\_CaCO3

are spatial fields of the export flux of *POC*, *POP*,  $CaCO_3$ , and variables with *focnsed* in place of *bio\_export* in its name the flux distributions to the sea-floor.

Derived from these:

misc\_sur\_rCaCO3toPOC  
misc\_sur\_rPOCtoPOP

are the ratios of  $CaCO_3/POC$  and  $POC/POP$ , respectively.

There are then some fields for surface and benthic tracer concentrations, such as of phosphate and dissolved oxygen.

- In the 3D **netCDF**, *bio\_\** are the 3D spatial distributions of particles setting down through the water column – *bio\_fpart\_\** is the particulate flux density, *bio\_fparttot\_\** the total flux associated with each grid point, and *bio\_fparnorm\_\** are the fluxes in the water column normalized to the export flux out of the base of the surface ocean layer.

Also see spatial distributions of dissolved carbon, nutrients, and oxygen etc – *ocn\_\**.

From just the *spin-up* experiment results you can explore and/or plot some of these fields. Or better – construct a control experiment following on from the provided *restart* and analyze/explore that.

## 6.1 Basic controls on biological productivity

The following exercises will utilize a very basic (but relatively uncomplicated and fast) representation of the biological pump – one with only a single nutrient ( $\text{PO}_4$ ) potentially limiting to biological export, considered. The scheme is described in full in *Ridgwell et al.* [2007]<sup>2</sup>. Note that this specific configuration of the (modern) ocean model accounts only for 8 layers in the ocean (but spanning the same 0 – 5000 water depth range as the 16-level model configuration used in evaluating AMOC stability). Although the limited vertical resolution of the 8-level model places additional constraints on the ability of the model to reproduce features such as oxygen minimum zones in the ocean, it does enable a much shorter run-time (via a longer time-step in the model as well as calculations only needing to be carried out for about 50% of the total number of 16-level ocean cells). Hence we are making trade-off between model fidelity (in reproducing observed distribution of tracers in the modern ocean) and speed (and hence the ability to more effectively 'play' with the Earth system in a relatively short amount of real time).

We will consider the following scientific questions as a starting point, and then devise some model experiments to address them:

1. What would the global carbon cycle look like with no biological production in the ocean? Or alternatively: how important is the biological pump in controlling atmospheric  $p\text{CO}_2$  and how much higher would  $p\text{CO}_2$  be today in the absence of an active marine biosphere?
2. Conversely, is the biological pump operating at its maximum efficient in the ocean today, and if not, how much lower would atmospheric  $p\text{CO}_2$  be if virtually all nutrients at the surface were consumed?
3. Associated with #1 and #2 – how does changing the biological pump affect the distribution and hence availability of dissolved oxygen in the ocean? (How more oxygenated would the ocean be with no export of organic matter into the ocean interior, and how much more severe would oxygen depletion be if plankton were able to utilize all the nutrients at the ocean surface?)
4. If carbon and nutrients are returned back to solution (remineralized) much closer to the ocean surface, does atmospheric  $p\text{CO}_2$  increase or decrease??? (The faster return of DIC to the ocean surface will tend to increase  $p\text{CO}_2$  while the return of nutrients will tend to enhance biological productivity and decrease  $p\text{CO}_2$  ... so the answer is not necessarily intuitive and hence why we build and run computer models.)
5. Conversely, what if carbon and nutrients are much more efficiently exported into the very deepest parts of the ocean – does  $p\text{CO}_2$  increase or decrease (and what happens to e.g.  $[O_2]$ )?
6. What role does the calcium carbonate ( $\text{CaCO}_3$ ) 'counter pump' play? Would  $p\text{CO}_2$  be higher or lower (and by how much) in the absence of calcifying organisms and the production of  $\text{CaCO}_3$  in the ocean?
7. How sensitive is the cycling of carbon and nutrients and hence the pattern of ocean oxygenation to changes in the large-scale circulation of the ocean?. For instance, would  $p\text{CO}_2$  be higher or lower (and by how much) with a weaker AMOC?
8. What about feedbacks with climate – how does changing global surface temperature (and hence ocean circulation patterns and sea-ice extent) affect biological productivity and the biological pump in the ocean? (Is this a positive or negative feedback on climate change?)

---

<sup>2</sup>Ridgwell, A., et al., Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling, *Biogeosci.* 4, 87-104 (2007).

Experiments modifying the biological pump in the ocean would ideally be run for a few thousand and perhaps as long as 5000 years in order to obtain a complete new (quasi) steady state of carbon and nutrient cycling in the ocean. However, the main impacts in the ocean and on atmospheric  $pCO_2$  often tend develop relatively rapidly (decades). Hence model experiments could be run for a few 10s or perhaps a few 100 years to see something 'happen'. A full 5000 years can also be run, and you can periodically download and viewing the results as the experiment progresses and not necessarily wait until the end of the experiment to determine a good answer. As always – if possible – run a representative (or perhaps the most extreme, such as shutting off the biological pump entirely) experiment for 5,000 or even 10,000 years, and answer for yourself how long different key facets of the system (ocean circulation, global export production, mean ocean  $[O_2]$ , etc.) take to adjust. Don't forget to submit your jobs to the cluster!

All the experiments in this first section will start from the same *re-start*<sup>3</sup> and are based on the same *user-config*<sup>4</sup>. To run an e.g. 100 year experiment (but it need not be this – see above) using the template *user-config* and provided *re-start* would look like:

```
./runcookie.sh cookie.CB.p_worbe2.BASES LABS LAB.6.1.EXAMPLE 100
cookie.CB.p_worbe2.BASES.ridgwelletal.SPIN
```

(When you run new experiments based on this, remember to copy and rename the provided *user-config*LAB.6.1.EXAMPLE in order to create new and unique experiments each time.)

Whatever you plan to do re. perturbation experiments to explore the working of the ocean's biological pump, the first thing to do, is to run a **control experiment** for however long you have decided to run all the actual experiment experiments for. (You need not wait for this to finish, but can submit it as a *job* to the cluster and get on with running some real experiment experiments.) The control experiment can simply be derived (copied) from LAB.6.1.EXAMPLE, with no further alterations in parameter values needed. In this *user-config*, the value of atmospheric  $pCO_2$  is not prescribed (there is no *forcing* defined) and hence  $pCO_2$  is free to respond to any change in parameters. In the case of the control – there are no changes in the biological pump parameters compared to the *spin-up*, so you should see no (or little) drift in  $pCO_2$  occurring. A substantive drift in the control experiment (e.g. 10s of ppm of  $CO_2$  in the atmosphere over 100 years) is your way of knowing that you have either used the wrong *re-start* or have accidentally modified key parameters in your control experiment *user-config* compared to those used to generate the *re-start*.

Plan in advance the output you want to view and make sure it is going to appear(!) In the *user-config* – check that the results/output select 'level' is going to give you the output you are expecting. In the example *user-config*, this is specified by:

```
bg_par_data_save_level=7
```

(Refer to Chapter 13 for more details on how different (2D and 3D) *time-slice* variable fields and *time-series* are selected to be included in the model output.)

Also note that in the example *user-config* provided, netCDF time-slices are only saved at the very end (final annual average) of the model experiment, regardless of how long that is:

```
bg_par_infile_slice_name='save_timeslice_NONE.dat'
```

You may want to request more frequent saving of the spatial fields – see Section 13.2 (and 13.3).

In the example *user-config*, *time-series* saving is set to every year.

<sup>3</sup>As per described in Ridgwell et al. [2007].

<sup>4</sup>The same *user-config* as used to generate the *re-start* (with the exception that the atmospheric  $CO_2$  concentration is no longer prescribed).

In terms of the specific 'questions' at the start (the numbering system is the same in the following), the parameters and parameter values to change and well as what to 'look for' are:

1. How to kill the biological pump in the ocean?

The parameter scaling the rate of nutrient uptake (and hence biological export) is:

```
# maximum rate of conversion of dissolved PO4 into
# organic matter by phytoplankton (mol kg-1 yr-1)
bg_par_bio_k0_P04=1.9582242E-06
```

(in units of  $\text{mol PO}_4 \text{ kg}^{-1} \text{ yr}^{-1}$ ). Setting this to zero would 'turn off' completely the biological pump, leaving you with an abiotic ocean. Alternatively you could initialize the ocean with a zero phosphate concentration.

What to look for? The value of atmospheric  $p\text{CO}_2$ . You might also confirm that biological export really is zero (so check the export fluxes).

To obtain these values you can refer to the respective time-series results files. Or, the summary text file: biogem\_year\_yyyy\_yyy\_diag\_GLOBAL\_AVERAGE.res where yyyy\_yyy is the mid-point of the annual average year saved. Each of these files contains summary information associated with each *time-slice*, including mean atmospheric composition, mean ocean composition, biological export fluxes. The same information can be found in various *time-series* files, but sometimes it is simply easier to find it here in one place!

2. Maxing out the biological pump?

Similar to above – nutrient uptake and the strength of the biological pump can be enhanced by increasing the value of the parameter `bg_par_bio_k0_P04` (maybe 10 times larger?). The question is then what does the ocean and atmospheric  $p\text{CO}_2$  look like if the biological pump is operating at its maximum rate and (almost) all nutrients are consumed at the surface.

You may find regions of nutrients that still never get fully consumed ... why?

3. For the question on dissolved oxygen ( $[O_2]$ ) availability in the ocean interior – the experiments are the same as #1 and #2 (above), except you are looking for how the mean and distribution of dissolved oxygen in the ocean changes. Mean ocean  $[O_2]$  values can be obtained from time-series of the summary file. Spatial distributions are recorded in the *netCDF* output. You might view horizontal or vertical slices (from 3D), or benthic distributions (2D). Think about where animals tend to live and where they could potentially be impacted if export from the ocean surface was much higher than 'today' (i.e., compared to the *spin-up*, or ideally, the same year in your control).

4. What would the ocean look like if organic matter was remineralized much closer to the ocean surface?

The parameter controlling the depth-scale (and vertical distribution) at which particulate organic matter (POM) is remineralized in the ocean interior is controlled by:

```
bg_par_bio_remin_P0C_eL1=550.5195
```

(units of  $m$ ). Reducing the value of this parameter forces a greater proportion of sinking POM to be remineralized closer to the surface, while a larger value pushes particulate organic matter (and associated carbon and nutrients) deeper down on average into the ocean interior.

What to look for? Largely as before – the summary file is useful for atmospheric  $CO_2$ , mean ocean  $O_2$ , and global (biological) export fluxes (as you might expect nutrients released closer to the ocean surface to result in greater surface ocean nutrient supply and hence fuel higher export). You might also consider the spatial patterns of nutrients and dissolved oxygen.

See Meyer *et al.* [2016]<sup>5</sup> for an example of a study testing these sort of model changes and its implications (and hence ideas of what else to look for, and for what a 'reasonable' parameter value to test might be).

5. Conversely, you can also increase the value of the depth scaling parameter in order to force a deeper mean depth of POM remineralization.

Alternatively – you can partition more of the total export into a POM form that is assumed to be resistent to degradation and is transported to the ocean floor completely unchanged (see: Ridgwell *et al.* [2007]):

```
bg_par_bio_remin_POC_frac2=6.4591110E-02
```

Increasing this value forces a greater fraction of the POM exported from the surface ocean to reach the ocean floor.<sup>6</sup>

In both (4) and (5), it may not be obvious which way atmospheric  $pCO_2$  responds – e.g. for (4) by returning nutrients more efficiently to the ocean surface, you increase export, and hence increase the fixation and removal of atmospheric  $CO_2$  from the ocean surface, drawing down atmospheric  $pCO_2$ . BUT, at the same time, carbon released from  $POC$  (as  $DIC$ ) through bacterial remineralization, also occurs at a shallower depth and is returned to the surface more efficiency ... tending to increase atmospheric  $pCO_2$  ... Just this problem – 2 opposing effects with an uncertain net impact – is exactly why (to find out the answer) you build and run numerical models of the system!

6. The export of  $CaCO_3$  from the ocean surface is calculated as a ratio to the export of  $POC$ . By default, this ratio varies with the carbonate chemistry (saturation state) of the surface ocean, following Ridgwell *et al.* [2007a] (and see also Ridgwell *et al.* [2007b] and Ridgwell *et al.* [2009]).

The scaling parameter controlling the molar ratio of  $\frac{CaCO_3}{POC}$  is:

```
bg_par_bio_red_POC_CaCO3=0.044372
```

Setting this zero will effectively turn back the clock 200Ma to a world prior to the evolution of planktic calcifiers (e.g. see: Ridgwell [2005]), or conversely, a future world in which they are all driven extinction from ocean acidification ...

What is the impact on atmospheric  $pCO_2$  of setting this to zero? From your knowledge of carbonate chemistry ... why does this happen? You might try e.g. doubling the value (and hence doubling the export of  $CaCO_3$  from the surface ocean). As well as  $pCO_2$ , you might also look at fields of  $ALK$  (or other carbonate chemsity parameters) in the ocean to see what other geochemical changes occur.

Because (rightly or wrongly) in the model,  $CaCO_3$  depends on surface ocean  $\Omega$  (w.r.t. calcite) you might explore what happens to  $CaCO_3$  production and export under a release of (fossil fuel)  $CO_2$  – simplest here is to use the same re-start, and create and add a forcing to the *user-config* to implement a large pulse or continuous  $CO_2$  flux to the atmosphere

---

<sup>5</sup>Meyer, K.M., A. Ridgwell, and J.L. Payne, The influence of the biological pump on ocean chemistry: implications for long-term trends in marine redox chemistry, the global carbon cycle, and the evolution of marine animal ecosystems, Geobiology, DOI: 10.1111/gbi.12176 (2016).

<sup>6</sup>See Ridgwell et al. 2007] for a description of how this all works.

7. While it is possible to add a freshwater 'hosing' forcing (as used in an earlier Chapter/exercise) to modify large-scale ocean circulation, there is a simpler way just to 'turn off' the AMOC in the standard modern continental configuration of **cookie**.

The consequence of using the simplified **EMBM** atmosphere in conjunction with no topography over land, is that the net moisture transport from the Atlantic to the Pacific in the real world, is not reproduced. (This net transport is primarily a consequence of the blocking of Westerly transport moisture across the North American continent by e.g. the Rockies, while low latitude Trade Winds travel relatively topographically unrestricted through Central America, meaning that moisture transport from the Pacific to North Atlantic is partly blocked, while the reverse lower latitude transport is now.) To correct for this, there is a built-in prescribed moisture transport (technically, a negative salinity transport), principally from the North Atlantic to the North Pacific. In the model, there is a scaling parameter for the magnitude this transport ... which can be set to zero ... hence removing the prescribed moisture transport and likely leading to a collapse of the AMOC.

To kill (collapse) (hopefully) the AMOC in the modern continental configuration of **cookie**, set:

`ea_28=0.0`

Stuff to look for – in addition to how the AMOC changes (e.g. plot the stream-function) how do the patterns of  $[PO_4]$  and  $[O_2]$ , particularly in the deep North Atlantic change? What impact does this have on global export production, and ultimately on atmospheric  $pCO_2$ ?

8. Finally, you might explore how in some (or all) of the above experiments, a changing climate in response to changing atmospheric  $pCO_2$ , in turn modulates the impacts through carbon-climate feedback.

By default in the *user-config*, climate- $CO_2$  feedback is disabled:

`ea_36=n`

i.e. changing atmospheric  $pCO_2$  does not influence the climate system (which remains implicitly forced with a preindustrial radiative forcing value). This allows us to avoid complications arising due to carbon-climate feedback and hence obtain a the underlying response of the system to changing the biological pump in the ocean.

You can re-enable climate feedback by setting:

`ea_36=y`

Impacts may include (but not be limited to): ocean stratification (rapid warming) and/or changes in convection at high latitudes, increased/decreased sea-ice extent that affects the ocean surface area available for biological export, changing solubility of oxygen in sea-water. Note that in this particular simple biological export scheme, there is no temperature-dependence of biological activity and hence export.

And that pretty much wraps up all the main knobs that it is possible to play with using the very basic biological pump scheme in **cookie**. (However, a few further ideas follow ...)

## 6.2 Comparing model vs. observations

Key to having any 'confidence' in both past and future biogeochemical (e.g. carbon, nutrient, oxygen) cycling and climate dynamics and sensitivity to perturbation in a model, is having confidence in the model's ability to adequately reproduce the relevant features of the modern ocean in some direct comparison made with observations. Note that it is not necessarily critical that every single feature in the modern ocean is faithfully accounted for – the degree to which the model needs to match observations is going to depend on the question. For example, if a model is only needed for making first order projections of changes in atmospheric  $pCO_2$  or  $pO_2$ , the details of the structure of biogeochemical cycling in the ocean may be less important, as long as the gross partitioning of carbon between deep ocean vs. surface and atmosphere, or bottom water oxygenation and carbon flux to the sediments, respectively, is reasonable. Other questions such as involving denitrification may require more explicit details of the distribution and intensity of oxygen minimum zones (OMZs) to be able to be reproduced in the modern ocean. However, even in the latter case, if structurally in the model OMZs tend to be structurally (e.g. as a property of the fundamental ocean grid and physics and/or biases in circulation) too weak (too high  $[O_2]$ ), it is always possible to adjust denitrification (in this example) to occur at the 'correct' rate by changing parameter values (e.g. of a minimum  $[O_2]$  value at which denitrification occurs). As long as the same structural biases are present in paleo simulations, there is no reason to believe that projected past ocean denitrification rates would be incorrect. (In fact, this adjusting of parameter values to correct for some bias and achieve an appropriate rate of some process, is ubiquitous throughout all Earth system modelling, including atmospheric physics and particularly cloud formation.)

---

As an example/practice in model-data comparison, the output of a higher vertical resolution ocean model simulation (16 vs. the previous 8 levels in the ocean) – cookie.CB.p\_worjh2.BASES.caoetal.SPIN – is provided here (see **READ.ME** downloads) along with observations of the distributions of  $[PO_4]$  and  $[O_2]$  in the modern ocean. The latter (observations) are interpolated and re-gridded to a high (1 degree in lon and lat) resolution grid and available as part of the World Ocean Atlas series of ocean climatologies (both physics and geochemical), and then re-gridded to the same **cookie** ocean grid as the provided (16 ocean level) experiment (which is the spin-up of *Cao et al.* [2009]).

While the comparison can be made more formally and statistically (with a little **MATLAB** or **python** code), much can be gained visually (and indeed this is typically how model-data comparisons are made in the literature). Fortunately, this visual comparison can be made in **Panoply** and both re-gridded observations and model output, being **netCDF** format, can be loaded into **Panoply**. When making visual comparisons, make sure that you use the same scale limits in both plots (these can be set manually). You can also create difference plots of model minus observations. For difference plots – be careful when changing depth or longitude slices, that you change to the same slice in both model and observations. Note that for difference plots the convention is to use a color-scale that goes from blue (negative) to red (positive) through white (or sometimes, a yellow). Set the magnitude of the scale limits equal so that zero (no difference between model and data) maps onto white (or yellow).

You may (and ideally should!) have ideas/thoughts on 'why' the model and observations diverge where they do. These thoughts may well involve processes in the ocean, e.g. is the dissolved oxygenation feature x due to insufficient or excessive organic matter remineralization in the ocean, or is the surface dissolved phosphate concentration in region y excessive/over-depleted due to insufficient/excessive biological productivity? Perhaps the most fundamental purpose or advantage

of models is that they provide you a way to test and attempt to answer such questions (which are basically questions of understanding global biogeochemical cycles in the first place and how tracer patterns in the ocean arise). So given a hypothesis for model-data mismatch, you might return to the previous section and see if a model parameter is listed that could be adjusted in value to test your hypothesis (e.g., changing remineralization depth or changing a control on biological export might improve or worsen the model-data fit – often it can improve in one location and degrade elsewhere).

Note that the first section of this chapter uses a faster 8-level ocean model configuration, whereas the model-data comparisons are on the basis of a 16-level ocean and so the results of testing parameter changes in an 8-level ocean should be regarded as qualitatively rather than quantitatively comparable.

---

If you wish to play with (modify and test hypotheses etc.) this configuration, and/or run a control experiment to make the comparisons with observations with, an example *user-config* (with fixed climate and no  $CO_2$ -climate feedback) is provided as: LAB.6.2.EXAMPLE and can be run"

```
./runcookie.sh cookie.CB.p_worjh2.BASES LABS LAB.6.2.EXAMPLE 100  
cookie.CB.p_worjh2.BASES.caoetal.SPIN
```

### 6.3 Temperature and the biological pump

So far, both 8- (non-seasonal) and 16- (seasonally forced) level configurations have not account for the role of temperature in biological (metabolic) activity and rates of carbon transformation.

A pair of example configurations are provided for you as per published in *Crichton et al. [2021]*<sup>7</sup>.

1. The *user-config* for the non temperature-dependent ('standard') control configuration of *Crichton et al. [2021]*:

LAB.6.3.STND

and the resulting *spin-up* which you can use as a *re-start*:

cookie.CB.p\_worjh2.BASES.crichtonetal.STND.SPIN

Note that this configuration is basically the same as in *Cao et al. [2009]*.

2. The *user-config* for the temperature-dependent (both biological uptake and water column remineralization) configuration of *Crichton et al. [2021]*:

LAB.6.3.TDEP

and the corresponding spin-up:

cookie.CB.p\_worjh2.BASES.crichtonetal.TDEPSPIN

For both, the *base-config* you need is the same: cookie.CB.p\_worjh2.BASES

---

One thing to do is simply to explore how the biogeochemical cycling (export flux magnitudes and patterns, nutrient and dissolved oxygen distributions at the surface and down through the water column) differs. You might then apply a warming perturbation by adjusting the scaling value for atmospheric  $pCO_2$  in the prescribed restoring *forcing*:

```
# *** FORCINGS ****  
...  
bg_par_atm_force_scale_val_3=280.0E-06
```

and see what 'happens' (contrast the response of non T-dependent vs. T-dependent configurations under the exact same perturbation). **And then ... try and deduce why whatever has happened, has happened ...**

Note that in these *user-configs*, atmospheric  $pCO_2$  is restored to a given value and this controls climate (rather than was the case with *ea\_36=n*).

---

<sup>7</sup>Crichton, K. A., J. D. Wilson, A. Ridgwell, P. N. Pearson, Calibration of temperature-dependent ocean microbial processes in the cGENIE.muffin (v0.9.13) Earth system model, GMD 10.5194/gmd-14-125-2021 (2021)

You could also swap out the fixed prescribed atmospheric  $pCO_2$  restoring forcing:

```
# specify forcings -- generic forcing of atmospheric pCO2 and d13C
# NOTE: the original paper used 280 ppm
bg_par_forcing_name="pyyyz.RpCO2_Rp13C02"
bg_par_atm_force_scale_val_3=280.0E-06
bg_par_atm_force_scale_val_4=-6.5
```

with an emissions forcing as per you user previously exploring the impact of fossil fuel  $CO_2$  release.

