

The Bumper Book of *.cookies*

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

Andy Ridgwell

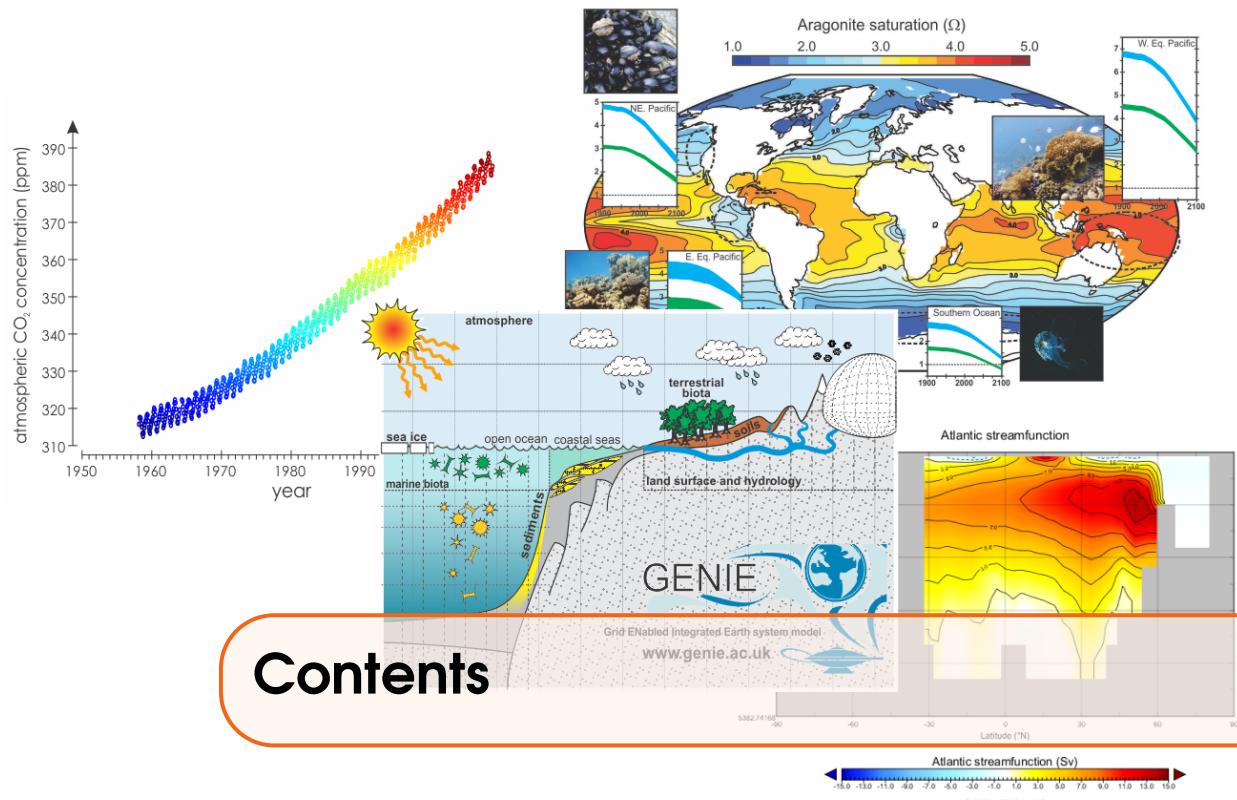
Copyright © 2017 Andy Ridgwell

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

0	Installation, Configuration, Basic Usage	1
0.1	(none)	2
1	Climate Dynamics & Experimental Design	3
1.1	(none)	4
2	Ocean circulation I – modern	5
2.1	(none)	6
3	Ocean circulation II (fake worlds)	7
3.1	(none)	8
4	Atmospheric dynamics	9
4.1	(none)	10
5	Fossil fuel CO₂ and 'ocean acidification'	11
5.1	(none)	12
6	Ocean biogeochemical cycles	13
6.1	(none)	14