For example, from Chapter 5, the basic emissions flux forcing was:

```
# specify forcings -- flux of CO2 to atmosphere
# NOTE: bg_par_atm_force_scale_val_3 == scaling of units from mol yr-1 to PgC yr-1
# NOTE: bg_par_atm_force_scale_val_4 == carbon isotopic composition of CO2 emissions
bg_par_forcing_name="pyyyz.FpCO2_Fp13C02"
bg_par_atm_force_scale_val_3=8.3333e+013
bg_par_atm_force_scale_val_4=-27.0
```

which you could paste into the *user-config* (in place of the restoring forcing) and modify as before for a more exciting perturbation of the system.

Experiments to configure and run that you might like to consider are:

1. Drive both 'standard' and T-dependent configuration with the same emissions scenario.  
This contrasts the response, including climate feedback, of 'standard' and T-dependent scheme and gives you some quantitative sense of the important of T-dependent processes in the ocean and the strength and sign of their associated feedbacks.
2. Run both the above pair of experiments with fixed climate (radiative forcing), which for the assumption made in the paper of 280.0 ppm  $pCO_2$  would require:

```
ea_36=n
ea_radfor_scl_co2=1.007
```

(where  $280/278 = 1.007$ )

This would help you isolate the strength and sign of the feedback of both systems (and  $pCO_2$ ) with climate.

3. Strictly ... you would run 4 controls – one for both 'standard' and T-dependent and for both fixed and  $CO_2$ -responsive climate. All 4 control experiments would be run with no carbon emissions.

What to look out for?

- Changes in  $pCO_2$  between standard and T-dep configurations and with and without carbon-climate feedback (ea\_36) enabled.
- Changes in biological export between standard and T-dep configurations in response to.
- Latitude-vertical slices of dissolved  $PO_4$  and  $O_2$  for the T-dep experiments would reveal how nutrient cycling and the intensity and depth of OMZs might respond to a warming climate.

## 6.4 Iron co-limitation of biological productivity

"Iron ( $Fe$ ) is a "micronutrient"; essential to the biochemistry of all cells and, in particular, for enzymatic activities associated with photosynthesis and with nitrogen fixation in the marine environment. Yet it is only required in very small quantities, as little as one part in 2000 compared to phosphorus or one in 200,000 compared to carbon!"

As with macronutrients such as phosphate ( $PO_4^{2-}$ ) and nitrate ( $NO_3^-$ ), supply of  $Fe$  to the surface ocean occurs through up-welling and mixing of ocean waters from below. However, dissolved  $Fe$  has a short lifetime in the oxygenated seawater environment.  $Fe^{II}$ , the most soluble state, is rapidly oxidized to  $Fe^{III}$ , which is highly insoluble and tends to precipitate out and be removed (scavenged) by particulate matter settling through the water column. The result is that in up-welling water, the ratio of dissolved  $Fe$  to that of highly soluble  $PO_4^{2-}$  is lower than the ratio required by phytoplankton cells to grow and divide. Consequently, phytoplankton in surface waters cannot fully utilize the abundant up-welled phosphate (or nitrate) unless more  $Fe$  is brought into the system.

Dust is important because mineral aerosols contain iron, primarily in the form of  $Fe$  oxides such as hematite and oxide-hydroxides such as goethite (found as coatings on other mineral grains). However, the present-day flux of aeolian  $Fe$  is not everywhere sufficient to correct the relative nutrient imbalance. For instance, dust fluxes to the Southern Ocean are among the lowest anywhere on Earth, and the aeolian  $Fe$  supply is too small to compensate for the depleted  $Fe$  relative to  $PO_4^{2-}$ . Consequently, phytoplankton cannot fully utilize available macronutrients, and dissolved  $PO_4^{2-}$  exists year-round at the ocean surface. Similar reasoning applies to the presence of excess  $PO_4^{2-}$  in the Eastern Equatorial Pacific. Although dust supply to the North Pacific often appears moderately high in global model simulations, fertilization experiments in the Northwest Pacific suggest that this region is still iron-limited. Hence, the geography of complete nitrate utilization at the surface will be controlled to a first order by the mass distribution of dust deposition."<sup>8 9</sup>

---

<sup>8</sup>Ridgwell, A., The Global Dust Cycle, in Surface Ocean–Lower Atmospheres Processes, Eds. C. Le Quéré and E. S. Saltzman, AGU Geophysical Monograph Series, Volume 187, 350 pp.

<sup>9</sup>Also see: Jickells, T. D., Z. S. An, K. K. Andersen, A. R. Baker, G. Bergametti, N. Brooks, Cao J. J., P. W. Boyd, R. A. Duce, K. A. Hunter, H. Kawahata, N. Kubilay, J. laRoche, P. S. Liss, N. Mahowald, J. M. Prospero, A. J. Ridgwell, I. Tegen, and R. Torres, Global Iron Connections Between Desert Dust, Ocean Biogeochemistry and Climate, *Science*, 308, p. 67 (2005).

All experiments in this section start from the same *re-start* and are based on the same *user-config*<sup>10</sup>:

```
./runcookie.sh cookie.CB.p_worjh2.BASESFe LABS LAB.6.4.EXAMPLE
100 cookie.CB.p_worjh2.BASESFe.FeMIP.SPIN
```

This is effectively the iron cycle configuration used and evaluated in *Tagliabue et al.* [2016]<sup>11</sup>.

The *forcing* used in both the experiment includes a prescribed dust flux to the ocean surface (the Mahowald part of the directory name string). This is necessary because the model configuration you are using includes a co-limitation of biological productivity by iron (*Fe*) in addition to phosphate (*PO<sub>4</sub>*). (The files associated with the dust forcing are: *biogem\_force\_flux\_sed\_det\_sig.dat* and *biogem\_force\_flux\_sed\_det\_SUR.dat* but you do not need to edit these files.) **BUT** – note that no prescribed value of atmospheric *pCO<sub>2</sub>* is given in the *forcing* for this *user-config* (although it was for generating the *re-start*) – this is so that any change you make to the marine iron cycle and hence to the biological pump in the ocean can be seen as an impact on atmospheric *pCO<sub>2</sub>*. This also helps make for a better control experiment, because if the *re-start* was somehow mis-matched with the *user-config* or you accidentally changed something you did not mean to (or notice), you could expect to immediately see atmospheric *pCO<sub>2</sub>* starting to drift. (Also note that the experiments need not be run for 100 years ... you may want longer or shorter, depending on how you are perturbing the model and what exactly you are 'looking for'.)

What to change? (aka, what experiment/playing around to do?)

1. Well ... firstly, you might check that biological productivity is in fact limited anywhere. You can try 2 different things:

- (a) Reduce the half-saturation constant for *Fe* limitation to zero (so phytoplankton growth is never limited by iron, at least, up to the point where it completely runs out):
 

```
# [Fe] M-M half-sat value (mol kg-1)
bg_par_bio_c0_Fe=0.10E-09
```

 and set this parameter to zero (or something very very small).

- (b) Flood the ocean with so much iron that it simply cannot be limiting anywhere.

The dust field itself is sort of difficult to do anything with (in terms of manipulations), so instead, you can change the assumed solubility of iron in dust – i.e. what fraction of iron delivered to the surface ocean in dust is assumed to dissolve and hence become bio-available. The parameter controlling solubility<sup>12</sup> is:

```
# aeolian Fe solubility
bg_par_det_Fe_sol=0.00291468
```

And change to ... ? Maybe try an order of magnitude (or more) higher value.

(The second modification is likely to have more impact and come closer to answering the question: what would export production and atmospheric *pCO<sub>2</sub>* look like if there was no iron limitation on biological productivity? In the first option, you might still completely run out of iron somewhere in the ocean and hence limit biological productivity.)

<sup>10</sup>Remember to copy and rename the file LAB.6.4.EXAMPLE in order to create new and unique experiments each time.

<sup>11</sup>Tagliabue, A., O. Aumont, R. DeAth, J.P. Dunne, S. Dutkiewicz, E. Galbraith, K. Misumi, J.K. Moore, A. Ridgwell, E. Sherman, C. Stock, M. Vichi, C. Völker, and A. Yool, How well do global ocean biogeochemistry models simulate dissolved iron distributions?, GBC DOI: 10.1002/2015GB005289 (2016).

<sup>12</sup>Note that this is a fractional solubility, not a % solubility, so a value of 0.001 equates to a 0.1 wt% solubility.

2. A second question/investigation (also utilizing a copy of *user-config LAB.6.4.EXAMPLE*) might be to explore the internal cycling of iron and hence the importance of dissolved iron supplied 'from below' via ocean up-welling and mixing, rather 'from above' and directly from dust.

The dissolved *Fe* content of the ocean interior is set by: ocean transport (iron from elsewhere), iron released through the remineralization of organic matter, and iron scavenged from solution and removed onto sinking particles. Of these, you have direct control over *Fe* scavenging, via:

```
# modifier of the scavenging rate of dissolved Fe
bg_par_scav_Fe_sf_POC=1.338130
```

that scales the rate at which *Fe* is removed from solution (this rate also depends on the sinking flux of particulate organic matter as well as the concentration of 'free' iron (not bound to organic ligands and hence assumed protected from scavenging).

Increasing this value will increase the loss rate of *Fe* from the ocean interior (and surface), while decreasing it will reduce the rate of loss. Consider an order of magnitude<sup>13</sup> change in value (in either direction) to create meaningful changes to the marine iron cycle.

What to look for? Obviously, start with *pCO<sub>2</sub>* (*time-series*) and *POC* export (which can be viewed as both *time-series* and spatially in the 2D **netCDF** output). (All of these are also included in the summary output files.) Also 2D distributions of *PO<sub>4</sub><sup>2-</sup>* and *Fe* are highly relevant and useful. For more involved outputs, also in the 2D **netCDF** output you can find the pattern of iron solubility (*misc\_sur\_Fe\_sol*) and the flux of dissolved iron to the ocean surface (*misc\_sur\_fFe\_mol*).

Also, in the 2D **netCDF** output – *misc\_sur\_PO4Felimbalance* – is a simple attempt to illustrate where *PO<sub>4</sub><sup>2-</sup>* (positive values) rather than *Fe* limitation of biological productivity occurs. For example, high positive values (on scale up to 1.0) can be found in the Equatorial Atlantic and Indian Oceans, suggesting plenty of dissolved iron but no phosphate. The Southern Ocean is more iron than phosphate limited, while the Pacific gyres (most negative values) are dominantly iron limited.

<sup>13</sup>It need not be any more than this. In fact, you might try a slightly smaller change, e.g. factor-5.

## 6.5 Further ideas

### 6.5.1 Further analysis ideas

In addition to the primary environmental properties of atmospheric  $pCO_2$  (and climate if you have the feedback enabled), biological export production (and hence the patterns and magnitude of food sources for marine ecosystems), and dissolved oxygen, you might also look at how the patterns of carbon isotopes ( $\delta^{13}C$ ) change in the ocean.

How did the different changes in the biological pump that you tested, alter (if at all) the patterns of  $\delta^{13}C$  in the ocean? Can you distinguish between the different biological pump changes, based on  $\delta^{13}C$ ? (Carbonate and organic carbon  $\delta^{13}C$  is a key paleoceanographic proxy and one people would ideally like to use in order to reconstruct changes in the past such as in the biological pump.)

### 6.5.2 Further experiment ideas

By default, the update and export in organic matter, of carbon, occurs in a fixed 'Redfield' ratio, with  $PO_4^{2-}$ . The classical value is 106 (i.e., for every mole of  $PO_4^{2-}$  taken up from the ocean, 106 moles of  $CO_2$  are removed (and fixed into organic matter). The parameter determining this is:

`bg_par_bio_red_POP_POC=106.0`

To change the default value (106.0), simply add a new line at the end of the *user-config* file specifying the value you want. A larger number means that  $PO_4$  is being utilized more efficiently and more organic matter is being produced for the same nutrient consumption.

You should see impacts on atmospheric  $pCO_2$  and the oxygenation of the ocean interior form changing this.

To test the effect of there being more (or less)  $PO_4^{2-}$  in the ocean in the first place, it is possible to increase the inventory of the ocean as a whole following a re-start, by:

`bg_ocn_dinit_8=1.0E-6`

which will add  $1 \mu mol kg^{-1}$  of  $PO_4^{2-}$  uniformly to the ocean. (A larger/smaller number will obviously increase the glacial nutrient inventory by more/less. A negative number will remove  $PO_4^{2-}$ .)

You might also play with the 'half saturation constant' for nutrient uptake (see Ridgwell *et al.* [2007]):

```
# [PO4] M-M half-sat value (mol kg-1)
bg_par_bio_c0_P04=2.1989611E-07
```

As set, this specifies that at a  $PO_4^{2-}$  concentration of about  $0.2 \mu mol kg^{-1}$ , growth (net export) is half the maximum possible.

You might, for instance, try setting the value to zero, which assumes maximum growth can continue right up nutrient completely running out. This would be similar to assuming all the primary producers in the ocean had extremely small cell sizes (and hence low half saturation values) and were adapted to oligotrophic conditions.

Or, test an ocean dominated by assumed very large cell sizes and phytoplankton that struggle under anything other than fully eutrophic (high nutrient) conditions. e.g. you might try values of  $1.0 \mu mol kg^{-1}$  ( $1.0E-06$ ) or even  $2.0 \mu mol kg^{-1}$  ( $2.0E-06$ ) and see what happens, particularly to the pattern of surface nutrient concentrations and export.

There is one more knob controlling organic matter export and cycling in the ocean that you could tweak and explore the effect of the assumption regarding how much organic matter produced is partitioned into particulate form that sinks, vs. dissolved, controlled by:

```
bg_par_bio_red_DOMfrac=0.66
```

which specifies 66% of organic matter production is diverted into dissolved form. See *Ridgwell and Arndt [2014]*<sup>14</sup>. And/or, you might adjust its mean lifetime in the ocean, which by default is 0.5yr:

```
bg_par_bio_remin_DOMlifetime=0.5
```

### 6.5.3 Other thoughts and suggestions

- If you want to combine forcings, you need to first update the file: `configure_forcings_ocn.dat` – this specifies which ocean flux forcing will be used – simply copy the relevant line from the equivalent file of the forcing to be added. You will also need to copy in the relevant ‘`_sig.dat`’ and ‘`_SUR.dat`’ files. Remember that in the *user-config* file, you will need to set the relevant flux scaling parameter for each different flux in the forcing.

- By default, the  $CO_2$ -climate feedback is ‘on’:

```
# set climate feedback
ea_36=y
```

Should you want to assess the impacts of geoengineering independently of changes in climate – the option is there. (Note that under some of the high end  $CO_2$  emissions scenarios, there may be a degree of collapse of the AMOC that will presumably affect the patterns of ocean acidification and oxygenation etc.)

- If you are having doubts that your experiment is actually ‘doing’ anything (different from the control) – remember to create anomaly maps (plots) to look for specific changes in e.g. saturation state, pH, or the water column inventory of anthropogenic  $CO_2$ . Even before this – plot anomalies of the flux you think you have applied, looking specifically at the region you think you have applied it to. For this, **cookie** saves the 3D distributions of dissolved Fe and  $PO_4$ . See Figures below.
- Always be aware of the caveats regarding this specific model (and models in general) – how much does it different form the ‘real world’ for the modern ocean, particularly in terms of patterns of carbonate saturation? Does it even simulate anthropogenic  $CO_2$  uptake adequately in the first place?

---

<sup>14</sup>Ridgwell, A., and S. Arndt, Why Dissolved Organics Matter: DOC in Ancient Oceans and Past Climate Change, in: Biogeochemistry of Marine Dissolved Organic Matter Eds. Hansell, D. A., and C. A. Carlson, Elsevier (2014).

**6.5.4****Biological pump background information**

A few notes and background information on controls on the parameterized controls on the biological pump in **cookie**.

---

**Remineralization depth**

In the model configuration that you have been using, the degradation of particulate organic matter sinking in the water column proceeds according to a fixed profile of flux with depth (there is no e.g. temperature control on the rate of bacterial degradation of sinking organic matter) with  $CO_2$  and  $PO_4$  released back to the seawater as the particulate flux decreases. The parameter that controls the (*e*-folding) depth scale of particulate organic matter is:

```
bg_par_bio_remin_POC_eL1=589.9451
```

Either edit this value (found under the heading: # -- REMINERALIZATION --) or add a new line at the end of the *user-config* file specifying the value you want. Units are *m*.

Read *Ridgwell et al.* [2007] for additional discussion of this parameter. See Figure 2-4 in *Ridgwell* [2001] ([http://www.seao2.org/pubs/ridgwell\\_thesis.pdf](http://www.seao2.org/pubs/ridgwell_thesis.pdf)) for an illustration of how the flux of particulate organic matter decreases with depth in the ocean, plus references therein.

There is also an associated parameter: `bg_par_bio_remin_POC_frac2`, which sets a fraction of organic matter that is assumed to settle through the water column completely un-altered (currently assigned a value of 0.045 == 4.5%), but this is arguably less useful to change than the remineralization length-scale of the more labile fraction (the other 95.5% of particulate organic carbon exported from the ocean surface).

---

 **$CaCO_3$ :POC rain ratio**

Kicked off by a classic 1994 *Nature* paper by *Archer and Maier-Reimer* (see: *Kohfeld and Ridgwell* [2009]), one potential means of changing atmospheric  $CO_2$  naturally at the last glacial involves changes in the export ratio between  $CaCO_3$  (shells) and  $POC$  (particulate organic matter). Such a change in ratio could come about through a variety of ways (e.g., via the 'silica leakage hypothesis' (see: *Kohfeld and Ridgwell* [2009]) and also through the direct effect of  $Fe$  on diatom physiology (see *Watson et al.* [2000] in *Nature* and also Supplemental Information). There are also ideas about an opposite ocean acidification effect, whereby the less acidic glacial (compared to modern) ocean led to increased calcification and  $CaCO_3$  export. Note that this response (higher saturation == greater rate of calcification) is encoded into your model configuration – see *Ridgwell et al.* [2007b].

In **cookie**, the  $CaCO_3 : POC$  rain ratio is controlled (technically: scaled) by the parameter:

```
bg_par_bio_red_POC_CaCO3=0.0485
```

The pattern of  $CaCO_3 : POC$  rain ratio is not uniform across the ocean (why? (see: *Ridgwell et al.* [2007, 2009])), and its pattern can be viewed in the (2D **BIOGEM**) netCDF variable: `misc_sur_rCaCO3toPOC`.

### 6.5.5 Advanced experiment ideas ...

In today's ocean, iron limitation of biological productivity is only regionally important (compared to e.g. in the first exercise you were changing the dissolved iron supply from dust everywhere). So one might legitimately ask whether specific regions are iron limited, and how important this is (or not) to atmospheric  $pCO_2$ . However, changing the dust field is not trivial, or rather it would be extremely time-consuming to e.g. edit the numbers to double all the dust fluxes in a specific region (e.g. the Southern Ocean).

As an alternative, we could re-configure the model to read in a field specifying the spatial pattern of iron solubility, set the values of iron solubility to a uniform value (0.1%), and then simply edit the values in the file to create whatever pattern of modified iron input you want.

To do this, you need to add the following lines to the \*end\* of the *user-config* file<sup>15</sup>:

```
# Replace internal dust Fe solubility field?
bg_ctrl_force_det_Fe_sol=.true.
# Filename for dust Fe solubility field
bg_par_det_Fe_sol_file='worjh2.det_Fe_sol.MahowaldUNIFORMsol.dat'
```

which tells **cookie** to use a prescribed field of solubility values rather than calculating it, and then directs **cookie** to the file: *worjh2.det\_Fe\_sol.MahowaldUNIFORMsol.dat* containing the spatial pattern of solubility values.

This file can be found in the directory: *genie-biogem/data/input* and can be edited to change (increase or decrease) the solubility of iron in dust in specific regions (hence changing the dissolved iron flux).

It is important to note that the values in this file are % rather than fractional solubility. By default, the file is populated with value 0.1 everywhere (at all ocean grid points) – 0.1%, or a fractional solubility of 0.001 (as per was the value of the parameter *bg\_par\_det\_Fe\_sol* before).

Lastly, we need to modify the iron scavenging rate:

```
# modifier of the scavenging rate of dissolved Fe
bg_par_scav_Fe_sf_POC=0.1
```

to match the change in iron solubility.

A little care will be needed in running experiments. The change in pattern of iron solubility and scavenging rate means that the iron cycle will be slightly different as compared to the *re-start* and when you start running it, the distribution of dissolved iron and hence iron availability to biology will start to change, in turn changing the biological pump and atmospheric  $pCO_2$ . You could either run a control and subtract the drift in whatever parameters you are interested in from the real experiment, or better, spin the adjusted configuration up. For the latter, you'd take the *re-start* as before, and maybe run e.g. 1000 years under the new iron parameter values and see whether the system is re-equilibrating. If this looks 'good', simply use this (e.g. 1000 year) experiment as your new *re-start*.

Increasing or decreasing the iron input to different regions of the ocean surface is simply then a matter of editing the values in *worjh2.det\_Fe\_sol.MahowaldUNIFORMsol.dat* in some pattern (hopefully informed by some hypothesis that you have formulated about iron limitation and the biological pump).

---

<sup>15</sup>Ideally, we would replace or delete a few parameter settings in the *user-config* file in order to avoid confusion or mistakes, but this way is much simpler.



Getting going with ECOGEM  
Ecosystem configuration  
Increasing ecological complexity  
Build it up, tear it down  
Mixotrophy  
Ecology in ... future oceans  
Ecology in ... past oceans  
Ecology in ... fake oceans!  
EcoGENIE 1.1  
Further ideas and investigations

## 7. Marine ecosystems and dynamics

In the chapter we address the role and nature of marine (plankton) ecosystems.<sup>1</sup>

**cookie** includes an explicit ecosystem component, including primary and export production as well as plankton biomass – **ECOGEM**<sup>2</sup> – that is designed as an alternative option to the 'biologically induced export flux' representation of export production (**BIOGEM**). The ecological model takes what is known as a size-structured approach to representing diversity of function in marine ecosystems, and is flexible in being able to be configured to represent any range of size classes of phytoplankton and zooplankton (and/or mixotrophs).

### Stuff to keep in mind...

- We will be working with highly idealised ecosystems in a relatively idealised (modern) ocean.
- The aim is to explore why the model behaves as it does.
- The assumption is that this will give us some insight into why the real world behaves as it does. Perhaps. (It is up to you to question the validity of this assumption.)

<sup>1</sup>Loosely based on original workshop material devised by Ben Ward <b.a.ward@soton.ac.uk>

<sup>2</sup>Ward *et al.* [2017] – Ward, B. A., Wilson, J. D., Death, R. M., Monteiro, F. M., Yool, A., and Ridgwell, A.: EcoGENIE 0.1: Plankton Ecology in the cGENIE Earth system model, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2017-258>, 2017.

**Read.me**

We will use some non ecology-enabled, biogeochemical cycles *restarts* from the previous Chapter (specifically: cookie.CB.p\_worjh2.BASES.caoetal.SPIN and cookie.CB.p\_worjh2.BASESFe.FeMIPSPIN) which comprise a stable ocean circulation and a cycle of nutrients already modified by the biological pump (although we are not going to explore changing the mechanistic control on biological productivity).

However, you will need to download:

- Observational data

```
$ wget --no-check-certificate  
http://www.seao2.info//cgenie/data/GENIE_observations.nc
```

## 7.1 Getting going with ECOGEM

Previously, you were running the standard 'biogeochemical' version of **cookie**<sup>3</sup>. In **BIOGEM**, the biological pump is driven by an implicit (i.e. unresolved) biological community. As in the real ecosystem, the biological uptake of carbon and nutrients (such as phosphorus and iron) is limited by light, temperature and nutrient availability. However, unlike the real ecosystem, any uptake is *directly* and *instantly* converted to particulate and dissolved organic matter (POM, DOM), of which the POM is exported to the ocean interior via (gravitational) settling, i.e.:

- surface inorganic nutrients  $\xrightarrow[\text{and export}]{\text{production}}$  POM and DOM

In contrast, in this chapter you we are going get started with the '**ECOGEM**' ecological modeling package<sup>4</sup>. This will allow us to extend the capabilities of **cookie** to examine a range of questions relating to the role of physiology and community structure in regulating the biological pump and hence atmospheric  $CO_2$  etc. In **ECOGEM**, biological uptake is again limited by light, temperature and nutrient availability ... but now it must pass through an explicit and dynamic intermediary plankton biomass pool before the net products of biological production can expressed as the production of POM and DOM:

- surface inorganic nutrients  $\xrightarrow{\text{production}}$  plankton biomass  $\xrightarrow{\text{export}}$  POM and DOM

The existence of a plankton biomass reservoir creates a delay term in the system such that the occurrence of warm temperatures in a sunlit surface with abundant nutrients, does not in itself guarantee immediate and massive carbon export. This is because the ecosystem biomass must build up first. This could be important for e.g. the timing of spring blooms.

Note that while the experimental configurations are based on those of *Ward et al.* [2018], here we use a slightly different modern continental configuration and physics tuning (and hence ocean circulation state) and a slightly different iron cycle tuning. **ECOGEM** itself has also been adjusted to reduce the carbon export relative to phosphorous and has photosynthesis suppressed under sea-ice.

### Running the model

We will start with the simplest possible configuration of **ECOGEM**, with just a single (small) phytoplankton class<sup>5</sup>. You can run an experiment (e.g. for 10 years) based on this configuration:

```
$ ./runcookie.sh cookie.CBE.p_worjh2.BASES LABS LAB.7.1.EXAMPLE 10
    cookie.CB.p_worjh2.BASES.caoetal.SPIN
```

(Here, although you are using a new *base-config* – *cookie.CBE.p\_worjh2.BASES* – it is identical to the 16-level phosphorous-only biogeochemical configuration used in the previous Chapter except now with the ecosystem model enabled (the 'E' in '*cookie.CBE.p\_worjh2.BASES*').)

The model will run as before ... except much slower ... :( ... as **cookie** has to calculate plankton growth and ecological interactions in addition to everything else it was doing previously.

<sup>3</sup>e.g. see *Ridgwell et al.* [2007]

<sup>4</sup>see: *Ward et al.* [2017] – Ward, B. A., Wilson, J. D., Death, R. M., Monteiro, F. M., Yool, A., and Ridgwell, A.: EcoGENIE 0.1: Plankton Ecology in the cGENIE Earth system model, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2017-258>, 2017.

<sup>5</sup>i.e. as if the ocean was populated with a small species of phytoplankton (photo-autotroph) and nothing else.

## Viewing 2D time-slice output

What to look for?

**ECOGEM** saves its own 2D (but not 3D) netCDF *time-slice* output (also in the results sub-directory of your experiment output folder).

variable	variable (long name)	Description	Application
eco2D_Plankton_C_x	C Biomass - Popn. #x	Biomass ( $mmolCm^{-3}$ ) of phytoplankton or zooplankton [OF GIVEN SIZE CLASS].	Relative dominance of different size classes. <sup>6</sup> Comparison to observations.
eco2D_Plankton_P_x	P Biomass - Popn. #x	Biomass ( $mmolPm^{-3}$ ) of phytoplankton or zooplankton [OF GIVEN SIZE CLASS].	(as above)
eco2D_Plankton_ChI_x	Chl Biomass - Popn. #x	Biomass ( $mgChlm^{-3}$ ) of phytoplankton [OF GIVEN SIZE CLASS].	Comparison to observations.
eco2D_Plankton_C_x	C Biomass - Total	Total ecosystem biomass ( $mmolCm^{-3}$ )	–
eco2D_Plankton_P_x	P Biomass - Total	Total ecosystem biomass ( $mmolPm^{-3}$ )	–
eco2D_Plankton_ChI_x	Chl Biomass - Total	Total ecosystem biomass ( $mgChlm^{-3}$ )	Comparison to observations, e.g., satellite retrieval.
eco2D_Size_*	–	Various size-related metric.	Analysis of ecosystem changes.
eco2D_Uptake_Fluxes_x	Uptake Fluxes x	Rate of nutrient (can carbon) uptake.	Primary productivity.
eco2D_xGamma_*	Limitation (BY NUTRIENT) (OF GIVEN SIZE CLASS)	Various size-related metric.	Analysis of ecosystem changes.
eco2D_Diversity_*	–	Nutrient limitation of growth.	Analysis of ecosystem controls.
eco2D_Nutrients_*	Nutrients *	Ocean surface nutrient concentrations including that of DIC, in units of $mmolm^{-3}$ .	(Duplicate of <b>BIO-GEM</b> fields ... but different units.)

Table 7.1: Summary of the main 2D time-slice output for ecology.

### Comparing to observations

Models are intended as an (as close as possible) approximation of the real world (whatever that is). It might, therefore, be useful to check if our approximation is in anyway realistic. We can do this by comparing the model output<sup>7</sup> to observations.

1. Download a compilation of key biogeochemical variables –  
[http://www.seao2.info/cgenie/data/GEnIE\\_observations.nc](http://www.seao2.info/cgenie/data/GEnIE_observations.nc) .
  2. Compare model output to key biogeochemical variables derived from ocean measurements.<sup>8</sup>  
The variables in the GEnIE\_observations.nc **netCDF** file include:
    - Phosphate – how well does **cookie** simulate surface nutrient availability, and to what degree does ecosystem complexity help explain observations?
    - Observed Chlorophyll is remote-sensed, and can be contrasted with the chlorophyll biomass simulated by **ECOGEM**.
  3. Does the model perform well with respect to reproducing these variables? If not, why not?
- 

<sup>7</sup>NOTE: **ECOGEM** only saves a limited number of surface (2D) data arrays. You can look at other variables (in 2D and 3D) by opening the corresponding **BIOGEM** netCDF files.

<sup>8</sup>Remember that you can create difference maps.

**7.2****Ecosystem configuration**

In the last section you ran a very simple configuration of the **ECOGEM** ecosystem model, and compared it to observations. In this section we are going to add a bit more ecological realism, with the aim of improving model performance (i.e. as contrasted against observations). We will start by adding a zooplankton population that should bring a degree of ‘top-down’ control to the phytoplankton population.

First, in the *user-config* file (LAB.7.1.EXAMPLE), note:

1. `bg_par_bio_prodopt="NONE"` – which effectively disables the simple biological export scheme in **BIOGEM**, replacing it with the explicit biology of **ECOGEM**. This is a necessary step whenever running **ECOGEM**, because we do not want the implicit and explicit biological schemes to be implemented simultaneously ... then:
2. One of the most important parameters specifies the *ecosystem configuration* file:

```
eg_par_ecogem_plankton_file = 'NPD.eco'
```

This points to a file (located in `~/cgenie.cookie/genie-ecogem/data/input/`) that specifies every plankton population (‘species’, if you like) that is accounted for in the model experiment. If you open that file in a text editor, you will see something akin to the following:

```
01          02      03
\|          \|      \|
-START-OF-DATA-
Phytoplankton    10.00   1
-END-OF-DATA-

\|          \|      \|
01          02      03

DATA FORMAT AND ORDER
-----
COLUMN #01: plankton functional type name
COLUMN #02: plankton diameter (micrometers)
COLUMN #03: number of randomised replicates

INFO: TRACER ASSIGNMENT RULES
-----
Plankton functional type one of: Prochlorococcus
                                Synechococcus
                                Picoeukaryote
                                Diatom
                                Coccolithophore
                                Diazotroph
                                Phytoplankton
                                Zooplankton
                                Mixotroph
```

NPD.eco note that only the lines in between the -START-OF-DATA- and -END-OF-DATA- tags are read by the computer. The rest is there solely for your guidance.

Each line that is entered in the computer-readable area tells the model to create a distinct plankton population in the model. The 'plankton functional type' of this population is specified in the first column, while the plankton diameter is specified in the second column. A '1' must always be placed in the third column<sup>9</sup>.

In this 'NPD' (single nutrient-plankton-detritus) configuration, we only have a 10 micron generic phytoplankton. The ecological and physiological traits of this population are assigned automatically according to the size and the functional type (here: photo-autotroph).

**NOTE:** The only PFTs available at the moment are Phytoplankton, Zooplankton and Mixotroph. The other groups currently have no real functionality associated with them.

**NOTE:** This file format is fussy, and you cannot have any empty lines in between the -START-OF-DATA- and -END-OF-DATA- tags – every line between these tags must have data (3 parameter values).

3. You can increase the ecological complexity of the model by adding another plankton population. Save the *ecosystem configuration* file under a new and highly intuitive name (such as NPZD.eco), and add another line specifying a 100 micron zooplankton. e.g.:

```
Zooplankton      100.00      1
```

It is important that the zooplankton is 10 times larger than the phytoplankton in terms of diameter. This is the optimal predator-prey length ratio in the default configuration. (You could maybe think about changing this value later on.)

4. To run the model with this new configuration, change the name of the *ecosystem configuration file* in the *user-config* file...

```
eg_par_ecogem_plankton_file ='NPZD.eco'
```

5. Save the new *user-config* file under a different name (e.g. LAB.7.2.NPZD). You can now execute the model at the command line<sup>10</sup> ...

```
$ ./runcookie.sh cookie.CBE.p_worjh2.BASES LABS LAB.7.2.NPDZ 10  
cookie.CB.p_worjh2.BASES.caoetal.SPIN
```

... or submit it to the cluster queue ...

6. Once you have completed the new simulation, compare the new results to the old simulation, and also in terms of its ability to reproduce observations. Has the addition of zooplankton to the model improved its behaviour or not? Look also at the global distributions of carbon biomass in the phytoplankton and zooplankton populations (again, a log scale might help).

*How have the zooplankton interacted with the phytoplankton to change the ecological dynamics in the model? Has global productivity (and export production) increased or decreased??*

10 years is clearly far too short a time to accommodate any change in the global cycle of nutrients (and carbon and oxygen), but may be sufficient to see much of the impact of changing the assumed ecosystem structure. (You can always run for longer to judge for yourself on what time-scales what components of the Earth system adjust and hence what the 'ideal' (practical) run-time for changing the ecosystem structure might be.)

<sup>9</sup>It doesn't 'do' anything, but the model still needs it ... why ... WTF?

<sup>10</sup>Don't forget to change the name of the *user-config* file here as well ...

### Visualising composite data

We can perhaps get a better handle on this question by looking at the ratio of phytoplankton-to-zooplankton biomass. Use **Panoply** to combine data arrays.

1. Open C Biomass - Popn. 001 (10.00 micron phytoplankton). Next, select C Biomass - Popn. 002 (100.00 micron zooplankton), and click the Combine Plot icon at the top of the **Panoply** window.
2. A box will open up asking you In which existing plot should I combine the variable. Click Combine.
3. The new map shows total zooplankton carbon biomass minus the total phytoplankton carbon biomass (see the label on the color scale). We want to look at the Z:P biomass ratio, so from Array(s), select Array 2 / Array 1.<sup>11</sup> You may find it helpful to look at the data on a log scale, with a scale range of 0.1 to 10. You might also like to change the Color Table: option to GMT\_polar.cpt.

### Questions:

- What does this plot say about the relationship of zooplankton and phytoplankton in different regions of the ocean?
- In what regions do zooplankton or phytoplankton dominate?
- What affect does a high Z:P ratio have on the biomass of the phytoplankton population?<sup>12</sup>

---

<sup>11</sup>Make sure that you are looking at the right year for both.

<sup>12</sup>For example, in terms of the chlorophyll concentration.

### 7.3 Increasing ecological complexity

In the last section, we looked at the results of some simulations based on ‘NPD’ (one nutrient-(phyto)plankton-detritus) and ‘NPZD’ (one nutrient-(phyto)plankton-zooplankton-detritus) type ecosystem models. Here we will begin to incorporate a bit more ecological complexity.

---

#### Plankton size classes

We are going to add a few more plankton size classes, so that we end up with small, medium and large phytoplankton and (small, medium and large) zooplankton.

1. Save the *ecosystem configuration* file under a new name (e.g. 3P3Z.eco), replacing the existing plankton populations with the ones described in Table 7.2.

2. To run the model with this new configuration, change the name of the *ecosystem configuration file* in the *user-config* file:

```
eg_par_ecogem_plankton_file='3P3Z.eco'
```

3. Save the new *user-config* file under a different name (e.g., LAB.7.2.3P3Z) and then run the new model at the command line (e.g. for 10 years).

Table 7.2: Plankton functional groups and sizes.

<i>j</i>	PFT	Diameter ( $\mu\text{m}$ )	<i>j</i>	Functional Type	Diameter ( $\mu\text{m}$ )
1	Phytoplankton	0.6	4	Zooplankton	6.0
2	Phytoplankton	6.0	5	Zooplankton	60.0
3	Phytoplankton	60.0	6	Zooplankton	600.0

---

#### Viewing 2D time-slice output

In the 2D *time-slice* netCDF data, there are now a lot more variables. There are all the same diagnostics as before, plus some new ones relating to the new plankton populations you have just added. There are also variables describing the size distribution and diversity of the photosynthetic community (non-phototrophic populations are ignored in these metrics). These were not included before, because there was only one phytoplankton population.

#### Size fractions

Variables “eco2D\_Size\_Frac\_...” give the chlorophyll biomass in the three size fractions:

1. picophytoplankton (diameter  $\leq 2 \mu\text{m}$ )
2. nanophytoplankton ( $2 < \text{diameter} \leq 20 \mu\text{m}$ )
3. microphytoplankton (diameter  $> 20 \mu\text{m}$ )

### Size metrics

Variables 'eco2D\_Size\_...' give metrics describing the phytoplankton size distribution.

- eco2D\_Size\_Mean: Geometric mean<sup>13</sup> phytoplankton diameter, weighted by carbon biomass
- eco2D\_Size\_Stdev: Geometric standard deviation<sup>14</sup> of phytoplankton diameter, weighted by carbon biomass.
- eco2D\_Size\_Minimum: Diameter of smallest phytoplankton contributing >0.1% of the total phytoplankton carbon biomass.
- eco2D\_Size\_Maximum: Diameter of largest phytoplankton contributing >0.1% of the total phytoplankton carbon biomass.

### Diversity metrics

Variables 'eco2D\_Diversity\_...' give metrics describing the phytoplankton diversity.<sup>15</sup>

- eco2D\_Diversity\_Threshold: the threshold diversity index. The number of species contributing >0.1% of the total phytoplankton carbon biomass [Barton et al., 2010]<sup>16</sup>.
- eco2D\_Diversity\_Berger: the inverse Berger(-Parker) index [Berger and Parker, 1970]<sup>17</sup>. The proportion of carbon biomass made up by all but the single most dominant population. For example, if the dominant population accounts for 40% of the total carbon biomass, inverse Berger (-Parker) index is 0.6.
- eco2D\_Diversity\_Simpson: the inverse (Gini-)Simpson index [Simpson, 1949]<sup>18</sup>. This is effectively the probability that two samples taken at random from the community will be from a different species (note that the probability of selecting a population is dependent on carbon biomass, not cell abundance). If we define the proportional biomass of each species as its relative contribution to the total carbon biomass in the community, the inverse Gini(-Simpson) index is calculated as one minus the sum of the squares of the proportional biomasses of each species.
- eco2D\_Diversity\_Shannon: the Shannon(-Wiener or -Weaver) index [Shannon, 1948]<sup>19</sup>. With the proportional biomass defined as above, the Shannon index is defined as the sum of [the proportional biomass multiplied by the logarithm of the proportional biomass] for each species.

---

<sup>13</sup>We use the geometric mean and standard deviation, because phytoplankton biomass is approximately log-normally distributed across the phytoplankton size range.

<sup>14</sup>The geometric standard deviation describes the *relative* size range of the phytoplankton. For a geometric standard deviation of  $\sigma$ , ~68.2% of the phytoplankton carbon biomass will be in cells no more than  $\sigma$  orders of magnitude smaller or larger than the geometric mean size.

<sup>15</sup>NOTE: The threshold index is a fairly crude measure of the total number of species in the community, relative to a small and arbitrary threshold of relative biomass. This index is not very sensitive to the relative biomass of individual species (although one very successful species can raise the absolute value of the threshold, thus lowering the diversity). The other three indices do more to quantify the evenness of the community. The more unequal the proportional abundances, the smaller the value of the index. If almost all the abundance is concentrated into one type and all the other types are very rare, the latter three indices can become very small. A community with fewer species, but with more evenly distributed biomass, may well have higher values for these three diversity indices.

<sup>16</sup>A D Barton, S Dutkiewicz, G Flierl, J Bragg, and M Follows. Patterns of diversity in marine phytoplankton. *Science*, 327(5972):1509–1511, 2010.

<sup>17</sup>W H Berger and F L Parker. Diversity of planktonic foraminifera in deep-sea sediments. *Science*, 168 (3937):1345–1347, 1970.

<sup>18</sup>E H Simpson. Measurement of diversity. *Nature*, 163:688, 1949.

<sup>19</sup>C E Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(379-423 and 623-656), 1948.

View the spatial distributions of some of these metrics, but bear in mind that they summarize the diversity of a phytoplankton community that includes just three species. They are probably not that revealing, so we will come back to them later. Instead, have a look some of the other metrics describing the model ecosystem.

**Questions:**

- What are the global distributions of the different size classes?
- How do the global biomass distributions compare to variables such as temperature<sup>20</sup>, or primary production (Uptake Fluxes C)?
- How does nutrient, light and temperature limitation vary between the size classes?
- Can you account for the distribution biomass between the size classes according to the different limiting factors?

---

**Create your own ecosystem**

Save the *ecosystem configuration* file under a new name and add some more plankton populations (whatever and as many as you like<sup>21</sup>). Update the **cookie** parameter eg\_par\_ecogem\_plankton\_file in the *user-config* and save this file under a new name. Run the new model.

**Questions:**

- How many populations can you get to coexist (i.e. each having a non trivial (>0.1%) biomass)?
- What effect do the new populations have on the community as a whole?
- What effect, if any, do your additions have on the strength of biological export?  
(e.g. look at bio\_fpart\_POC in fields\_biogem\_3D.nc.)

---

<sup>20</sup>Ocean temperature is saved in fields\_biogem\_3D.nc

<sup>21</sup>Just bear in mind that the more populations you put in, the slower the model will run!

**7.4****Build it up, tear it down****A fully size-structured ecosystem**

We are now going to switch to a more diverse version of the size-structured ecosystem model. This configuration has 8 size classes of phytoplankton, and 8 size classes of zooplankton, as shown in Table 7.3.

1. Save the *ecosystem configuration* file under a new name, replacing the existing plankton populations with the ones described in Table 7.3.
2. Create a new *user-config* (e.g., LAB.7.4.8P8Z) and point to the new *ecosystem configuration* file, e.g., (e.g. 8P8Z.eco).
3. Run the new model for at least 20 years (this will probably take about 10 minutes).

Table 7.3: Plankton functional groups and sizes.

<i>j</i>	PFT	Diameter ( $\mu\text{m}$ )	<i>j</i>	Functional Type	Diameter ( $\mu\text{m}$ )
1	Phytoplankton	0.2	9	Zooplankton	0.6
2	Phytoplankton	0.6	10	Zooplankton	1.9
3	Phytoplankton	1.9	11	Zooplankton	6.0
4	Phytoplankton	6.0	12	Zooplankton	19.0
5	Phytoplankton	19.0	13	Zooplankton	60.0
6	Phytoplankton	60.0	14	Zooplankton	190.0
7	Phytoplankton	190.0	15	Zooplankton	600.0
8	Phytoplankton	600.0	16	Zooplankton	1900.0

**Ecosystem characteristics**

We can now begin to look at the size and diversity metrics in a more meaningful way.

1. View:

- (a) the total carbon biomass
- (b) the carbon uptake flux (i.e. primary production)
- (c) the geometric mean size

Make sure that in each case that you are looking at the last year of model output. You may also find it useful to adjust the color scale, or to change to a logarithmic colour scale.

Looking at the maps, we can perhaps pick out three different “biomes” in terms of their community properties:

- (a) The low-latitude oligotrophic gyres are relatively unproductive, and support some of the lowest annual mean biomass in the surface ocean. In these regions the mean phytoplankton size is very small.
- (b) Sub-polar latitudes between  $40^\circ$  and  $50^\circ$  (either N or S) are much more productive, and support very high annual mean biomass. These communities also have the highest mean sizes of any region.
- (c) The polar oceans are also highly productive (except perhaps the high Arctic), and support relatively high annual mean biomass. These communities are made up (in the model, at least) of slightly smaller phytoplankton than we see in the subpolar regions.

2. What can we find out about the community structure in these regions? Open up some of the other metrics describing the community (standard deviation of size distribution, size fractionation, diversity, limiting factors). What can you find out about the community structure within each region, in terms of coexistence and exclusion?
  - Does the community span a broad or narrow size range?
  - How many size classes are coexisting in each biome?
  - What is the smallest and largest size class in each biome?
  - How much biomass is concentrated in each size fraction (picoplankton, nanoplankton and microplankton)?

Overall – what factors do you think are most important in terms of dictating the global distribution of each size class?

To find out the answers to these questions, you are going to pull the model apart, and then put it back together. At each stage the aim is to bring in a different limiting factor, so that you can see its effect on the model behaviour.

---

### The fundamental niche

The first step is to find out the impact of abiotic factors on the distribution of different phytoplankton sizes. In other words, we need to find out what the distribution of the phytoplankton would be in the absence of any ecological interactions, such as resource competition and predation. This is effectively their ‘fundamental niche’.

The fundamental niche is a fairly abstract concept, and not something that can be measured in the real world. In model world, however, we can get a useful estimate of the fundamental niche by making a few simple changes to the model.

1. First of all, you can remove all predation, simply by removing the zooplankton from the *ecosystem configuration* file.
2. Next, you also need to remove all competition for nutrients and light. This involves tweaking the model equations so that the phytoplankton are not nutrient limited, and do not attenuate light. To do this, all you need is to add the following line to the *user-config* file:  
`eg_fundamental = .TRUE.`
3. If you now run the model you should have a community of eight phytoplankton size classes that are growing solely as a function of the incoming light and the temperature. This growth will be balanced balanced the basal (i.e. non-grazing) mortality. As there is no feedback between the ecosystem and the environment, populations that can survive will grow exponentially and without limit, potentially reaching astronomical abundance in very little time. Populations that cannot survive will rapidly decline to almost nothing.