7	Marine ecosystems and dynamics	15
7.1	(none)	16
8	The geological cycle of carbon	17
8.1	(none)	18
9	The land surface and carbon-climate feedbacks	19
9.1	(none)	20
10	Climates of Past Worlds	21
10.1	(none)	22
11	Proxies and Reconstructing the Past	23
11.1	(none)	24
12	Synthesis: carbon in the Earth system	25
12.1	(none)	26
13	cookie model output	27
13.1	(none)	28
14	Introduction to model results analysis	29
14.1	(none)	30
15	muffinplot	31
15.1	(none)	32
16	muffindata	33
16.1	(none)	34
17	cookiegen	35
17.1	Introduction	36
17.1.1	cookiegen	36
17.1.2	Installing cookiegen	36
17.2	Quick Start Guide	37
17.3	Running cookiegen	40
17.4	Configuring cookiegen	41

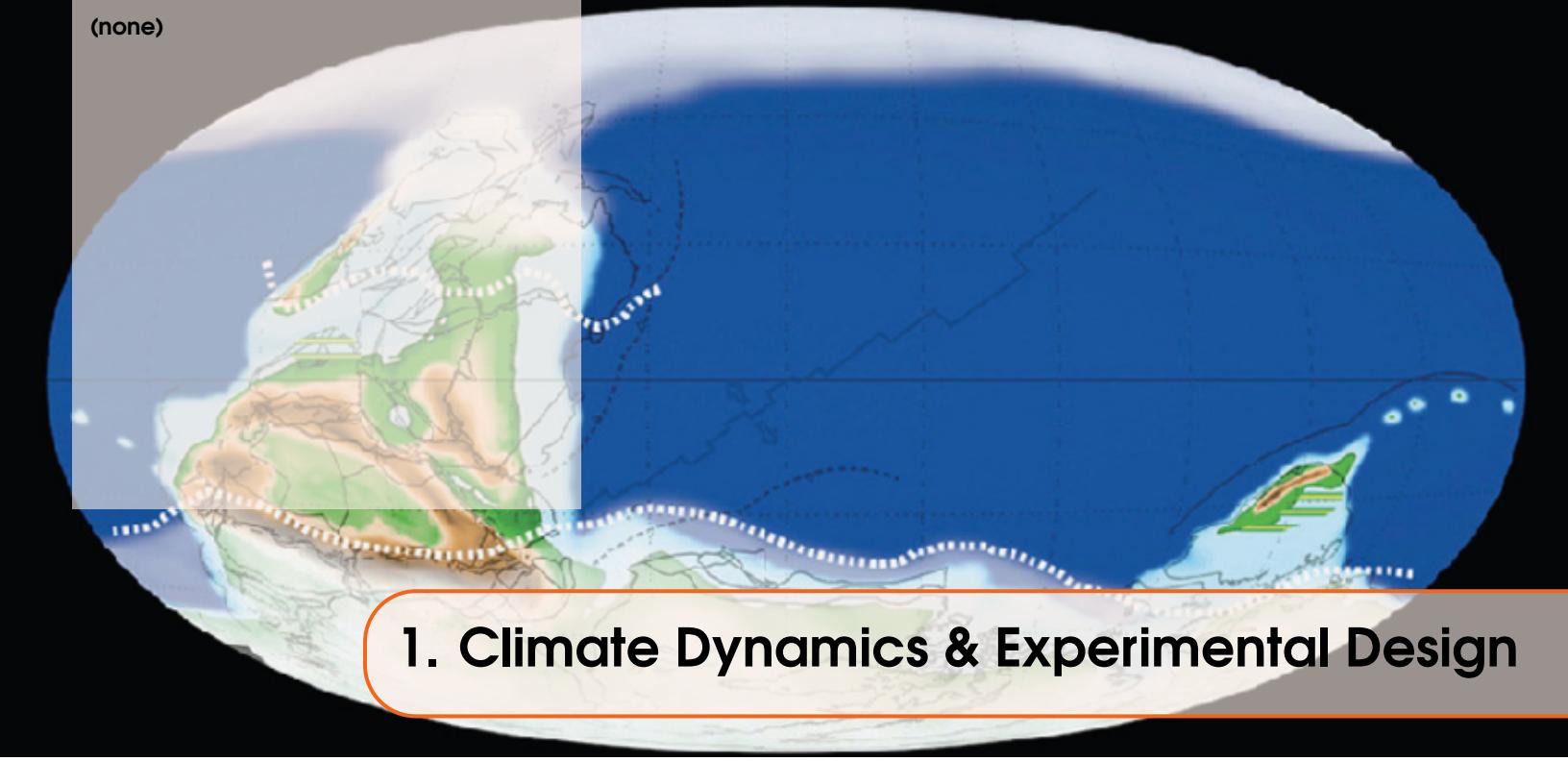
17.5	Overview	41
17.5.1	cookiegen configuration examples	44
17.6	Creating base- and user-configs for cookie	45
17.7	cookiegen parameter settings – details	59
17.8	Appendix – pre-generated Worlds	62
17.8.1	Conceptual worlds	62
17.8.2	Idealized-continent worlds	62
17.8.3	Modern worlds	64
18	FAQ	65
18.1	(none)	66
19	HOW-TO (technical)	67
19.1	(none)	68
20	HOW-TO (experiment)	69
20.1	(none)	70
21	cookie Development	71
21.1	(none)	72