The regions in which each plankton shows positive growth defines its fundamental niche. This is a function of abiotic conditions only, and is the absolute limit of its geographical range. Look at the carbon biomass distribution in each size class (set the data range in each case from 0 to  $1\text{mmolCm}^{-3}$ ).

- How and why does the fundamental niche vary with size?
- Could the limits of the fundamental niche explain some of the patterns seen in the full model?

### Resource competition.

The next step is asses the impact of resource competition. We are first going to do this in the absence of any zooplankton grazing.

1. All you need to do at this stage is to re-enable nutrient and light competition. To do this, simply delete ‘eg\_fundamental = .TRUE.’ from the *user-config* file (or comment out the line to disable it), and save under a new name. Leave the *ecosystem configuration* file as it is.
2. You should have a community of eight phytoplankton size classes that are competing for nutrients and light, again as a function of temperature. This is a much more realistic simulation, as feedbacks between the ecosystem and the environment serve to limit the size of the phytoplankton populations. Examine the model to find out:
  - What size classes are able to persist when resource competition is enabled?
  - Why are different size classes more or less abundant in different areas?
  - How does the distribution of each size class compare to the fundamental niche?
  - What are the reasons for any differences?

Phytoplankton biogeography at this stage begins to approximate the realised niche, which defines the range of conditions that support a population in the presence of ecological interactions. Note that at this stage, however, we have ignored the effects of any predator-prey interactions, as the zooplankton grazers are still missing.

### Resource competition + one generalist zooplankton

The previous simulation is clearly unrealistic (although, hopefully, informative). You are now going to add back in just a single zooplankton class, that grazes equally on all plankton (including itself).

1. Add a 100 micron zooplankton into the *ecosystem configuration* file, and save under a new name. Also update the *user-config* file to reflect the change, and save under a similar name.
2. You need to modify the model so that the zooplankton eats all prey with equal preference. This can be done by adding the following lines to the *ecosystem configuration* file.

```
eg_ns=1
eg_pp_sig_a=1.0e99
```

NOTE: For aficionados, the first parameter disables prey-switching (i.e. predators no longer preferentially attack the most abundant prey). The second parameter increases the width of the grazing kernel (i.e. predators can attack a range of prey across a huge size range with equal preference).

3. The addition of zooplankton to the model community should give a more accurate approximation of the realised niche.
  - Does the addition of a single zooplankton grazer enable more or less coexistence?
  - What factors might be responsible for any shifts in biogeography?

**Resource competition + one “switching” zooplankton**

You began with a full food-web containing 8 phytoplankton and 8 zooplankton size classes. The diversity of zooplankton clearly has an effect on the phytoplankton community that is not seen in the previous experiment. This effect can be imitated with just one generalist zooplankton if we instruct it to graze preferentially on the most successful prey.

Re-enable this ‘prey switching’ effect by changing the following control parameter to a 2:

```
eg_ns=2  
eg_pp_sig_a=1.0e99
```

Compare this simulation to the first experiment (8 phytoplankton and 8 zooplankton) to see how the inclusion of prey switching increases coexistence through the ‘kill-the-winner’ mechanism.

- How does nutrient limitation change with phytoplankton size, and how might zooplankton be affecting this?
- Look at the C:P biomass ratio in the community as a whole, and compare to your estimates from the NPZD model (Lesson 1).
- How does the C:P ratio vary with size? How does having a diverse community affect the coupling of carbon and limiting nutrients?

---

**Further questions to answer**

- What sets the fundamental niche, and how does it change with size?
- How is the fundamental niche modified by resource competition?
- What species are favoured in terms of nutrient competition?
- How is the outcome of competition affected by...
  - Abiotic conditions?
  - Increased mortality (through generalist grazing)?
  - Density-dependent mortality (through specialist grazing)?
- Do these experiments tell you all you need to know?  
What other modifications can you think of making?

**7.5****Mixotrophy**

Try adding some mixotrophs to the phytoplankton and zooplankton already present in the community. These will have exactly half the nutrient uptake traits of phytoplankton of a similar size, and half the prey capture traits of zooplankton if a similar size. To do this:

1. Save the previous *ecosystem configuration* file under a new name.
2. Edit the new *ecosystem configuration* file and add an additional line to add a mixotroph:

Mixotroph      xxx      1

where xxx is the class size of the mixotroph. (And remember to save it.)

3. Update the *user-config* to point to the new *ecosystem configuration* file, and save again under a new name. (It is generally a good idea to make a note of the name and goal of each experiment as you set it up.)
4. Run the new model for at least 10-20 years.

Further questions to explore/answer:

- How does this effect the mean and standard deviation of cell size?  
(Size and diversity metrics will be calculated for phytoplankton and mixotrophs together)
- How does mixotrophy effect the C:P ratio of organic matter?
- How does the realised niches of mixotrophs compare to the fundamental niches of phytoplankton?

Also: try replacing all the phytoplankton and zooplankton with a range of sizes of mixotrophs. How does the simulation differ from one with the same size range of separate phytoplankton and zooplankton classes?

For instance – you might also try having an ecosystem comprising just a single small (e.g.  $10\mu m$ ) mixotroph (no phytoplankton and no zooplankton), and perhaps an ecosystem with one small mixotroph  $10\mu m$  and one additional mixotroph, 10 times larger  $100\mu m$ . Compare the productivity of such an ocean compared to some of your previous simplified ecosystem configurations (such as a single small  $10\mu m$  phytoplankton, and a configuration with a single small phytoplankton and a single larger  $100\mu m$  grazer).

## 7.6 Ecology in ... future oceans

What might happen to primary production and the biological pump as (future) climate continues to change? What might happen to the distribution of species (here: size classes) in the ocean? Does it 'matter' e.g. for marine bio geochemical cycling and feedback on atmospheric  $pCO_2$ ? (And how different is the Earth system response when using a complex ecosystem model rather than the highly simplified representation of biological export you used previously?)

There are all good questions for a model!

The best place to start is with the full complexity ecosystem structure with 8 size classes for both phytoplankton and zooplankton (as described in *Ward et al. [2018]*). However, you do not have a *restart* for this – so far, you were adding an ecosystem to a distribution of nutrients and carbon cycling derived from a much simpler description of biological export production.

1. First, create a *restart* that includes a complete ecosystem. You could run this for 10,000 years (it would take 3 days!). But you might get away with e.g., 1,000 years following on from the non-ecosystem *restart*. Copy your previous full-complexity ecosystem *user-config* (LAB.7.4.8P8Z) to e.g., LAB.7.6.8P8Z and run it:

```
$ ./runcookie.sh cookie.CBE.p_worjh2.BASES LABS LAB.7.6.8P8Z 1000
cookie.CB.p_worjh2.BASES.caoetal.SPIN
```

Even 1,000 years might not be long enough ... check the *time-series* of POC export and surface nutrient concentrations and judge whether steady-state has been reached or not.

2. You then need to run a historical transient experiment (e.g., as per in Chapter 5). To adapt your *spin-up* experiment as a historical transient, add the following lines to the *user-config*.

- Under # \*\*\* DATA SAVING \*\*\*, make sure you change the existing line to the following, or add these at the end of the parameter block:

```
bg_par_data_save_sig_timeinterval=0.0
bg_par_infile_slice_name='save_timeslice_historicalfuture.dat'
bg_par_data_save_slice_timeinterval=0.0
bg_par_infile_sig_name='save_timeseries_historicalfuture.dat'
```

These simply ensure you are saving output at appropriate years/intervals.

- Under # \*\*\* FORCINGS \*\*\* and in place of any existing forcing lines:

```
# historical forcing of pCO2
bg_par_forcing_name='pyyyzz.RpC02_Rp13C02.historical2010'
```

This prescribes increasing atmospheric CO2 following historical observations.

- And then under # \*\*\* MISC \*\*\*:

```
# change start year
bg_par_misc_t_start=1765.0
```

Starts the experiment at year 1765.

Run this experiments, starting form the end of your *spin-up*, for 245 years, which will take it up until year 2010.

3. To further adapt your *user-config* to run future CO2 scenarios – see Section 5.2.3.

What to look for? The questions (at the start) should help guide you in this, and might include:

- Patterns of primary productivity and export (where do they increase, decrease), plus time-series of global totals (what is the global impact of warming?).
- Patterns of ocean oxygenation (and patterns of dysoxia and anoxia).
- Surface ocean nutrient distributions, and nutrient limitation.
- Spatial patterns of plankton size – either simply for mean size, or look at the patterns of biomass of specific size classes. (You might follow a the preferred habitat/location of a particular size class with latitude to see whether it experiences a 'range shift'.)
- Diversity etc.

You might also contrast with emissions scenarios that you might have run previously employing the simplified biological export scheme in **BIOGEM**, rather than the complex ecosystem model of **ECOGEM**. (Make sure you are comparing between the same emissions scenarios). The question to be answered would be: does including a complex representation of marine ecology 'matter'?

## 7.7 Ecology in ... past oceans

Here we consider a series of examples<sup>22</sup> of the **ECOGEM** model of marine ecology, applied to published paleo configuration of **cookie** and used to ask questions of how different could marine carbon cycling (and atmospheric  $pCO_2$ ) and oxygenation have been in the (relatively recent) past and how do model projections compare with proxy observations.

Three examples come from the Last Glacial Maximum, with direct comparison being made to post glacial time (here, the late Holocene), with the intention to explore the role of climate cooling, altered ocean circulation, and increased iron supply to the ocean in modifying plankton distributions and size structures. Secondly, from around the time of the Paleocene-Eocene Thermal Maximum (PETM) some 55 Myr ago, now both considering a different paleogeography, a warmer ocean, and then transient warming on top of that. And thirdly, marine ecology and carbon cycling in the aftermath of the impact and extinction event at the end of the Cretaceous, some 66 Myr ago, and return to our earlier exploration of what happens if you lose the larger plankton in the ocean.

In general, any of the different *user-configs* you have created and experimented with, you be applied to a paleo *base-config*. Primarily, only the atmospheric the atmospheric CO<sub>2</sub> forcing values needs changing in order to simulate e.g., a warmed past climate state.

---

<sup>22</sup>Either published or in the works or existing in some publication fantasy ...

### The Last Glacial

The Last Glacial Maximum (LGM) (ca. 19 to 23 ka) was characterized by lower sealevel and higher ocean salinity, colder ocean temperatures, and a reorganized meridional overturning circulation in the Atlantic. The latter 2 changes in particular should have impacted marine ecosystems and indeed, observations suggest range migration (temperature-tracking) and shifts in the zone of highest productivity in the Southern Ocean, amongst other impacts. The aim of this particular investigation, is to assess what these ecological changes are (at least in model world).

Provided is a configuration of LGM ocean circulation that has been tuned to fit observations of benthic carbon isotopes, which provide an observational constrain on large-scale ocean circulation. To this, you'll an ecosystem (in place of the default **BIOGEM** 'induced export' scheme). Also provided is a **cookie** configuration for late Holocene ('HOL') (0-6 ka), created in exactly the same way and also tuned to respective (0-6 ka) benthic carbon isotope data. Configuration HOL provides a point of comparison (or control) for you to compare the ecology in a colder, LGM ocean against.<sup>23</sup>

The *user-configs* you need to use, or copy-rename, can be found in the directory:

genie-userconfigs/MS/odalenetal.CP2019

(and you can run your experiments from here<sup>24</sup>, or better, copy-edit-rename the *user-configs* you run experiments with, from the LABS sub-directory).

Read the *readme.txt* file for instructions for the basic set of command line parameters needed, but remember that you may be running your *user-config* with a different name and likely from a different sub-directory. Spin up the following experiments (and then experiment with them later)<sup>25</sup>:

- (1) cookie.CB.Glteiiaa.BASESFeTDTL\_rb *base-config*  
+ cookie.CB.Glteiiaa.BASESFeTDTL\_rb.SPIN *user-config*
- (2) cookie.CB.Glteiiva.BASESFeTDTL\_rb *base-config*  
+ cookie.CB.Glteiiva.BASESFeTDTL\_rb.SPIN *user-config*

Submit these to the cluster ... remembering that you will need to recompile **cookie** between (1) and (2) (and potentially also re-compile before running (1) if you have not used that particular *base-config* immediately prior). Run the spin-ups for 10,000 years.

Neither of these configurations currently have an ecosystem enabled, but they will provide a baseline against which you can contrast a pair of experiment that include explicit ecology. What you need to do know is to add in/enable **ECOGEM**. To do this, you need to modify both *base-* and *user-config* files of both the HOL and LGM model configurations.<sup>26</sup>

#### 1. *base-config*

First, you need to enable the **ECOGEM** module.

At the top of the HOL *base-config* file cookie.CB.Glteiiaa.BASESFeTDTL\_Crb<sup>27</sup> you will see the line:

ma\_flag\_ecogem=.FALSE.

Simply change this to .TRUE.

---

<sup>23</sup>i.e. you run pairs of HOL and LGM experiments and contrast between them, rather than necessarily comparing to previous modern configurations.

<sup>24</sup>If you use them in this directory, make sure at the command line, you replace LABS with MS/odalenetal.CP.2019.

<sup>25</sup>Noting that the long file-names differ only in a single 'V' vs. an 'G' ...

<sup>26</sup>Strongly recommended that that you copy-rename both sets of files and edit the copies.

<sup>27</sup>And similarly for the LGM one.

## 2. *user-config*

Next, some deletions and additions are needed in the *user-config* provided (which only has implicit biological export enabled).

In the section of the file under the heading:

```
#  
# --- BIOLOGICAL NEW PRODUCTION -----  
#
```

you are going to delete everything there (in that section) and replace it with:

```
# biological scheme ID string  
bg_par_bio_prodopt="NONE"
```

The only other thing you need then is a section of code that defines all the **ECOGEM** ecological parameters.

Go open file `cookie.CBE.p_worjh2.BASESFeTDTL.FeMIRSPIN`, which you can find in the *user-config* sub-directory MS/wardetal.2018, and you'll see under the heading:

```
#  
# --- ECOGEM -----  
#
```

a long list of parameter settings (down to the next section headed DATA SAVING). Simply copy-paste this entire section (including ECOGEM header lines if you like), anywhere in your *user-config* file. At the very end of the file would do just fine.<sup>28</sup>

And lastly ... under:

```
#  
# --- MISC -----  
#
```

(and the end of the section) add the following lines

```
# kraus-turner mixed layer scheme on (1) or off (0)  
go_imld = 1  
# set mixed layer to be only diagnosed (for ECOGEM)  
go_ctrl_diagmld=.true.  
# add seaice attenuation of PAR  
eg_ctrl_PARseaicelimit=.true.  
# relative partitioning of C into DOM  
eg_par_beta_POCtoDOC=0.75
```

This enables a 'mixed layer depth' scheme in the ocean circulation model that **ECOGEM** needs to calculate light limitation during esp. intervals of high latitude / wintertime deep mixing, but then only allows **ECOGEM** to 'see' the depth and does not allow physical mixing in the ocean. Then limitation of photosynthesis added to sea-ice covered areas. Finally, an adjustment is made of carbon vs. phosphorous in exported *POM*.

This ... should do-it, i.e. you have added the same tuned ecosystem model as described in *Ward et al. [2018]* to your LGM / HOL *user-configs*.

Now you are ready to run new LGM and HOL experiments, with a full ecosystem in each<sup>29</sup>.

---

<sup>28</sup>Try and ensure that there is blank line at the very end of the file just in case **cookie** has any problems reading it in.

<sup>29</sup>Remembering that presumably both your *base-config* and *user-config* file names are different as compared to running the non **ECOGEM** version described in the readme.

Obvious questions to investigate include not only how ecosystems and patterns of biological export may have differed between LGM and HOL, but also how patterns of nutrients ( $PO_4$  and  $Fe$ ) may have differed ... and also distributions of dissolved  $O_2$  in the ocean (and in the interior, rather than across the ocean surface). A water mass ventilation age tracer has also been simulated and the results of this will help in understanding how global circulation patterns differ between the 2 time intervals.

You can also compare between with and without **ECOGEM** versions as the only thing that changes between pairs of HOL or pairs of LGM experiments is the biological scheme<sup>30</sup>.

---

<sup>30</sup>Actually, this is not true, as in the **ECOGEM** enabled experiments, the mixed layer scheme in the ocean circulation model is also activated and which has an impact on ocean circulation. You could then create a 3rd set of HOL+LGM experiments, using the basic **BIOGEM** implicit export biological scheme, but now also setting the `go_imld` parameter to a value of 1

### Warm climates of the past

In this practical we are going to look at the ocean as it *might* have been just over 55 million years ago, at the Paleocene-Eocene Thermal Maximum (PETM) – in short a lot warmer, and with a somewhat different continental configuration and hence ocean circulation. The exercise is based on a recent 'ECOGENIE' (**cookie** configured to include **ECOGEM**) publication<sup>31</sup> which you should read first.

We are going to make the rather strong assumption that the ecosystem is structured according to exactly the same rules as in the modern ocean, and simply run the same ecological model configuration but in a new climate and ocean environment. However, because we don't really have any good (or any!) data constraints on the iron supply to the early Eocene ocean via e.g. dust, the ecological configuration does not include iron as a limiting nutrient. Bear that in mind when thinking about your results.

The *user-configs* you need to use, or copy-edit-rename, can be found in the directory:  
genie-userconfigs/MS/wilsonetal.2018  
(and you can run your experiments from here, or better, copy-edit-rename the *user-configs* you run experiments with, from the LABS sub-directory).

Read the *readme.txt* file for instructions for the basic set of command line parameters needed, but note that you may be running a *user-config* with a different name and likely from a different sub-directory. You want to spin up the following experiments first (and then experiment with them later):

- (1) Modern
- (3) Early Eocene CO<sub>2</sub> and Climate

Submit these to the cluster, but remember that you will need to recompile cookie between (1) and (3) (and re-compile before (1) if you have not been using the *base-config* cookie.CBE.worjh2.BASES immediately prior to this). Run the spin-ups for 10,000 years.

When you have completed the pair of spin-ups, see what you can diagnose and learn about how ecosystems and the spatial pattern of ecology and biological export differ between a colder modern ocean and the warmer Eocene ocean.

For example, in the warmer Eocene world:

- What has happened to the mean plankton size in different regions?
- What has happened to the fundamental niches in different size classes?
- What has happened to the realised niches?
- Is the system more or less productive?
- Has carbon export gone up or down?

See what you can find out about the two systems and think about the mechanisms that might be responsible for the differences ...

---

<sup>31</sup>Wilson, J.D., F.M. Monteiro, D.N. Schmidt, B.A. Ward, and A. Ridgwell, Linking marine plankton ecosystems and climate: A new modeling approach to the warm early Eocene climate, *Paleoceanography and Paleoclimatology*, 33, 1439–1452, DOI: 10.1029/2018PA003374 (2018).

Note that to be more comparable with the Eocene, this particular modern configuration also lacks iron co-limitation. We could also try and remove the effect of the Eocene being warmer than the modern so as to concentrate just on the effect of a different continental configuration and hence ocean circulation. The experiment (included in the *user-config* sub-directory and briefly described in the *readme* file:

(2) Late Paleocene Early Eocene Paleogeography

does exactly this, i.e. attempts to 'remove' the effect of higher ocean temperatures by running an Eocene experiment at  $\times 1CO_2$ .

Conversely, we could run modern at  $\times 3CO_2$  and then compare to the  $\times 3CO_2$  Eocene experiment. This alternative comparative experiment is provided as:

(S1) Modern with 3 x CO<sub>2</sub>

and will also require a 10,000 year long spin-up ... Note that the 2 strategies, while slightly different, are attempting the same thing (i.e. isolating the effect of paleography and ocean circulation from coeval climate change and warming).

Finally – having run some or all of these spin-ups and contrasted the results (focussing on ecological patterns and biological export, but remembering also to assess differences in ocean circulation), you can investigate the impact of geologically rapid warming a-la the PETM, on the system (esp. ecological patterns and biological export plus ocean circulation). There are two immediately obvious possible approaches:

1. You could use a  $\times 1CO_2$  spin-up as a re-start – either or both of modern and Eocene continental configurations – and run the  $\times 3CO_2$  experiment from this.

This will give you an instantaneous warming – much faster than occurred associated with PETM onset, and also faster even than modern anthropogenic warming. However, it does provide a nice simple idealized perturbation to investigate and analyse.

An experiment duration of 100, or even 10 years, might be sufficient.<sup>32</sup>

Conversely and for fun, you could also start from a  $\times 3CO_2$  spin-up, and run a  $\times 1CO_2$  experiment, to achieve a rapid cooling. Investigate the differential ecological response to rapid cooling vs. rapid warming.

2. Secondly, in the *user-configs*, you might note near the bottom is a *forcing* that determines the value of atmospheric CO<sub>2</sub>:

```
# specify forcings
bg_par_forcing_name="pyyyzz.RpC02_Rp13C02"
```

followed by a line that specifies either  $\times 1CO_2$  or  $\times 3CO_2$ , e.g.

```
bg_par_atm_force_scale_val_3=278.0E-06
```

or

```
bg_par_atm_force_scale_val_3=834.0E-06
```

(followed by a line specifying the carbon isotopic composition of the atmosphere).

---

<sup>32</sup>But running for a full 10,000 years would enable you to follow not only the initial rapid warming perturbation, but also the long-term recovery and adjustment to a new steady state.

Similar to as you have seen before for a fossil fuel  $CO_2$  emissions biogem\_force\_restore\_atm\_pCO2\_sig.dat which contains:

```
-START-OF-DATA-
0.0      1.0
999999.0 1.0
-END-OF-DATA-
```

Referring back to the instructions for changing fossil fuel  $CO_2$  emissions *forcing*, you should be able to modify this (or better, copy-rename a new forcing directory and edit the file biogem\_force\_restore\_atm\_pCO2\_sig.dat in that) to create a prescribed time-dependent change in atmospheric  $CO_2$ .

Note that in the format of biogem\_force\_restore\_atm\_pCO2\_sig.dat, the values in the second column scale the value of the parameter `bg_par_atm_force_scale_val_3` in the *user-config*. Hence, in biogem\_force\_restore\_atm\_pCO2\_sig.dat, setting a value of 2.0 in place of 1.0, will double the applied  $CO_2$  forcing. Exactly as per in the fossil fuel  $CO_2$  emissions exercises, pulses, ramps, etc etc can be constructed to control the rate and shape of the applied change in atmospheric  $CO_2$ .

Either way – assess the important and impact of the rate of  $CO_2$  rise and hence warming. (Equally, you might explore rapid cooling and how that fundamentally differs in impact from warming.)

**Marine ecology following the end Cretaceous Impact**

## 7.8 Ecology in ... fake oceans!

We can also consider the question of the causes and consequences of different ecologies in the ocean in a more general way and return to our hypothetical or 'fake' oceans.

For any of your 'fake' worlds' that you generated earlier, you can add carbon and nutrient cycling (if that was not already included), plus a marine ecology. To do this, you first need to create a new *base-config* that includes all the carbon cycling and nutrients needed by **ECOGEM** (and **BIOGEM**), then you need to create/configure a suitable *user-config*.

1. *base-config* – First, you need to enable the **ECOGEM** module.

At the top of the your *base-config* file, change the **ECOGEM** 'flag' to true:

```
ma_flag_ecogem=.TRUE.
```

Next, it is likely that you did not add any biogeochemical tracers when you created your fake world, and the *base-config* section:

```
# ****
# TRACER CONFIGURATION
# ****
```

will probably look like:

```
# the total number of tracers includes T and S
# T and S do not need to be explicitly selected and initialized
# ****
# Set number of tracers
GOLDSTEINTRACSOPTS='$(DEFINE)GOLDSTEINTRACS=2',
# list selected biogeochemical tracers
# <<<                                >>>
# list biogeochemical tracer initial values
# <<<                                >>>
```

Go find the *base-config* file cookie.CBE.p0055c.BASES in the configs sub-directory of genie-main. Copy all of the section headed

```
# ****
# TRACER CONFIGURATION
# ****
```

into your *base-config*, replacing the original contents with only 2 tracers selected.

2. *user-config* – Next, you need to define some biogeochemical cycling PLUS and ecosystem.

Go find the *user-config* file: wilsonetal.p0055c.8P8Z.pal.3x in genie-userconfigs sub-directory MS/wilsonetal.2018.

Easiest, is to simply re-use (copy-rename) the *user-config* file: wilsonetal.p0055c.8P8Z.pal.3x. The only parameters you might want to adjust<sup>33</sup>, other than a different ecological structure (and eg\_par\_ecogem\_plankton\_file), is atmospheric  $CO_2$ , which is set to  $\times 3CO_2$  by:

```
bg_par_forcing_name="pyyyyy.RpCO2_Rp13C02"
bg_par_atm_force_scale_val_3=834.0E-06
bg_par_atm_force_scale_val_4=-4.9
```

(and atmospheric  $\delta^{13}C$  to -4.9).

Strictly, the scaling for the air-sea gas exchange coefficient, for a fake world, should be:

```
# re-scale gas transfer coefficient ...
bg_par_gastransfer_a=0.722
```

(changing the value from 0.5196 to 0.722).

---

<sup>33</sup>Note that ocean alkalinity is also set for an Eocene world.

Those changes – enabling **ECOGEM** and adding ocean (and atmosphere) tracers to your *base-config*, and then taking a paleo **ECOGEM user-config** as a template to work from, should get you going with an ecology in your fake world.

If you also want to diagnose ocean circulation better and add a ventilation tracer, then in the *base-config*, increase the number of selected tracers by 2 (under # Set number of tracers) and add the following 2 lines to the list of selected tracers:

```
gm_ocn_select_48=.true. # r  
gm_ocn_select_49=.true. # b
```

and ... in the *user-config*, add (anywhere, but e.g, in the MISC parameter section):

```
# add ventilation tracers  
bg_ctrl_force_ocn_age=.true.
```

## 7.9 EcoGEnIE 1.1

As outlined in the Sections: *Choosing a template user-config* and *Configuring cookie experiments*, a number of recommended changes to the configuration of **ECOGEM** as published by *Ward et al.* [2018] (EcoGEnIE 1.0).

These changes have been incorporated into an example *user-config* (in the *genie-userconfigs* sub-directory *EXAMPLES*: *cookie.CB.world4.BASESFeTDTL.ECOGEM\_NEW.SPIN* (and paired with the *base-config*: *cookie.CB.world4.BASESFeTDTL*), and are:

- Under \*\*\* REMINERALIZATION \*\*\*:

```
# set 'instantaneous' water column remineralization
bg_par_bio_remin_sinkingrate_physical=9.9E9
bg_par_bio_remin_sinkingrate_reaction=125.0
```

which instantaneously remineralizes all particulate organic matter throughout the water column according to the remineralization profile and/or reaction rates. This is activated via a 'very large' value for *bg\_par\_bio\_remin\_sinkingrate\_physical*. At the same time, reaction rates (including scavenging) are calculated as if the sinking rate was finite and equivalent to the value of: *bg\_par\_bio\_remin\_sinkingrate\_reaction* ( $m\ d^{-1}$ ).

- Under \*\*\* MISC \*\*\*:

```
# set mixed layer to be only diagnosed (for ECOGEM)
go_ctrl_diagmld=.true.
# add seaice attenuation of PAR
eg_ctrl_PARseaicelimit=.true.
# relative partitioning of C into DOM
eg_par_beta_POCtoDOC=0.75
```

which firstly substitutes a diagnosed mixed layer depth, rather than actually applying mixed layer physics to ocean circulation, then limits light available under sea-ice in proportion to the fractional sea-ice cover in that grid cell, and lastly, re-partitions carbon from POM to DOM and is tuned to produce an approximately Redfield ratio (106) of 104.7 : 1 in *C* : *P* of exported POM., and then:

```
# maximum time-scale to geochemical reaction completion (days)
bg_par_bio_geochem_tau=90.0
# extend solubility and geochem constant T range (leave S range as default)
gm_par_geochem_Tmin = -2.0
gm_par_geochem_Tmax = 45.0
gm_par_carbchem_Tmin = -2.0
gm_par_carbchem_Tmax = 45.0
```

which firstly, limits the maximum consumption of any particular reactant by a single reaction, to an imposed lifetime (90 days), with the remaining parameters extending the valid temperature range for solubility and geochem constants to  $-2 - 45^{\circ}\text{C}$ . Note that the valid salinity range is left unchanged.

For reference, the original (i.e. as per in *Ward et al.* [2018]) but reformatted, user-config, is provided as: *cookie.CB.world4.BASESFeTDTL.ECOGEM\_NEW.SPIN* (also in the *genie-userconfigs* sub-directory *EXAMPLES*).

## 7.10 Further ideas and investigations

### The role of iron limitation

Up to this point, we have included both phosphate and iron as limiting nutrients. You might want to know more about the importance (or not) of iron availability in determining patterns of biological productivity.

1. You can get a more exact picture of the nutrient limitation terms via the netCDF output variables: `eco2D_xGamma_Fe_001` and `eco2D_xGamma_P_001`.

These two variables take values of between 0 and 1. A 1 indicates that the factor is not limiting to growth. A 0 indicates the factor is completely preventing growth.

- In what regions are iron and phosphorus more or less limiting to growth?
- In regions where neither is limiting, what other factors might be important?

2. Plankton stoichiometry plays a critical role in determining which nutrient is most limiting to growth. You can increase the plankton  $Fe : C$  ratio by increasing the minimum and maximum iron quotas. Look at the parameters `eg_qminFe_a` and `eg_qmaxFe_a` in the *user-config* file.

- What happens to the ecosystem if you increase these parameters by a factor of 2, 5 or 10?
- How does a change in these parameters affect the model behaviour?
- What has changed in terms of the patterns of nutrient limitation?
- What has happened to the concentration of the limiting and non-limiting nutrient?

NOTE: Rather than changing the parameter and subsequently forgetting what you started with ... instead, you might copy/paste a new version of the line in question, and comment out the original by placing a '#' at the beginning of the line. For example:

```
#eg_qminFe_a = 3.0e-6
eg_qminFe_a = 6.0e-6
```

changes the minimum iron quota by a factor of 2, whilst keeping a record of the original setting (inactivated by the #). Or, you might do something like:

```
eg_qminFe_a = 6.0e-6 # 3.0e-6
```

that reduces the total number of lines you end up with in the *user-config* file.

3. Nutrient supply ratios are also important in determining the limiting nutrient.

The `bg_par_det_Fe_sol_exp` parameter determines the solubility of atmospheric iron inputs in seawater. Decreasing the value of `bg_par_det_Fe_sol_exp` will therefore decrease the iron-to-phosphorus supply ratio.

- What happens to the ecosystem if you decrease `bg_par_det_Fe_sol_exp` by e.g. 10, 20 or 50%?

4. Lastly, you can turn off entirely the iron requirement of plankton:

```
# include cellular quota: Fe -- no Fe limitation enabled
eg_useFe = .FALSE.
eg_fquota = .FALSE.
```

- What happens to the patterns of biological productivity as well as global total export?

(none)



## 8. The geological cycle of carbon

**8.1 (none)**

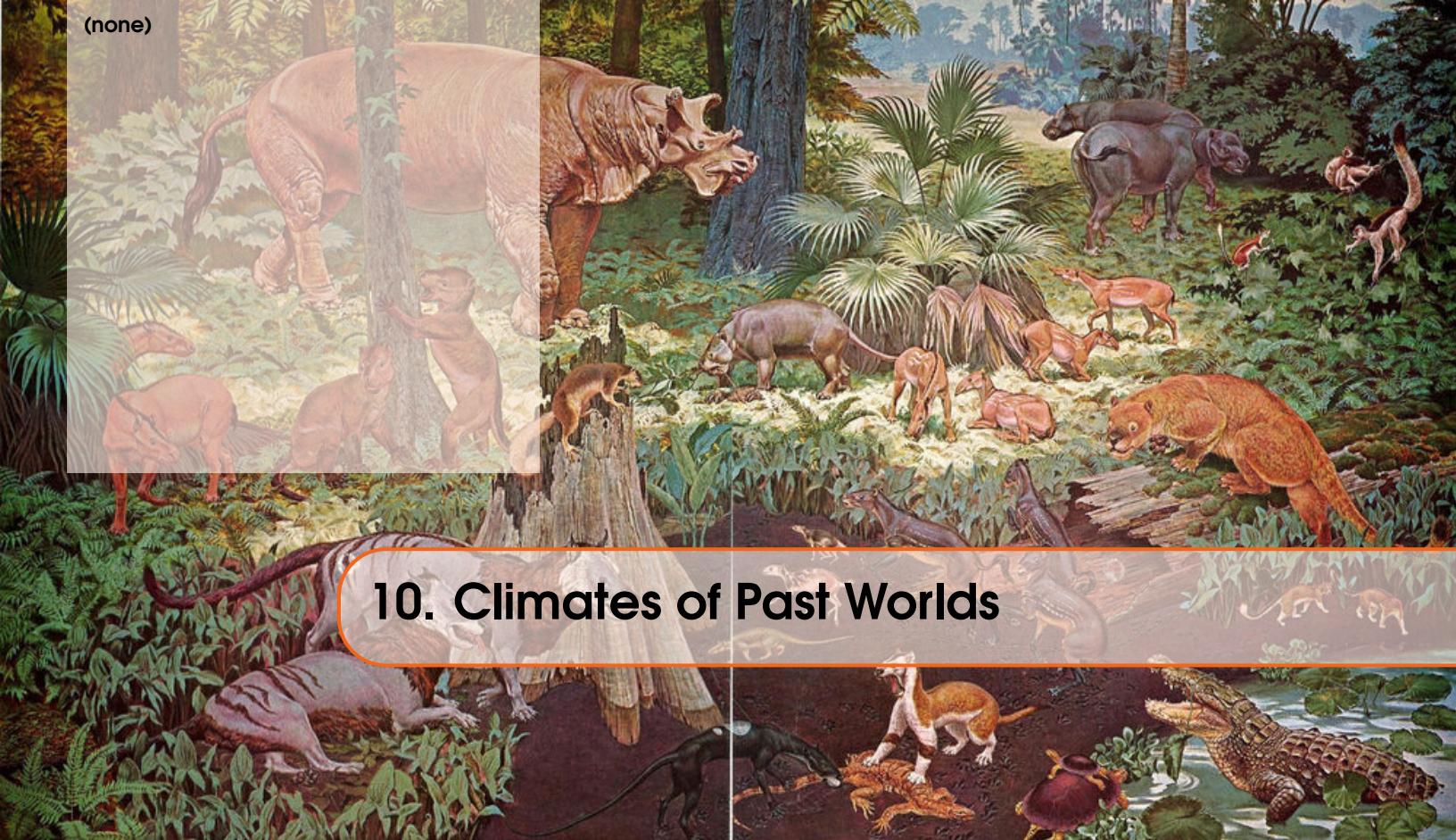
(none)



## 9. The land surface and carbon-climate feedbacks

**9.1 (none)**

(none)



## 10. Climates of Past Worlds

**10.1 (none)**

## 11. Proxies and Reconstructing the Past

**11.1 (none)**

(none)

## 12. Synthesis: carbon in the Earth system

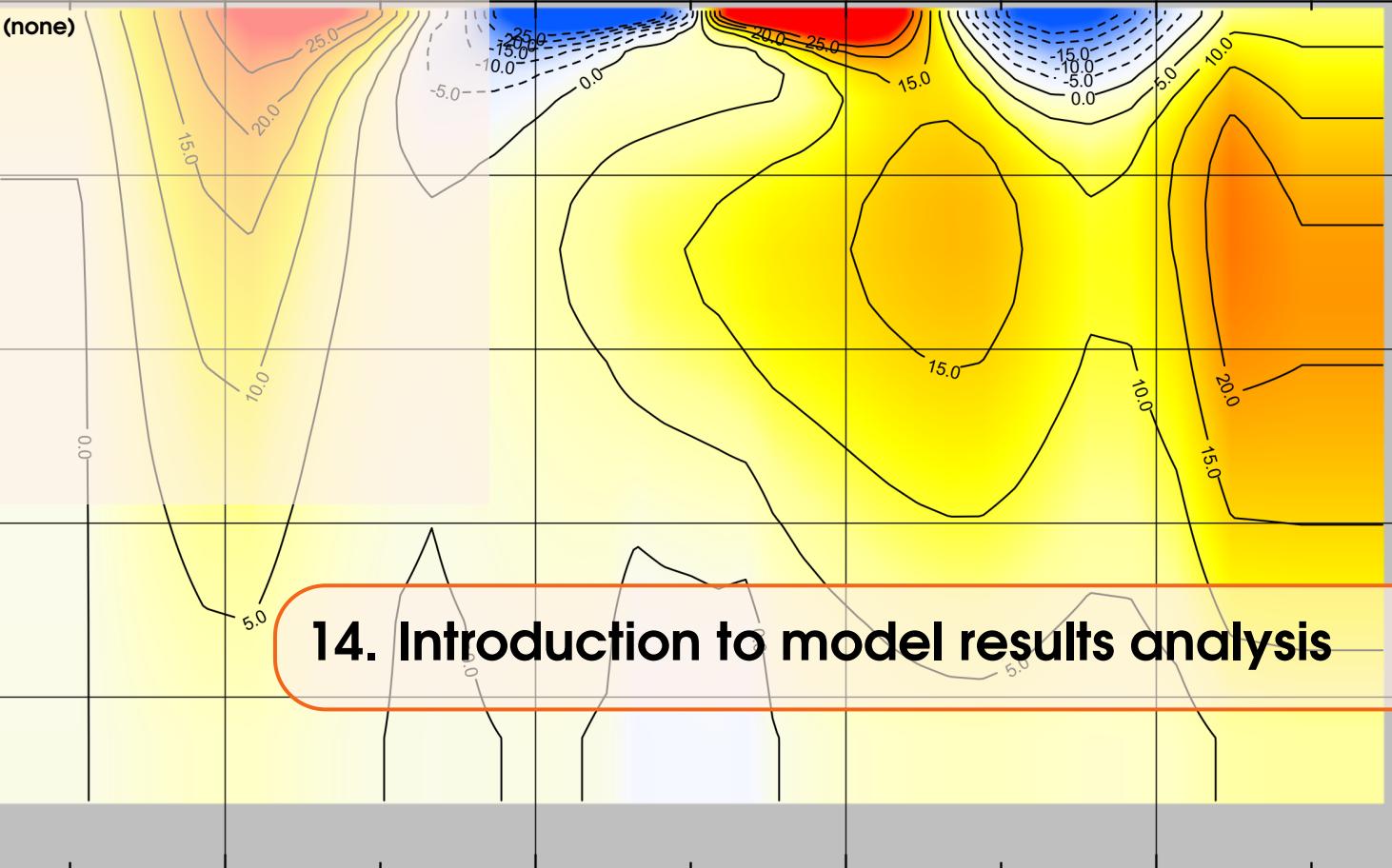
**12.1** (none)

(none)

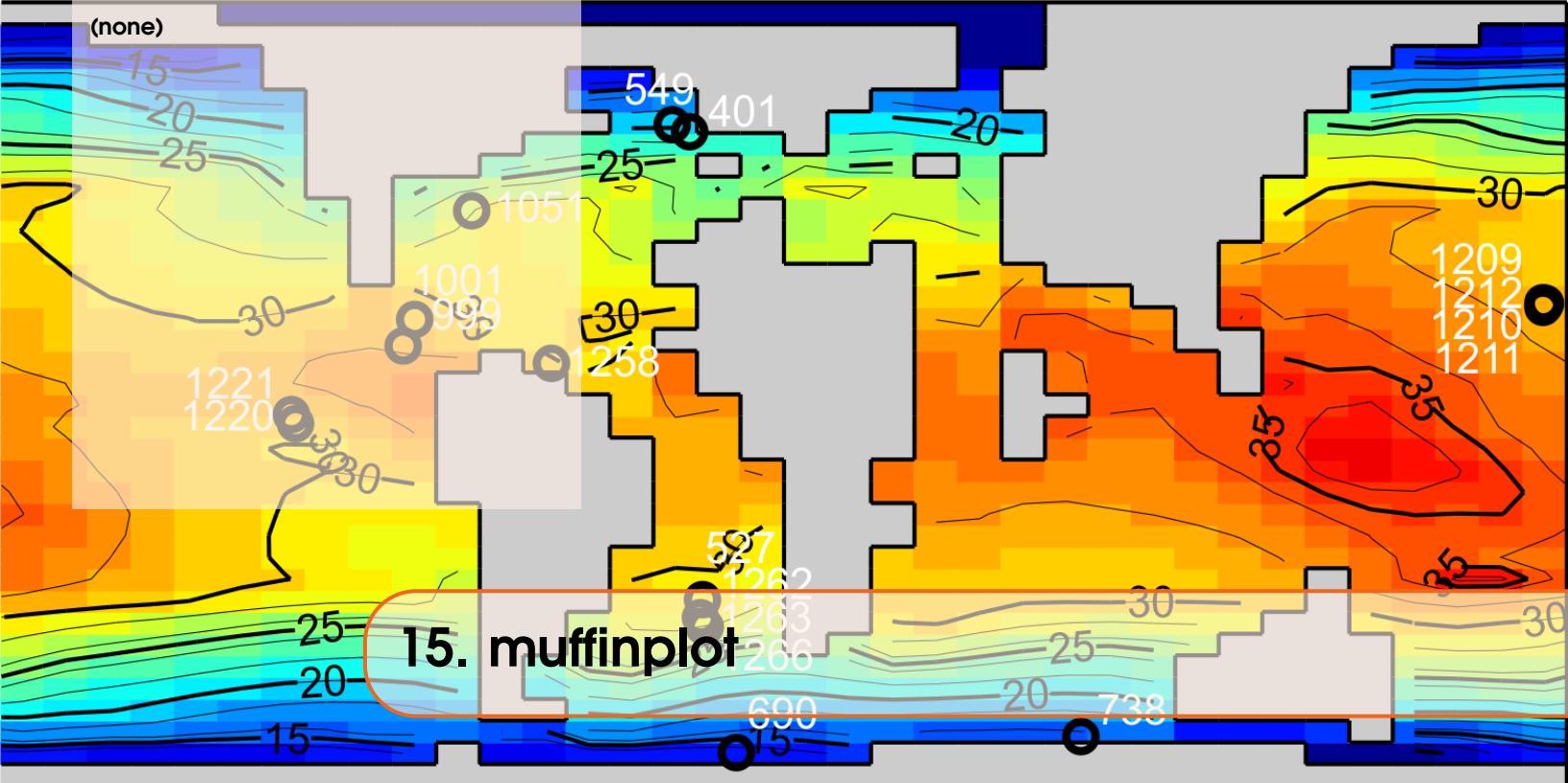


### 13. cookie model output

**13.1 (none)**



**14.1 (none)**



**15.1 (none)**

(none)

## 16. muffindata

**16.1 (none)**

**Introduction**  
  cookiegen  
  Installing cookiegen  
**Quick Start Guide**  
  Running cookiegen  
  Configuring cookiegen  
**Overview**  
  cookiegen configuration examples  
**Creating base- and user-configs for cookiegen**  
  cookiegen parameter settings – details  
**Appendix – pre-generated Worlds**  
  Conceptual worlds  
  Idealized-continent worlds  
  Modern worlds

## 17. cookiegen

## 17.1 Introduction

### 17.1.1 cookiegen

**cookiegen** is a collection of **MATLAB**<sup>1</sup> functions that can create all the primary configuration files required by the cGENIE (**cookie**) Earth system model. **cookiegen** was originally designed to take the output from a fully coupled GCM, particularly of past climates with different continental configurations, and then re-grid the output needed by the **cookie** model (in the form of files of boundary conditions, all saved in their respective correct format). However, **cookiegen** can also be used to 'draw' conceptual alternative Earths (in terms of continental configuration) and modify existing configurations.

### 17.1.2 Installing cookiegen

The **Summary** code is hosted on **GitHub**:

<https://github.com/genie-model/cookiegen>

There are 2 ways to get your mitts on the **cookiegen** code:

1. By downloading an archive file, containing all the code etc. For this – click on the **green** Clone or Download button, and select Download ZIP.  
You then unpack/unzip the files and directory structure where you want it.  
This [archive download] is the simplest and perfectly workable way to proceed<sup>2</sup> (and is the recommended option).
2. Or you can *clone* the repository to where you intend to run **cookiegen**. Note that you will need a git client installed on your computer. There are GUI clients for git, or this can be done at the command line:

```
$ git clone https://github.com/genie-model/cookiegen.git
```

By doing this, you have created your own code repository (and an identical copy of the one hosted on GitHub). As part of the `git clone` command, you also automatically *check out* (from your very own personal repository) a copy of the code.

Note that you download or clone **cookiegen** to the computer that you have **MATLAB** installed on and will use to run **cookiegen** (i.e. not necessarily to the computer where **cookie** itself is run).

---

<sup>1</sup>As a **MATLAB** code, (only) very basic familiarity with using **MATLAB** at the command line is required. And a copy/license of the **MATLAB** software ... A sufficient grasp of **MATLAB** can be gained by going through the following sections/subsections of the matlabbananas textbook, which can be found on GitHub (look for the PDF file compiled from the latex source – BANANAS.pdf):

- Section 1.1 – The **MATLAB** interface and command line.
- Section 1.6 – Changing directories and file search paths.
- Section 2.2 – Using *functions* ... of which **cookiegen** is one.

<sup>2</sup>Note that this way, you will be unable to easily update the code with whatever new developments or bug fixes occur in the future, nor can propagate back any code changes that you might have made and might want to become part of the official **cookiegen** code (i.e. downloading the zip file becomes a one-off installation that loses its formal connection to the **GitHub** repository).

**17.2****Quick Start Guide**

There are 5 steps in using **cookiegen** to create and ready a new model configuration to be run. The basic steps and quick start methodology is outlined below.<sup>3</sup>

1. Select an existing **cookiegen** configuration – these live in the CONFIGS folder. The most common options are:

- EXAMPLE\_BLANK – Generates a continental configuration starting from an all-ocean (no land) ('water-world') grid. Requires the user to edit in the location of any land (and optionally also modify seafloor topography).
- EXAMPLE\_MASK\_INPUT – Generates a continental configuration from a simple text file containing a grid of '1's (ocean) and '0's (land). By default, this file lives in INPUT.<sup>4</sup>
- EXAMPLE\_GCMHadCM3L\_modern – Generates a continental configuration derived from the model output of a coupled GCM experiment for the present-day.
- EXAMPLE\_GCMHadCM3L\_paleo – Generates a continental configuration derived from the model output of a coupled GCM experiment for the late Permian. (So, identical in operation to EXAMPLE\_GCMHadCM3L\_modern.)

Copy and rename the configuration file closest to what you are after. Make sure that at the minimum, you specify a new 'name' for the continental configuration ('World') you are about to generate – in the configuration file, the parameter name for this is: par\_wor\_name – This must be an 8-character long string.<sup>5</sup>

2. Run the **cookiegen** function, passing the name of a file containing the parameters defining the creation of your new configuration, i.e.,

```
>> makecookie('FILENAME')
```

where FILENAME is the name of your configuration file.

Depending on whether opt\_user is set to true, an editor window will open (firstly for changes to the land-sea mask to be made, then later for the ocean bathymetry to be edited) and which requires manual intervention (mouse clicks from you) until **cookiegen** can continue and complete. Refer to the instructions that appear on the MATLAB command line.

Using the default settings, all the files you need will appear in the OUTPUT folder and the model will be run-able after just 3 more short steps:

3. A sub-folder will appear in OUTPUT with the name of your configuration.  
(The name of your configuration is an 8-character string which is defined in the **cookiegen** configuration file [parameter name: par\_wor\_name].)  
Copy the entire sub-folder to the genie-paleo folder of your **cookie** installation.

---

<sup>3</sup>In practice and particularly for more specific usage, some decisions regarding the *base-config* (tracer selection) and *user-config* and some copy-pasting will be required (detailed later).

<sup>4</sup>The file example provided is: wordrake.txt.

<sup>5</sup>You can use under-scores, but not spaces and not '-'. Do not use fewer than or mroe than 8 characters ...

4. In the example **cookiegen** configuration files, a generic *base-config* file suitable for a 16-level (36x36) ocean has been specified [parameter name: `par_cfgid`]. This file has been copied to `OUTPUT` (from `DATA`) and renamed incorporating the 8-character 'World' name you specified (and with the file extension `.config`).

Copy this file to the `genie-basconfigs` folder of your `cookie` installation.

- Also the example **cookiegen** configuration files, a generic *user-config* suitable for running a basic climate-only (no carbon cycle) experiment has been specified [parameter name: `par_usrid`]. This file has been copied to `OUTPUT` (from `DATA`) and renamed incorporating the 8-character 'World' name you specified (and with the file extension `.SPIN`).

Copy this file to the `genie-userconfigs` folder (or sub-folder) of your `cookie` installation.

If you named your 'World' e.g., MY\_WORLD and transferred the *user-config* into the LABS sub-directory of genie-userconfigs, you would run a generic climate-only experiment (for 10 years):

```
./runcookie.sh cookie.C.MY_WORLD.NONE.LABS cookie.C.MY_WORLD.NONE.SPIN 10
```

Note that in changing to a new *base-config*, the model will automatically re-compile before actually running the experiment.

If you want something beyond a generic/default climate-only configuration and experiment, steps 4 and 5 (above) need to be slightly modified. Some suggestions follow (also read the rest of the Chapter!!).

+ 'age' To better diagnose the large-scale circulation of the ocean, you can add the capability to simulate 'ventilation age' – see **Section 2.4**. To do this:

- From cGENIE.baseconfigs, open the file TRACERCONFIG.AGE.txt – select and copy all the lines. Open the *base-config* file that cookienet generated, delete the lines in-between:

1

and the paste in the contents of TRACERCONFIG.AGE.txt that you have just copied.

LAB.2.1.EXAMPLE would also work.

In place of the automatically-generated

- which you can find in the sub-directory: CGENIE.userconfigs.

5. Replace the default tracer definition text block in your base configuration file with the following:

- 3. Replace the default tracer definition text block in your base config with the contents of TRACERCONFIG.CARBONATECHEMISTRY.txt**  
**(Refer to the the +age modification instructions (above) for further details.)**

6. Use the example *user-config* USERCONFIG.CARBONATECHEMISTRY.SPIN (from the CGE-NIE.userconfigs sub-directory).  
LAB.5.1.EXAMPLE (or LAB.5.1.historical or LAB.5.1.emissions) would also work.

<sup>6</sup>If you have not transferred this file to the cluster, do this now. Either way, it is recommended that you change the name form the default generated file, e.g. change NONE → AGE in the file-name.

+ biopump If you want a simple single nutrient (P) biological pump in the ocean, following *Ridgwell et al.* [2007] (for an 8-level non-seasonal ocean) or *Cao et al.* [2009] (for a 16-level seasonally-forced ocean):

5. Replace the default tracer definition text block in your *base-config* with the contents of TRACERCONFIG.Ridgwelletal.2007.txt (for an 8-level ocean) or TRACERCONFIG.Caoetal.2009.txt (16-level ocean).  
(Refer to the the +age modification instructions (above) for further details.)
  6. From the CGENIE.userconfigs sub-directory, in place of the automatically-generated *user-config*, use: USERCONFIG.Ridgwelletal.2007.txt or USERCONFIG.Caoetal.2009.txt. LAB.6.1.EXAMPLE (8-level) or LAB.6.2.EXAMPLE (16-level) would also work.
-

### 17.3 Running cookiegen

To use the **cGENIE.cookie** model configuration generator **cookiegen** – at the command line in **MATLAB**, simply type:

```
>> makecookie('FILENAME')
```

where **FILENAME** is the name of a configuration file with an **.m** file-name extension, that specifies the required settings (see subsequent section).

**FILENAME.m** must be present in the directory where you installed **cookiegen** (the directory where e.g. where the file (**makecookie.m** is located). Or, if you would like to store it elsewhere, you will need to 'add' its location **MATLAB**'s search path by the **MATLAB** function **addpath**.

The **cookiegen** model configuration generator then starts, and depending on the specific settings in the configuration file, may require no user input, or may require user input (either because this option was requested, or because a re-gridding issue arose that requires manual intervention to resolve). A series of plots are created (and saved) as the configuration generation progresses together with the **cookie** model configuration files themselves. All the various steps plus details of how the contents of the configuration files are generated are reported at the command line, and saved in an ASCII format **\*.log** file for future reference.

Depending on the specific configuration file settings, **cookiegen** has 5 main modes of operation. The most common options are:

#### 1. Configuration based on re-gridding climate output from a GCM.

The most common usage of **cookiegen**, and enabling a new (typically paleo) configuration to be derived from the output of a GCM experiment. Currently options for utilizing 4 different GCMs are provided: HadCM3(l) ('hadcm3' or 'hadcm3l'), FOAM ('foam'), CESM ('cesm'), and ROCKEE-3D ('rockee').

2D wind-stress and speed + zonal planetary albedo, are derived from GCM climate fields.

#### 2. From a blank (all ocean) initial template.

This option – '' (empty string) – enables a topography to be drawn by hand within **cookiegen** and hence represents an interactive alternative to (3) (editing a land-sea mask in a text editor). Idealized zonal wind-stress and speed, and zonal planetary albedo, are used.

#### 3. From a simple land-sea mask file.

This option – 'mask' – will create a new configuration from any specified land-sea mask, whether 'real' or hypothetical. It provides a different way of 'drawing' continents – now via a text file beforehand, rather than 'by hand' drawing during the running of **makecookie**.

Idealized zonal wind-stress and speed, and zonal planetary albedo, are used.

Less common options are:

#### 4. Derivation based on an existing model topography '.k1' (or simplified .k2) file.

This option – 'k1' or 'k2' – allows a topography to be re-created, or adapted/altered, all directly from an existing model 'k1' file (and hence existing model configuration).

Idealized zonal wind-stress and speed, and zonal planetary albedo, are used.

#### 5. Derivation based on re-gridding a high resolution input topography.

In this option – 'mat' – a land-sea mask and ocean bathymetry are re-gridded from a high(er) resolution **.mat** (**MATLAB**) topography file.

Idealized zonal wind-stress and speed, and zonal planetary albedo, are used.

## 17.4 Configuring cookiegen

### 17.5 Overview

When **makecookie** is run, a configuration file with a specified filename<sup>7</sup> is loaded. The configuration file has a simple plain text (ASCII) format, but is given a .m extension, enabling the values of a number of controlling parameter values to be set directly in **MATLAB**.<sup>8</sup> The configuration file parameters control facets of **makecookie** behavior such as the primary mode of operation, input and output filenames, what types of configuration files you want to generate, as well as there being a number of parameters controlling the finer details of re-gridding and configuration file generation, including whether to enable user-input or not.

The configuration files can be edited with the **MATLAB** editor (indeed, as a .m file, they are inherently a **MATLAB** format file and associated with this program), or any text editor<sup>9</sup>. The configuration file is divided up into a series of sections of different parameter options. The main (most commonly used) parameters are summarized as follows (and are more fully described later). Highlighted in blue are the most commonly changed parameters.

```
% *** CONFIG NAME AND INPUT DATA SETTINGS ****%
par_wor_name
Defines the name of the model configuration. This must be a string and it must 8
characters long.
e.g. par_wor_name='my_world'
par_gcm
Defines the format of the input. The string is blank for an interactive user-defined
world.10
- par_expid
Defines the name of the GCM experiment if a GCM input is selected, or the land-sea
'.dat' if using the mask input option (or the '.k1' or '.k2' file). By default, these files (or
folder, in the case of GCM model input) live in the INPUT sub-directory.
- par_cfgid
Specifies an optional template base-config. If specified, 'makecookie' will copy the
base-config file – automatically add the required parameter block defining the new
continental configuration – and save this to the output directory (par_pathout). If
empty, par_cfgid='', a base-config file will have to be manually selected and the
continental configuration parameter block copied-pasted into it.
- par_usrid
Specifies an optional template user-config. If specified, 'makecookie' will copy the
user-config file and save it to the output directory (par_pathout).
- par_age
Defines the geologic age of the configuration (in Ma). This is used to modify the value
of the solar constant, as well as initial ocean concentrations of Mg, Ca, SO4.
```

<sup>7</sup>i.e. the single parameter passed to **makecookie** when it is invoked at the command line.

<sup>8</sup>Note that the parameter filename is passed as a string without the .m extension (which is implicitly assumed). An error message will be generated if the file does not exist or has the incorrect extension.

<sup>9</sup>But make sure they retain a .m rather than .txt filename extension

<sup>10</sup>If you utilize one of the example configuration files in the CONFIGS sub-directory, you will not need to change this.

```
% *** INPUT + OUTPUT SETTINGS ****%
- par_pathin
    Specifies the sub-directory where the input files can be found.11
- par_pathout
    Specifies the sub-directory where the cookie configuration files will be saved to.12
- par_plotformat
    Specifies the format of the plots that are created by makecookie. If you need better
    quality figures saved, set this parameter to pdf. 13
- opt_user
    Selects whether or not you wish to have the chance to manually edit the land-sea mask
    and the seafloor topography. Valid options are true and false. (Even if this option is
    selected, you need not do any editing.)
```

---

```
% *** GRID -- HORIZONTAL ****%
- par_max_i [default: 36]
    Defines the number of grid points in the longitude ('i') direction.
    This is typically 36 or 18, e.g. par_max_i=36
- par_max_j [default: 36]
    Defines the number of grid points in the latitude ('j') direction.
    This is typically 36 or 18 and is typically the same number as for the i direction, e.g.
    par_max_j=36
- opt_equalarea [default: true]
    This specifies whether or not an equal area grid is used/assumed. (A value of false
    results in the latitude grid being defined in equal increments of latitude).
- par_lon_off [default: -180.0]
    This specifies the longitude offset of the model grid. The basic modern configurations
    have an offset of -260.0. The standard paleo value is -180.0.
```

---

```
% *** GRID -- vertical ****%
- par_max_k
    Defines the number of layers in the ocean circulation model.
    This is almost always either 16 (e.g., par_max_k=16) or 8.
- par_max_k_shallow
    Defines the k (ocean level) value of the shallowest allowed seafloor. This is primarily
    used to exclude cells where the ocean model grid is only a single layer thick when
    re-gridding from a GCM or high resolution topography.
- par_min_k [default: 1]
    Defines the k (ocean level) value of the deepest part of the ocean. Choosing a value of 3
    for a 16-level flat-bottom ocean gives approximately the modern ocean volume.
- par_max_D [default: 5000.0]
    Sets the maximum ocean (scale) depth (in m).
    This is almost invariably 5000.0, but some series of paleo configurations use 5500.0 m.
```

<sup>11</sup>This need not be changed.

<sup>12</sup>This need not be changed.

<sup>13</sup>Note that makecookie will take much longer to run when saving vector graphics.

```
% *** MODULE SUPPORT ****%
- opt_makegold [default: true]
  Create the ocean circulation model files?
  If you only want to (re-)generate sediment model (SEDGEM) or land surface scheme
  (ENTS) files, you might want to set this to false.
- opt_makeseds [default: false]
  Generate sediment model (SEDGEM) files?
- opt_makeents [default: false]
  Generate land surface scheme (ENTS) files?

% *** BOUNDARY CONDITION SIMPLIFICATIONS ****%
- opt_makezonalwind
  Determines whether cookiegen applies an idealized zonal wind boundary conditions.
  For non-GCM-derived configurations this is the only option. For GCM-derived configu-
  rations, this allows the assumption of zonal wind boundary conditions to be evaluated
  against a 2D re-gridded wind product.
- opt_makezonalalbedo
  Determines whether cookiegen applies an idealized planetary albedo boundary condi-
  tion. For non-GCM-derived configurations this is the only option. For GCM-derived
  configurations, the default (true) is for a zonal planetary albedo boundary condition
  derived from the GCM and consistent with the standard model configurations. A false
  will create and apply a 2D re-gridded planetary albedo product.
- opt_makerndsseds
  Determines whether cookiegen creates a randomized seafloor topography for the purpose
  of calculating pressure-dependent carbonate chemistry and carbonate saturation at the
  sediment surface. Otherwise, seafloor topography is the same as that used by the physical
  ocean circulation model.
```

### 17.5.1 cookiegen configuration examples

Several example configurations are provided in the CONFIGS sub-directory. These illustrate parameter settings and the primary ways of using **cookiegen**. They also serve as useful templates for creating your own **cookiegen** configuration files. In alphabetical order, but with the commonly used examples highlighted in blue, they are:

- EXAMPLE\_BLANK – Starts from a blank ('water-world'), all-ocean (no land) template. User input (allowing continents and seafloor topography to be 'drawn') is automatically activated whether you want it or not!<sup>14,15</sup>
- EXAMPLE\_GCMHadCM3L\_modern – Takes the continental configuration and climate simulation output from a fully coupled GCM experiment (files in directory GCMHadCM3L in the **cookiegen** sub-directory INPUT) and derives full **cookie** model boundary conditions. User input (allowing continents and seafloor topography to be 'edited') is activated by default.
- EXAMPLE\_GCMHadCM3L\_modern\_adjusted – A variant on the above with an additional parameter setting included in the configuration file:  
`par_mask_name='mask.modern_adjusted'`  
 This mask file specifies grid locations to be swapped between land and ocean and hence forgoing the need to make adjustments to the land-sea mask interactively.
- EXAMPLE\_GCMHadCM3L\_paleo – Takes the paleo continental configuration and climate simulation output from a fully coupled GCM experiment (files in directory GCMHadCM3L in the **cookiegen** sub-directory textsf INPUT) and derives full **cookie** model boundary conditions. User input (allowing continents and seafloor topography to be 'edited') is activated by default.
- EXAMPLE\_K1file\_INPUT – Starts from a cookie model configuration 'k1' topography defining file: work2in\_.k1, which is stored in the **cookiegen** sub-directory INPUT. The k1 file defines a late Permian continental configuration (and bathymetry). User modification is enabled by default (so MATLAB pauses for modifications to be made), but this can be disabled (opt\_user=false).
- EXAMPLE\_K2file\_INPUT – Starts from a cookie model configuration 'k2' topography defining file: work2in\_.k2, which is stored in the **cookiegen** sub-directory INPUT. The k2 file defines an early Eocene continental configuration (and bathymetry). User modification is enabled by default (so MATLAB pauses for modifications to be made), but this can be disabled (opt\_user=false).
- EXAMPLE\_MASK\_INPUT – A conceptual world (loosely following the literature) generated from a land-sea mask (wordrake.dat) which is stored in the **cookiegen** sub-directory INPUT. The land fraction is minimal (as it is trying to reproduce the sort of zero-area numerical barrier used in the literature). This configuration has a barrier to ocean circulation, from the N pole down, with a high southern latitude gateway (a Drake Passage like feature).

---

<sup>14</sup>Because it assumes that you are going to go on and draw something ...

<sup>15</sup>Edit the **cookiegen** code to be able to turn off.

## 17.6 Creating base- and user-configs for cookie

Having created a new **cookie** configuration using the **MATLAB** function `cookiegen.m`, unless you are using automatically-generated ocean-only *base-* and *user-configs* files (see earlier), you are now going to need to create these configuration files before you can actually run anything<sup>16</sup>:

1. You need to create a *base-config*, which includes the parameter settings of your new world. There are 4 sub-steps involved in this:
  - (a) Choosing a template *base-config* file, which specifies appropriate physics parameters.
  - (b) Copy-pasting a list of parameter settings that has been automatically-generated by `cookiegen` into the template *base-config* file.
  - (c) Defining the number of dissolved and particulate tracers in the ocean (and gases in the atmosphere) that you want and entering this into the *base-config*.
  - (d) Ensure that the appropriate science modules are enabled in the *base-config*.
2. You need to create a *user-config*, whether simply to test the configuration or to provide the basis of a series of experiments.

A number of templates are provided and more detailed instructions follow. However, note that if you have similar *base-config* and *user-config* files already, then you can simply adapt these. Another option is to utilize *base-config* and *user-config* files associated with any of the published experiments listed in `genie-userconfigs/MS`, or any of the described EXAMPLE.\* example configurations.

And ...

3. Ensure that all files are in their correct location in the **cookie** directory structure.
- 

### 1a. Choosing a template *base-config*

If you do not want to adapt an existing *base-config* (see suggestions above) or want to start 'fresh', a pair of alternative template files are provided with **cookiegen** (copy and rename the file, and then edit). The template files are missing the definitions of the world settings and tracers, which will be rectified in steps 1b and 1c, respectively. They also only specify a basic set of science modules (addressed in step 1d). Otherwise, they differ only in the tunable physics parameter values, under the heading:

```
# ****
# PHYSICAL CLIMATE CALIBRATION
# ****
```

The available template files are:

- `BASECONFIG.08lvl.config`

This is based on the physics calibration as used in e.g. *Ridgwell et al. [2007]* and *Ridgwell and Hargreaves [2007]* with the parameters tuned for an 8-level ocean modern world (specifically: `worbe2`). For fake and paleo worlds with an 8-level ocean, this is the recommended template.<sup>17</sup>

<sup>16</sup>Note that steps 1c, 1d, and 2 are not independent of each other and need to be consistent – e.g. if you define Fe co-limitation of biological productivity in step 2, you need to have selected the appropriate tracers in step 1c. And if you specify biological productivity calculated by **ECOGEM** in step 2, you need to have selected the **ECOGEM** module in step 1d.

<sup>17</sup>If you want exactly the same physics as per e.g. *Ridgwell et al. [2007]*; *Ridgwell and Hargreaves [2007]* , use: `cgenie.eb_go_gs_ac_bg.worbe2.BASES.config` as a template.

The differences compared to the published usages are:

- The ocean starts warmer ( $5^{\circ}C$ )<sup>18</sup> than the default ( $0^{\circ}C$ ). This helps the ocean circulation come to an equilibrium state more rapidly, particularly for greenhouse climates:

```
# temp0 -- start with a warm ocean
go_10=5.0
# temp1 -- start with a warm ocean
go_11=5.0
```

Note that both parameters should be changed together (one sets the Northern Hemisphere temperature, and one the Southern ... or something ...).

- There is no (salt / freshwater) flux adjustment, with the scaling parameter set to zero:

```
# SclFWF -- scale for zero freshwater re-balancing
ea_28=0.0
```

Note that the built-in (salt / freshwater) flux adjustment is designed to maintain a strong AMOC and the code is specific to the original modern continental configurations (e.g. worbe2, worjh2). A (salt / freshwater) flux adjustment using this parameter and associated code should never be used for paleo or fake worlds.<sup>19</sup>

- The sea-ice diffusivity parameters are adjusted for improved stability<sup>20</sup>:

```
# reduced sea-ice eddy diffusivity
gs_11=1000.000
# set a fractional sea-ce coverage threshold for preventing advection
gs_par_sica_thresh=0.99
```

- Insolation forcing is seasonal<sup>21</sup>:

```
# set seasonal cycle
ea_dosc=.true.
go_dosc=.true.
gs_dosc=.true.
```

Note that all the respective parameter for all three physical climate components (ocean circulation model, sea-ice model, EMBM atmosphere) needs to be set.

- An additional option is provided for toggling between isopycnal/diapycnal (.true.) and horizontal/vertical (.false.) mixing schemes:

```
# it is recommended that it is turned OFF (= .false.) for 'fake' worlds
go_diso=.true.
```

In all modern and 'realistic' paleo configurations, isopycnal/diapycnal mixing is used. However, this scheme can lead to unwanted negative tracer values. In extreme biogeochemical configurations, particularly at low atmospheric  $pO_2$  and when sharp horizontal transitions in redox state in the ocean occur, the magnitude of negative values can become unacceptable.

Work with fake worlds, where no ocean like that may have ever existed on Earth and hence there is no specific ocean circulation pattern to try and reproduce, the recommendation is to use the horizontal/vertical (.false.) mixing scheme.

---

<sup>18</sup>These parameters could potentially be increased further (to say ...  $10^{\circ}C$ ).

<sup>19</sup>Instead, a salt forcing in the desired deep-water formation location, balanced by a freshwater flux elsewhere, can be defined.

<sup>20</sup>The original parameter setting was:

gs\_11=6200.000

<sup>21</sup>The original tuned 8-level ocean circulation model configuration had no seasonal cycle (options set to .false.).

- BASECONFIG.16lvl.config

This is based on the physics calibration as used in e.g. *Cao et al.* [2009] with the parameters tuned for a 16-level ocean modern world (specifically: `worjh2`). For fake and paleo worlds with an 16-level ocean, this is the recommended template.<sup>22</sup> This includes the same changes as per made for the 8-level ocean circulation model based configuration:

- The ocean starts warmer ( $5^{\circ}\text{C}$ ).
- There is no (salt / freshwater) flux adjustment.
- The sea-ice diffusivity parameters are adjusted for improved stability<sup>23</sup>.
- Again as per for the 8-level ocean template *base-config*, an option is highlighted for toggling between isopycnal/diapycnal (`.true.`) and horizontal/vertical (`.false.`) mixing schemes:

```
# it is recommended that it is turned OFF (=.false.) for 'fake' worlds
go_diso=true.
```

As per for the 8-level ocean configuration, it is recommended to implement simple horizontal/vertical (`.false.`) mixing.

Finally, it should be noted that compared to the *Cao et al.* [2009] configuration, there is no modification of atmospheric diffusivity over the Southern Ocean and Antarctica, which was added to improve the seasonal sea-ice properties in the Southern Ocean in the `worjh2` modern configuration.<sup>24</sup> In the *Cao et al.* [2009] configuration, these parameter settings were:

```
#diffusivity scaling factor
ea_diffa_scl=0.25
#grid point distance over which scalar is applied (j direction)
ea_diffa_len=3
```

---

<sup>22</sup>If you want exactly the same physics as per e.g. *Cao et al.* [2009], use:  
`cgenie.eb_go_gs_ac_bg.worjh2.BASES.config` as a template.

<sup>23</sup>The original parameter setting was:  
`gs_11=3573.718`

<sup>24</sup>The parameters for this were:  
`# diffusivity scaling factor`  
`ea_diffa_scl=0.25`  
`# grid point distance over which scalar is applied (j direction)`  
`ea_diffa_len=3`

but by default, reduced atmospheric diffusivity is disabled.

### 1b. Defining your 'world' in the *base-config*

Having chosen a *base-config* file to work with, you are now going to copy-paste in the block of parameters that defines the **cookiegen**-generated world.

In the template *base-config* files there is a highlighted (<<< >>>) line:

```
# ****
# GRID & BOUNDARY CONDITION CONFIGURATION
# ****
# insert the automatically generated cookiegen parameter list here
# ****
# <<<                                >>>
# ****
```

Copy and paste the contents of the **cookiegen** output file<sup>25</sup>: config\_yymmdd.txt into the template file where indicated – immediately above, immediately below, or simply replacing the entire <<< >>> line so the block of text becomes something like:

```
# ****
# GRID & BOUNDARY CONDITION CONFIGURATION
# ****
# insert the automatically generated cookiegen parameter list here
# ****
#####
### cGENIE .config file parameter lines generated by cookiegen v0.9.99 on: 251112 ###
# INPUT FILE PATH
ea_1='.../cgenie.cookie/genie-paleo/paleo___',
go_1='.../cgenie.cookie/genie-paleo/paleo___',
gs_1='.../cgenie.cookie/genie-paleo/paleo___',
# Grid resolution
...
...
# Ocean Ca, Mg, SO4 concentrations (modern defaults, mol kg-1)
bg_ocn_init_35=10.280E-03
bg_ocn_init_50=52.820E-03
bg_ocn_init_38=28.240E-03
#####
# ****
```

<sup>25</sup>In the filename, yymmdd is the date of the configuration creation.

### 1c. Defining tracers in the *base-config*

As per for defining the specific world, there is an explicit section of the template *base-config* that needs to be edited in order to increase the number of tracers represented in the model beyond T and S (in the ocean, and temperature and humidity in the atmosphere):

```
# ****
# TRACER CONFIGURATION
# ****
# the total number of tracers includes T and S
# T and S do not need to be explicated selected and initialzied
# ****
# Set number of tracers
GOLDSTEINTRACSOPTS='$(DEFINE)GOLDSTEINTRACS=2'
# list selected biogeochemical tracers
# <<<                                >>>
# list biogeochemical tracer initial values
# <<<                                >>>
# ****
```

The template *base-config* files come with just 2 ocean tracers defined – temperature and salinity (i.e., there is no carbon cycle or ocean nutrients *etc.* enabled at this point). These are implicit and essential to ocean circulation and hence are not listed. However, the count of total number of tracers in the ocean (which determines the compiled tracer and tracer-related array size) includes them, i.e.

```
# Set number of tracers
GOLDSTEINTRACSOPTS='$(DEFINE)GOLDSTEINTRACS=2'
```

So how to fill out the list of tracers? A complete list of all gaseous (in the atmosphere), dissolved (ocean only), and particulate (ocean and also sediments) tracers can be found in the files: *tracer\_define.atm*, *tracer\_define.ocn*, and *tracer\_define.sed*, respectively (in directory *genie-main/data/input*). However, starting from scratch and knowing which ones to add to create a consistent and sufficient set is not trivial. Instead, and as before, you could take a published configuration from one of the *genie-useroncigs/MS* subdirectories (refer to the *base-config* employed in the *README* file) or an EXAMPLE.

Either copy-paste the entire # TRACER CONFIGURATION section, or individually update the three subsections (or copy-paste the entire block, then edit):

1. # list selected biogeochemical tracers

First list (set equal to .true. all the gaseous (in the atmosphere), dissolved (ocean only), and particulate (ocean and also sediments) tracers.

2. # Set number of tracers

Count up how many ocean tracers there are, and add 2 for temperature and salinity (which you do not need to explicitly list), and update (the value of 2) in the line:

```
GOLDSTEINTRACSOPTS='$(DEFINE)GOLDSTEINTRACS=2'
```

3. # list biogeochemical tracer initial values

Then, if you wish for any of the gaseous and dissolved (there is no initialization option for particulate/solid tracers) not to be initialized at zero, list their initial values.

Units of gas partial pressure are *atmospheres (atm)*. Units of dissolved tracers in the ocean are *mol kg<sup>-1</sup>*.

Note that if you employ a *re-start*, then these values are over-written.

A series of template tracer lists, taken from published configurations, are also provided:

- TRACERCONFIG.ABIOTIC.txt

This is rather trivially ... the same as in the *base-config* templates and defines an ocean (and atmosphere) with no biogeochemical tracers, just T and S in the ocean.

You might use this if you were only interested in questions of ocean circulation and climate, although even then, you might want to add an age tracer to the ocean (see HOW-TO on diagnosing how the model works) which you would do by un-commenting the line that defines 3 total tracers (and comment the line defining just 2), and un-comment the line setting: gm\_ocn\_select\_48=true.

- TRACERCONFIG.Ridgwelletal.2007.txt

This is the set of tracers used in *Ridgwell et al.* [2007] and defines a  $PO_4$ -based biological pump, with  $^{13}C$  in all the carbon pools, and the capability for accounting for sulphate-reduction with dissolved  $O_2$  is depleted.

This is one of the simplest but most versatile tracer sets, particularly for paleo, when simulating  $^{13}C$  may be important. It is usable with any  $PO_4$ -only based biological scheme, including **ECOGEM** with *Fe* cell quotas disabled (*Wilson et al.* [2018]). Other publications include:

- *Ridgwell and Schmidt* [2010]
- *Ridgwell and Hargreaves* [2007]
- *Crichton et al.* [2020]

- TRACERCONFIG.Caoetal.2009.txt

This is the set of tracers used in *Cao et al.* [2009] and defines a  $PO_4$ -based biological pump, with  $^{13}C$  in all the carbon pools.

Unlike TRACERCONFIG.Ridgwelletal.2007.txt, there are no  $SO_4$  or  $H_2S$  tracers defined in the ocean and hence no potential for sulphate-reduction. Excess oxygen consumption then results in the creation and transport (and subsequent destruction) of negative concentrations of  $O_2$ . Kinetically (and in terms of large-scale patterns of oxygenation), this is very similar to creating and then re-oxidizing hydrogen sulphide – see *Meyer et al.* [2016].

This set of tracers also includes radiocarbon  $^{14}C$  in all the carbon pools, for diagnosing ocean circulation (and water mass ages). Additionally, both *CFC* – 11 and *CFC* – 12 are included for tracing deep-water formation. See *Cao et al.* [2009] for how these tracers are used.

- TRACERCONFIG.Odalenetal.2019.txt

This is the set of tracers used in *Odalen et al.* [2019] and defines a biological pump with both  $PO_4$  and *Fe* co-limitation. There are no  $SO_4$  or  $H_2S$  tracers defined in the ocean and hence no potential for sulphate-reduction, but  $^{13}C$  is included in all the carbon pools.

This set of tracers includes a variety of numerical/color tracers for diagnosing pre-formed properties<sup>26</sup> – here of *DIC*, *ALK*,  $O_2$ ,  $PO_4$ , and  $\delta^{13}C$ .

The iron cycle is the original configuration, where 3 separate tracers are included: *Fe* (free dissolved iron), *L* (free ligands), and *FeL* (iron bound to ligands). Because the equilibrium between these tracers is re-solved every time-step (and used to determine the *Fe* concentration for scavenging), there is redundancy, and in practice, only 2 tracers are needed. Subsequent iron cycle scheme hence use only 2 tracers (e.g. see: TRACERCONFIG.Wardetal.2018.txt). The 3-tracer scheme was also used in *Tagliabue et al.* [2016].

---

<sup>26</sup>See: 'diagnose how the model works' HOW-TO

- TRACERCONFIG.Wardetal.2018.txt

This is the set of tracers used in *Ward et al.* [2018] and defines a biological pump with both  $PO_4$  and  $Fe$  co-limitation of biological productivity (in ECOGEM).  $SO_4$  or  $H_2S$  tracers as well as  $^{13}C$  are included. There are no circulation diagnostic and/or numerical/color tracers. The configuration of the marine iron cycle differs from that of *Odalen et al.* [2019] and *Tagliabue et al.* [2016] in that only two tracers are now explicitly represented and transported in the ocean –  $TDFe$  (total dissolved iron, including iron bound to ligands) and  $TL$  – total dissolved ligand concentration (both free and bound to iron). At each time-step, the concentration of 'free' iron (and iron bound to ligands) is solved for and used to calculate iron removal through scavenging.

---

### 1d. Enabling science modules in the *base-config*

The final step in configuring your *base-config* file, is to ensure that all the science modules you wish to include are included (and those you do not want, are not). The template *base-config* files specify the basic/simply climate system combination of **GOLDSTEIN** ocean circulation model, **GOLDSTEIN** sea-ice model, and the **EMBM** atmospheric model. In addition, the atmospheric geochemistry module (**ATCHEM**) and the ocean biogeochemistry module (**BIOGEM**) are selected. These settings look like:

```
# ****
# GENIE COMPONENT SELECTION
# ****
# make .TRUE. the cGENIE modules to be included
# ****
ma_flag_ebatmos=.TRUE.
ma_flag_goldsteinocean=.TRUE.
ma_flag_goldsteinseaice=.TRUE.
ma_flag_biogem=.TRUE.
ma_flag_atchem=.TRUE.
ma_flag_sedgem=.FALSE.
ma_flag_rokgem=.FALSE.
ma_flag_gemlite=.FALSE.
ma_flag_ecogem=.FALSE.
# ****
```

Note that **ATCHEM** and **BIOGEM** have to be selected together, as do **SEDGEM** and **ROKGEM** (if you want an 'open system because **ROKGEM** provides the weathering input). **ECOGEM** is a replacement for the implicit biological export scheme in **BIOGEM** and hence requires **BIOGEM** to be selected (and hence also **ATCHEM**). If you want the **ECOGEM** marine ecosystem model, set `ma_flag_ecogem=.TRUE.`.

---

## 2. Choosing a template *user-config*

The final file creation step is for/of a *user-config* file. Once again – you could take a published configuration from one of the genie-userconfigs/MS sub-directories or an EXAMPLE, copy and rename it and then edit in any changes you need.

A series of template *user-config* files, some directly derived from published configurations (esp. for the MODERN *user-configs*), are also provided as part of the **cookiegen** release. Where alternative parameter choices exist in the *user-config*, these are highlighted and commented out (###).

Also in the MODERN template *user-configs*, are some suggestions to parameter changes that align the MODERN *user-configs* with the corresponding PALEO *user-configs* (where a correspondence exists). These suggested alternative parameter choices are also included commented out (###) and are:

- Under \*\*\* REMINERALIZATION \*\*\*:

```
#### set 'instantaneous' water column remineralization
###bg_par_bio_remin_sinkingrate_physical=9.9E9
###bg_par_bio_remin_sinkingrate_reaction=125.0
```

which enables instantaneous water-column remineralization of POM – see USERCONFIG.PALEO.BIOGEM.PO4.SPIN for a full description.

- Under \*\*\* MISC \*\*\*:

```
#### maximum time-scale to geochemical reaction completion (days)
###bg_par_bio_geochem_tau=90.0
#### extend solubility and geochem constant T range (leave S range as default)
###gm_par_geochem_Tmin = -2.0
###gm_par_geochem_Tmax = 45.0
###gm_par_carbchem_Tmin = -2.0
###gm_par_carbchem_Tmax = 45.0
```

The first of these imposes a maximum reactant uptake time-scale. The second extends the temperature limits imposed on gas solubility and carbonate chemistry dissociation parameters. See: USERCONFIG.PALEO.BIOGEM.PO4.SPIN for a full description.

In all the PALEO templates, a series of recommendations are made that differ from some of the defaults in the MODERN template *user-configs*. These include changes to:

1. biological scheme
2. inorganic matter export ratios
3. water column remineralization
4. reactant consumption limitation
5. extended temperature ranges of solubility and geochemistry constants

These are all described in detail for USERCONFIG.PALEO.BIOGEM.PO4.SPIN.

As mentioned above – the same/similar settings also appear in the MODERN templates (but commented out) should one which to align all the model assumptions across model and paleo applications.

The provided template *user-config* files are:

- USERCONFIG.ABIOTIC.TRACER.SPIN

This is as very basic template for an ocean (and atmosphere) with no carbon cycle.

You might use or adapt this if you were only interested in questions of ocean circulation and climate, although even then, you might want to add an age tracer to the ocean<sup>27</sup> (see HOW-TO

---

<sup>27</sup>If so, un-comment the 'optional' parameter setting.

on diagnosing how the model works).

While the template *user-config* could be used with any degree of complexity of tracers, it is designed for use with TRACERCONFIG.ABIOCIC.txt, with or without the 3rd tracer in the *base-config* needed for diagnosing ventilation age: gm\_ocn\_select\_48=.true.

- USERCONFIG.MODERN.BIOGEM.PO4.SPIN

This provides the *user-config* parameter settings for the pre-industrial *spin-up* experiment described in *Cao et al.* [2009], as well as, commented-out, the settings for the configuration of *Ridgwell et al.* [2007]. (The differences in parameters primarily relate to the two difference vertical resolutions and the calibrations performed on them.) Refer to the EXAMPLES for e.g. the parameter changes needed to the *Cao et al.* [2009] historical transient experiment.

Suitable template tracer sets include: TRACERCONFIG.Ridgwelletal.2007.txt and TRACERCONFIG.Caoetal.2009.txt, depending on whether you intend to make use of the additional diagnostic ocean circulation tracers or not.

- USERCONFIG.MODERN.BIOGEM.PO4Fe.SPIN

This provides the *user-config* parameter settings for a modern (pre-industrial) marine iron cycle (biological productivity limited by both *PO<sub>4</sub>* and *Fe*). It is based on a configuration used by *Odalen et al.* [2019] as well as in *Tagliabue et al.* [2016] and was calibrated for the worjh2 world configuration. An alternative set of biological uptake and *Fe* cycle parameters are provided (commented out) which are also in use [unpublished work] and was calibrated for the worlg4 world configuration.

Suitable template tracer sets include: TRACERCONFIG.Odalenetal.2019.txt and TRACERCONFIG.Wardetal.2018.txt, depending on the *Fe* tracer scheme selected.

- USERCONFIG.MODERN.ECOGEM.PO4Fe.SPIN

This is basically, the *user-config* for the modern **ECOGEM** experiment following *Ward et al.* [2018] and bar some comments and tidying up is the same as wardetal.2018.ECOGEM.SPIN that can be found in genie-userconfigs/MS/wardetal.2018.

The only difference are a couple of optional suggestions that will align the configuration with the corresponding recommended paleo configurations (in addition to the POM remineralization and geochemical constant temperature range suggestions listed earlier):

- Under \*\*\* REMINERALIZATION \*\*\*:

```
#### DOC lifetime (yrs) -- following Doney et al. [2006]
###bg_par_bio_remin_DOMlifetime=0.5
```

which aligns the lifetime of DOM with the *Doney et al.* [2006] value adopted in the original **BIOGEM** configuration.

- Under \*\*\* MISC \*\*\*:

```
#### add seaice attenuation of PAR
###eg_ctrl_PARseaicelimit=.true.
#### relative partitioning of C into DOM
###eg_par_beta_POCtoDOC=0.70
```

which firstly limits light available under sea-ice in proportion to the fractional sea-ice cover in that grid cell, and secondly, re-partitions carbon from POM to DOM and is tuned to produce an approximately Redfield ratio (106) of 104.7 : 1 in *C* : *P* of exported POM.

The intended corresponding template tracer is: TRACERCONFIG.Wardetal.2018.txt.

- USERCONFIG.PALEO.BIOGEM.PO4.SPIN

This is a recommended template *user-config* for paleo applications. Several parameters have been changed or added compared to published paleo configurations (these recommendations need not be adopted and alternatives are given commented out in the file):

- 1. biological scheme**

```
# biological scheme ID string
bg_par_bio_prodopt="bio_P"
# biological uptake time-scale
bg_par_bio_tau=63.3827
# [PO4] M-M half-sat value (mol kg-1)
bg_par_bio_c0_PO4=0.10E-06
```

This differs from most of the published paleo applications in that it has a biological export scheme that is temperature-dependent and more responsive to changes in ocean  $PO_4$  inventory (a similar configuration was used by *Meyer et al.* [2016]).

- 2. inorganic matter export ratios**

While it would be perfectly acceptable to utilize the default carbonate saturation-dependent scheme for the ratio of exported  $CaCO_3$  compared to  $POC$  (see Ridgwell et al. [2007, 2009]), published paleo implementations of **cookie** have tended to utilize a simple spatially uniform and invariant with time,  $CaCO_3 : POC$  ratio. This is enacted via:

```
# fixed CaCO3:POC
bg_opt_bio_CaCO3toPOCrainratio='FIXED',
# underlying export CaCO3 as a proportion of particulate organic matter
bg_par_bio_red_POC_CaCO3=0.200
```

with the first parameter specifying a fixed and invariant ratio, and the second parameter specifying what the ratio actually is. As an alternative to the first parameter, one can set the power in the carbonate saturation parameterization, to zero, i.e.:

```
# exponent for modifier of CaCO3:POC export ratio
bg_par_bio_red_POC_CaCO3_pP=0.0
```

The default/recommended paleo value of 0.2 derives from the model-data analysis of *Panchuk et al.* [2008], and as such is both subject to the specific caveats of that study, and may not necessarily be appropriate for later in the Cenozoic (or earlier in the Mesozoic). For deeper time – prior to ca. the mid Mesozoic – the recommended parameter value is 0.0, on the basis that pelagic calcifiers had not yet evolved and the surface ocean export of biogenic  $CaCO_3$  is effectively zero – e.g. see *Ridgwell* [2005].

- 3. water column remineralization** In the default and previously published (in all applications) remineralization scheme, particulate matter only sinks a finite distance each time-step. At the next time-step, it starts from the ocean layer it reaches and sinks further (determined by the default sinking rate of  $125\text{ m d}^{-1}$ )<sup>28</sup>. This can make it difficult to e.g. check flux mass balances because there is a lag between export and particulate matter reaching the sediment surface.

Alternatively, a ‘cleaner’ approach is to instantaneously remineralize all particulate organic matter throughout the water column according to the remineralization profile and/or reaction rates. This is activated via `bg_par_bio_remin_sinkingrate_physical` which is simply assigned a very ‘large’<sup>29</sup> value for the sinking rate. At the same

<sup>28</sup>It always has to travel at least 1 ocean layer downwards on each time-step – more for faster sinking rates and/or thinner layers nearer the surface.)

<sup>29</sup>Sufficient that the deepest ocean model layer can be reached within a single time-step.

time, reaction rates (including scavenging) are calculated as if the sinking rate was finite and equivalent to the value of `bg_par_bio_remin_sinkingrate_reaction` ( $md^{-1}$ ). Hence:

```
# set 'instantaneous' water column remineralization
bg_par_bio_remin_sinkingrate_physical=9.9E9
bg_par_bio_remin_sinkingrate_reaction=125.0
```

#### 4. reactant consumption limitation

It is good practice not to allow complete consumption of any particular reactant by a single reaction. This is because multiple processes may be competing for the same reactant, whilst at the same time, ocean circulation is transporting the reactant away. The result can be (small) negative tracer values.

```
# maximum time-scale to geochemical reaction completion (days)
bg_par_bio_geochem_tau=90.0
```

By setting the value of `bg_par_bio_geochem_tau`<sup>30</sup>, the maximum consumption of any particular reactant by a single reaction is limited by an imposed lifetime (days). Here, the value is set to 90 d.

#### 5. extended temperature ranges of solubility and geochemistry constants

As described in *Ödalen et al.* [2018]: "The dissociation constants used in the cGENIE calculations of solubility for  $CO_2$  in seawater follow *Mehrbach et al.* [1973], which are only defined for waters between 2 and 35°C. Hence, the expression for  $CO_2$  solubility in the model is restricted so that all water below 2°C has the same  $CO_2$  solubility (similarly for all water above 35°C)."

The recommended parameter changes:

```
# extend solubility and geochem constant T range (leave S range as default)
gm_par_geochem_Tmin = -2.0
gm_par_geochem_Tmax = 45.0
gm_par_carbchem_Tmin = -2.0
gm_par_carbchem_Tmax = 45.0
```

extend the valid range to  $-2 - 45^{\circ}C$ . Note that the valid salinity range is left unchanged.

Finally, it should be noted that the parameter settings for activating temperature-dependent remineralization (as per *John et al.* [2014] and *Crichton et al.* [2020]) are given (and commented out (###)):

```
# *** Crichton et al. [2020] temperature-dependent remin ****
###bg_ctrl_bio_remin_POC_fixed=.false.
###bg_par_bio_remin_POC_K1=9.0E11
###bg_par_bio_remin_POC_Ea1=54000.0
###bg_par_bio_remin_POC_K2=1.0E14
###bg_par_bio_remin_POC_Ea2=80000.0
###bg_par_bio_remin_POC_frac2=0.008
```

Note that in *Crichton et al.* [2020], a different biological scheme is used, with a faster time-scale of nutrient uptake (but different light limitation) – also given commented out in the *user-config*<sup>31</sup>.

- USERCONFIG.PALEO.BIOGEM.PO4Fe.SPIN

This is very similar to USERCONFIG.PALEO.BIOGEM.PO4.SPIN, with the exception that it adds iron co-limitation of biological productivity and a marine iron cycle to the  $PO_4$ -only paleo settings, including the recommend changes to sinking and water column remineralziation,

<sup>30</sup>By default it is set to zero which is interpreted as allowing complete reactant consumption as per previously.

<sup>31</sup>There is no strict necessity to use the alternative biological export scheme.

reaction completion time-scales, and gas solubility.

A simplified 2-tracer ( $TDFe$ ,  $TL$ ) iron system is assumed:

```
# iron tracer scheme
# NOTE: the base-config requires TFe and TL tracers
bg_opt_chem_Fe='hybrid'
```

while the solubility and iron scavenging rates come from the `worjh2` calibrated configuration (see: `USERCONFIG.MODERN.BIOGEM.PO4Fe.SPIN`).

The only particularly novel addition, is that of a fake and globally uniform dust field in order to supply dissolved iron to the ocean surface:

```
bg_par_forcing_name="pyyyyy.RpC02_Rp13C02.DUST"
```

This is configured to give the same total dust supply to the ocean surface as per in the dust field used in *Ward et al.* [2018], but is supplied evenly across the global ocean surface.

Note that there is no (*user-config*) scaling parameter for sediment flux forcings and hence the total flux forcing appear explicitly in the time-series file: `biogem_force_flux_sed_det_sig.dat`. Also note that for a different total ocean area, the flux per  $m^{-2}$  will change as only the global total flux to the ocean surface is conserved.

- `USERCONFIG.PALEO.ECOGEM.PO4.SPIN`

This is based on the configurations published by *Wilson et al.* [2019], with iron quotas and usage disabled in **ECOGEM**.

In its modern tuning, **ECOGEM** has a high carbon export and high  $C/P$  ratio. This leads to a less well oxygenated modern ocean than observations. In removing iron and configuring  $PO_4$  as the sole limiting nutrient for paleo experiments where the dust flux field is not *a priori* known, the  $C/P$  ratio increases still further and with it, a further depletion of ocean oxygen. This situation is rectified for paleo applications (under 'recommended') by:

1. **diagnosing, not applying a mixed layer**

```
# set mixed layer to be only diagnosed (for ECOGEM)
go_ctrl_diagmld=.true.
```

and

2. **re-partitioning carbon from POM to DOM**

(leaving nutrients etc. unchanged):

```
# relative partitioning of C into DOM
eg_par_beta_POCtoDOC=0.70
```

The alternative calibrated value is given should  $Fe$  limitation is included. For dust and aeolian iron supply, simply follow the instructions and parameter values for `USERCONFIG.PALEO.BIOGEM.PO4Fe.SPIN`.

In addition:

3. **sea-ice light limitation** The original **ECOGEM** model did not include any limitation of biological productivity by sea-ice cover. This is included now by:

```
# add seaice attenuation of PAR
eg_ctrl_PARseaicelimit=.true.
```

which reduces light proportionally to the local fractional sea-ice cover.

Additional recommend changes for sinking and water column remineralization, reaction completion time-scales, and gas solubility, are the same as per for the **BIOGEM PALEO** template *user-configs*, described previously.

### 3. Ensuring files are in their correct location and parameter settings are consistent

1. First, ensure that you have remembered to copy/transfer the entire configuration file sub-directory created by **cookiegen**<sup>32</sup> to: cgenie.cookie/genie-paleo.
2. Next, your new *base-config* will need to be transferred to: genie-baseconfigs.
3. Your *user-config* can go 'anywhere' ... ish. Putting it directly in genie-userconfigs<sup>33</sup> may eventually make things over-crowded. Better is to create a sub-directory of genie-userconfigs (or even a sub-directory of this) and place it there<sup>34</sup> or simply in: genie-userconfigs/LABS
4. Checking the consistency between the tracers defined in the *base-config* and the experiment parameter requested in the *user-config* is the trickiest part. Refer to some of the published configurations (best) and/or EXAMPLES for guidance.

Remember that if you don't want any additional tracers, the tracer total line in the *base-config* should be:

```
GOLDSTEINNTRACSOPTS='$(DEFINE)GOLDSTEINNTRACS=2'
```

If you want a single color (age) tracer, you need:

```
GOLDSTEINNTRACSOPTS='$(DEFINE)GOLDSTEINNTRACS=3'
```

and the additional line:

```
gm_ocn_select_48=.true. # colr - 'RED numerical (color) tracer'
```

Sometimes it is best simply to start the experiment running and see what happens – there are checks (not comprehensive) in the code to see if you have the necessary tracers selected. Read any warnings that are reported as the experiment initializes, but often you can ignore these.

Ensure that the total number of defined tracers equals the number of selected tracers you have listed (plus temperature and salinity) in the *base-config*.

5. Finally, if you have previously run an experiment using a different *base-config*, you will need to briefly interactively run an experiment with the new *base-config* in order that **cookie** knows to re-compile. (Remember that you cannot go directly to submitting jobs to the cluster when changing *base-configs*.)

Good luck!

---

<sup>32</sup>Which by default will appear in a directory: OUTPUT

<sup>33</sup>The directory parameter passed to runcookie.sh is then just /

<sup>34</sup>If you place the *user-config* file into a sub-directory myexperiments, the directory parameter passed to runcookie.sh is then myexperiments

## 17.7 cookiegen parameter settings – details

Compared to **muffingen**, **cookiegen** adopts a somewhat simplified set of example configuration files and some of the parameters that you might want to adjust, and not explicitly listed in the revised files. The following Sections describe several different ways of adjusting the generated continental configurations including use of some of the 'hidden' parameters.

### Zonal wind-stress

For non GCM-based configurations, no prior wind fields exist. **cookiegen** hence creates and configures an idealized zonal wind-stress field (from which wind velocity and wind speed is derived). The zonal wind-stress can take alternative strengths, depending on whether a high latitude gateway (in either hemisphere) exists. This can be prescribed directly, or **cookiegen** can be enabled to 'choose' whether or not a high latitude gateway exists and hence whether or not to apply a strong or weak zonal flow. The parameter options (in the **cookiegen** configuration file) are:

- `par_tauopt=0;`  
**cookiegen** chooses whether or not to apply a strong or weak zonal flow, and in which hemisphere, depending on whether it thinks a high-latitude gateway exists or not.
- `par_tauopt=1;`  
A weak zonal flow is applied in both hemispheres (e.g. assuming land prevents the existence of an ocean-only circumpolar high latitude pathway in both hemispheres). The presence of a pole-to-pole super-continent (or 'ridge-world') would be an example of when this option might be selected<sup>35</sup>.
- `par_tauopt=2;`  
A strong zonal flow is applied in both hemispheres (e.g. assuming there is no land to prevent the existence of an ocean-only circumpolar high latitude pathway in both hemispheres). The presence of an Equatorial-only super-continent (or 'water-world') would be an example of when this option might be selected<sup>36</sup>.
- `par_tauopt=3;`  
'grey' world – an intermediate strength zonal flow is applied in both hemispheres. For hedging your bets of for use across a wide range of different continental configuration where you do not want to manually chose and prescribe either strong (2) or weak (1) and don't dare trust **cookiegen** to make the choice for you (option 0).

For reference – the modern world has a mix of strong (southern) and weak (northern) hemisphere zonal flows.

Note that when deriving a **cookie** configuration from GCM output, selecting a zonal wind profile will result in the generation of an idealized profile rather than calculate the zonal average from the GCM fields. (This is different from **cookiegen** behavior for planetary albedo, where selecting the `opt_makezonalalbedo` option will generate an idealized albedo field, whereas otherwise, a zonal mean albedo field is still generated and used by default, but with the values derived from a 2D re-gridded GCM field.)

<sup>35</sup>But note that in the ridgeworld EXAMPLE provided, the automatic assignment option (#0) is instead set to be consistent with the series of fake .dat based configurations.)

<sup>36</sup>But note that in the waterworld EXAMPLE provided, the automatic assignment option (#0) is instead set to be consistent with the series of fake .dat based configurations.)

### Ocean depth (and maximum levels)

The standard configurations of cookie assume a maximum ocean depth of 5000 m. (In fact, until recent code changes, the scale-depth in the **GOLDSTEIN** ocean model component was hard-coded as 5000.0 m.) This depth is then divided up into 16 (typically) or 8 ocean levels with a logarithmic distribution of ocean layer thicknesses.

The number of layers in the ocean is set by the parameter:

`par_max_k`

and the maximum (scale) depth is:

`par_max_D`

You need not have all `par_max_k` layers spread over depth `par_max_D`. You can specify the minimum (deepest) depth level used in the cookie depth grid via the parameter:

`par_min_k`

By default this has a value of 1. Setting a high value truncates the ocean floor. e.g. for a flat-bottom ocean, a common value is 3, meaning that the ocean model consists only of layers 16 down through 3, giving it a maximum depth of about 3500 m. This, in conjunction with a modern-like land fraction, gives a modern-like ocean volume (given that the modern ocean has an average seafloor depth of ca. 3500 m).

If you require a deeper (maximum depth) ocean than 5000 m, be aware that retaining the same (e.g. 16) value of `par_max_k` means that the same number of ocean layers are distributed over the greater ocean depth, and hence, your surface layer will be deeper, potentially changing biological productivity and biogeochemical cycling. There is hence a facility provided for maintaining the thickness depth distribution in the upper ocean while adding additional layers to the bottom of the ocean to create a greater (than 5000.0) maximum ocean depth. This is implemented by setting the parameter:

`par_add_Dk`

(the additional number of depth (k) levels) to a non-zero value. The effect of this is to assume that the maximum ocean depth parameter (`par_max_D`) corresponds to the specified maximum number of ocean layers (`par_max_k`) minus the number of additional layers (`par_add_Dk`)<sup>37</sup>. The maximum ocean depth is then re-calculated based on `par_max_k` minus `par_add_Dk` number of levels, corresponding to depth `par_max_D` (with the same logarithmic distribution of layer thicknesses with ocean depth applied to the additional layers deeper than `par_max_D`).

For example, setting:

`par_add_Dk=2`

in conjunction with a total number of ocean levels:

`par_max_k=18`

and:

`par_max_D=5000.0`

together configures an 18-level ocean model with the first 16 (18-2=16) levels spanning 5000.0 m, with an automatically re-calculated maximum ocean depth (`par_max_D` value) of 6922.2705 m and the additional 2 levels spanning this additional depth (1922.2705 m).<sup>38</sup>

---

<sup>37</sup>(rather than the maximum number of ocean levels as usually applied)

<sup>38</sup>Note that there is no way, unless the logarithmic function for the distribution of layer thicknesses with ocean depth is

When **cookiegen** writes out the parameter settings (file: config\_?????.txt), the scale depth of the ocean is also written out. By default this will be equal to the value of `par_max_D` (5000.0 m), unless, the value of `par_add_Dk` is non-zero, when the above described re-calculation of the maximum ocean depth is made (and written out in the parameter file).

---

---

changed, to retain the same modern/default layer thickness distribution for the upper ocean layers and have a maximum ocean depth that is different to 5000.0 m unless you carry out the calculation yourself for the depth occupied by e.g. 18 levels. (You cannot calculate how many layers occupies e.g. 6000.0 m and still retain the upper ocean depth distribution as the depth profile does not fit neatly into 6000.0 m.)

## 17.8 Appendix – pre-generated Worlds

A series of conceptual ('fake') worlds have been generated and are provided for your convenience :o) These fall into 2 categories:

### 1. Conceptual worlds

These include 'water world', 'ridge world', 'drake world', etc. and associated variants, e.g. as published in fully coupled GCMs<sup>39</sup>. They are designed to test controls on large-scale ocean circulation and do conform to an Earth-appropriate ocean volume or total cratonic area.

### 2. Idealized-continent worlds

These involve idealized super-continental configurations but aim to retain modern ocean volume and total cratonic area characteristics.

Both genie-paleo configuration file sets (sub-directories), and example *base-configs* are provided as part of the cookie code base. The details of these two series of fake worlds are as follows. Remember that the .m file configurations that create each of the fake world configurations are saved in the respective genie-paleo sub-directory, and hence both the specific details of how the world was created are available, as well as the ability to exactly recreate the world.

### 17.8.1 Conceptual worlds

- [TO BE ADDED]

### 17.8.2 Idealized-continent worlds

The idealized continents come both without (i.e. the configuration is either land, or deep ocean) or with, continental shelves, and in the latter case, come with varying number of 'steps' down to deeper water (and hence to the abyssal plain).

The fake world series 'k1' configuration names start with one of two, 4 character strings:

- fkh\_ == 'high' resolution ( $36 \times 36 \times 16$ )
- fkl\_ == 'low' resolution ( $18 \times 18 \times 16$ )

(No low vertical resolution, e.g. 8 level version is currently provided.)

The next 2 characters delineate the shape/position of the world:

- 1e == a single Equatorial-centered super-continent
- np == a single North polar-centered super-continent
- pp == a pole-to-pole super-continent

The 7th character is the 'series' represents:

- 0 == An earlier generation of worlds, with shelves stepping down in increments of only 1 ocean level.

Note that in generating the wind-stress in this series, the presence/absence of high latitude gateways is automatically identified (meaning that the zonal wind-stress profiles differ between np and pp in the magnitude of wind-stress in the Southern Hemisphere) – par\_tauopt=0.

---

<sup>39</sup>Example of aquaplanets and drake worlds in the MITgcm with simplified atmosphere:  
Rose, B. E. J., Ferreira, D. & Marshall, J. The role of oceans and sea ice in abrupt transitions between multiple climate states. *J. Clim.* 26, 2862–2879 (2013).  
Ferreira, D., Marshall, J. & Rose, B. Climate determinism revisited: Multiple equilibria in a complex climate model. *J. Clim.* 24, 992–1012 (2011).

- 1 == A subsequent generation, with shelves stepping down in increments of 2 ocean levels.  
Note: An intermediate wind-stress magnitude is applied to both hemispheres in all configurations (i.e. no distinction is made between hemisphere with or without high latitude gateways) – `par_tauopt=3` .
- 2 == A modification of #1, generated with a newer version of **cookiegen**, and with an imposed limit of 1000 – 5000m for the range of **SEDGEM** open ocean depths (the previous version included depths extending to the shallow sub-surface).  
Note that `par_tauopt=0` is used.  
Also note that in pp, shelves are only created on the Eastern margin of the basin so as to retain consistency in total shelf area with np.

The 8th character (digit) specifies the number of shelves, which for the initial series is:

0. (none)
1. level 15
2. level 15 + 14
3. level 15 + 14 + 13
4. level 15 + 14 + 13 + 12
5. level 15 + 14 + 13 + 12 + 11

and for the newer series:

0. (none)
1. level 15
2. level 15 + 13
3. level 15 + 13 + 11
4. level 15 + 13 + 11 + 9

All the oceans are ca. 3500m deep, i.e. truncated at ocean level 3, in order to produce approximately the present-day ocean volume.<sup>40</sup>

Example *base-configs* are provided for some, but not all of the idealized worlds, and follow the same naming convention and have full filenames of the form: cookie.CB.\*\*\*\*\*.BASES.config, where \*\*\*\*\* is the 8-character fake world 'k1' configuration name (see above). A basic set of tracers is provided (BASES).

Other than the rate at which the shelves step down<sup>41</sup>, the main difference between the different generations of worlds, is in the applied zonal wind field strength. In the earlier fake world series, the 'zonal wind-stress generation option' (`par_tauopt`) was 0. This attempts to automatically determine the presence of high latitude ocean gateways and if they exist, apply a stronger zonal wind field in that hemisphere (otherwise, a wind stress appropriate for a modern northern-hemisphere configuration with no zonal gateway is applied). To reduce the number of boundary conditions that change and co-vary between different fake worlds, in the second generation of idealized continent worlds, an intermediate strength zonal wind field is applied to all, regardless of the presence of absence of high latitude circumpolar ocean gateways. This is 'zonal wind-stress generation option' 3. Note that for the same resolution, wind-fields can be substituted/replaced, and using the saved .m configuration file, worlds can be re-generated with different wind-stress options if desired. The depth range of the ocean floor in **SEDGEM** (which informs the pressure used in the *CaCO<sub>3</sub>* stability calculation) is slightly restricted in #3 (the upper limit increased to 1000m).

---

<sup>40</sup>To change this, simply search-and-replace the 3 in the .k1 file with e.g. 1.

<sup>41</sup>(and that the earlier generation was for a pole-to-pole super-continent only)

**17.8.3****Modern worlds**

To both help illustrate how the configuration files are organized and accessed for paleo worlds, and to help make modifying older (modern) configurations easier, the 2 basic (published) modern configurations are provided in a 'genie-paleo' format (i.e. the format generated by **cookiegen**):

- p\_worbe2 == The 8-level  $36 \times 36$  ocean configuration of *Ridgwell et al.* [2007] and including the geological carbon cycle (**SEDGEM+ROKGEM**) configuration files of *Ridgwell and Hargreaves* [2007] (also: *Lord et al.* [2015]).
- p\_worjh2 == The 16-level  $36 \times 36$  ocean configuration of *Cao et al.* [2009] and including the geological carbon cycle (**SEDGEM+ROKGEM**) configuration files of *Archer et al.* [2009] (also: *Winkelmann et al.* [2015]).

Corresponding example *base-* and *user-config* files are provided:

- cookie.CBSR.p\_worbe2.BASES.config plus: cookie.CBSR.p\_worbe2.BASES.SPIN1  
(in: genie-userconfig/EXAMPLES), which runs the first-stage spin-up described in *Ridgwell and Hargreaves* [2007] (and also used in *Lord et al.* [2015]).
- cookie.CBSR.p\_worjh2.BASES.config plus: cookie.CBSR.p\_worjh2.BASES.SPIN1  
(in: genie-userconfig/EXAMPLES), which runs the first-stage spin-up used in *Archer et al.* [2009] and *Winkelmann et al.* [2015].

as well as some reduced tracer and biogeochemical complexity variants, such as cookie.CB ... and cookie.C ...

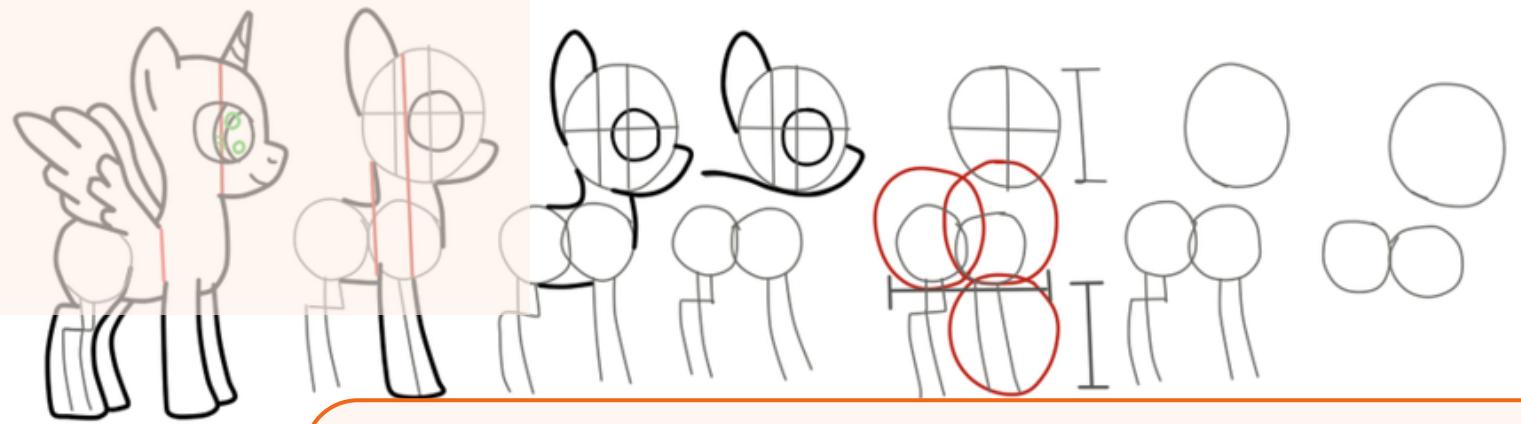
To make modified (e.g. wind fields) versions of either of the 2 modern 'paleo' format configurations:

1. Copy and rename the (8- or 16-level ocean configuration) genie-paleo sub-directory, giving it an 8 (or 6) character string name.
2. In this new sub-directory, edit the files (wind stress, wind velocity, wind speed, albedo, and/or topography) you want to modify. (These files do not need to be renamed.)
3. Copy and rename the corresponding *base-config* file (disabling the parameter settings for **SEDGEM** and **ROKGEM**, and/or enabling **ECOGEM**, as required).  
Then, in the *base-config* file, edit all instances of the genie-paleo subdirectory parameters to match the new genie-paleo sub-directory you assigned in step (1). (Parameters, e.g.: ea\_1='../../cgenie.cookie/genie-paleo/p\_worbe2'.
4. Copy-rename the example *user-config*, or modify your own existing one.  
Note that when employing the geological carbon cycle, the resolution of the sediment (and weathering) grids and associated boundary condition files are now defined in the *base-config* rather than the *user-config* (as originally done).



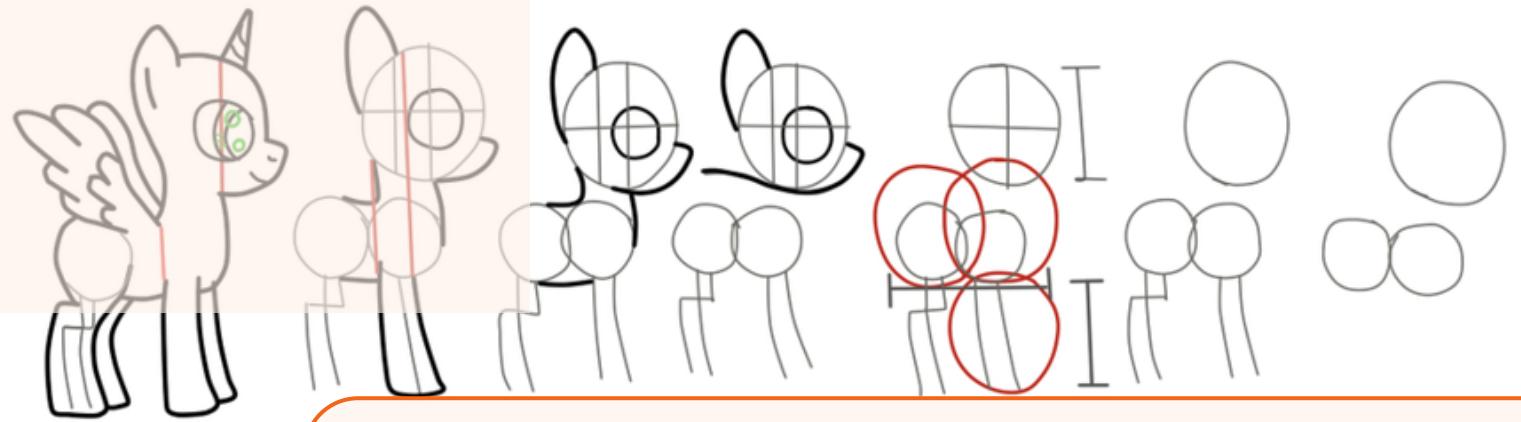
## 18. FAQ

**18.1 (none)**



## 19. HOW-TO (technical)

**19.1 (none)**



## 20. HOW-TO (experiment)

**20.1 (none)**

(none)

## 21. cookie Development

**21.1 (none)**