(none)



0. Installation, Configuration, Basic Usage

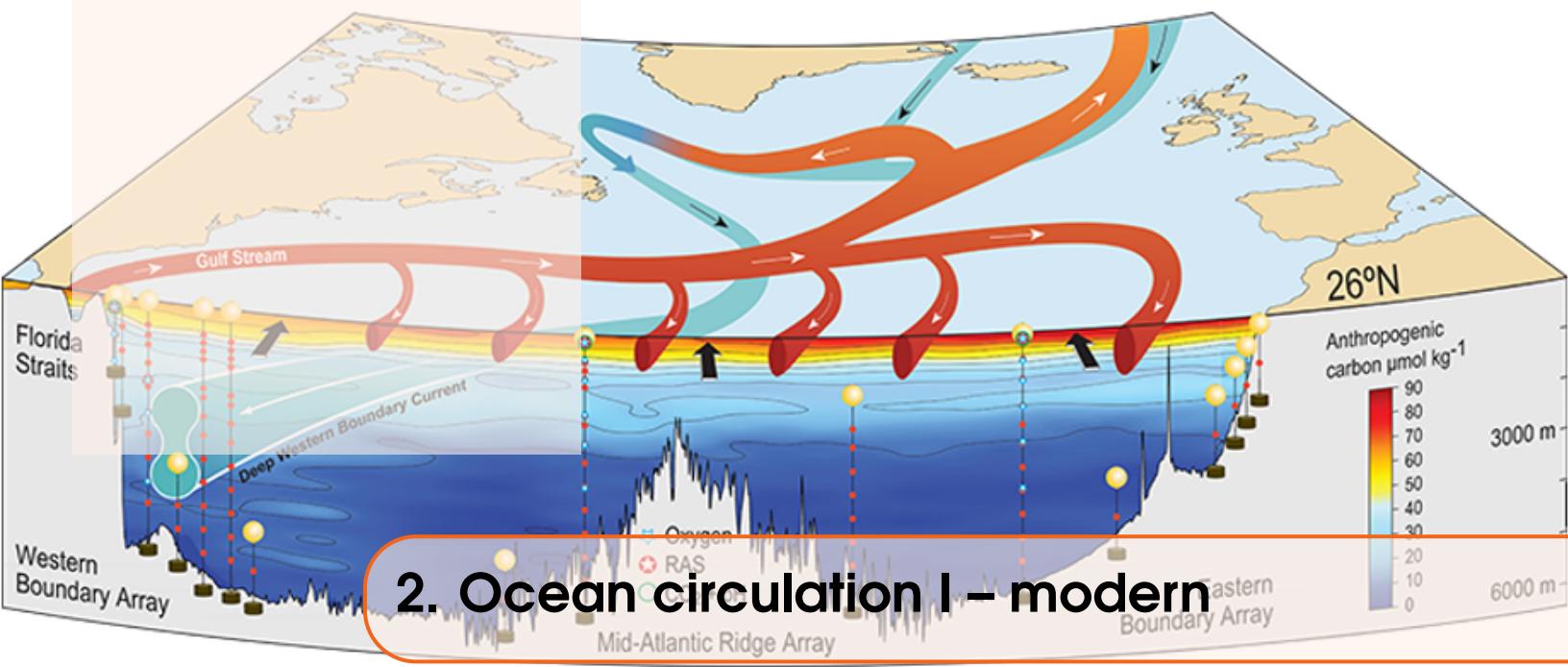
0.1 **(none)**



1. Climate Dynamics & Experimental Design

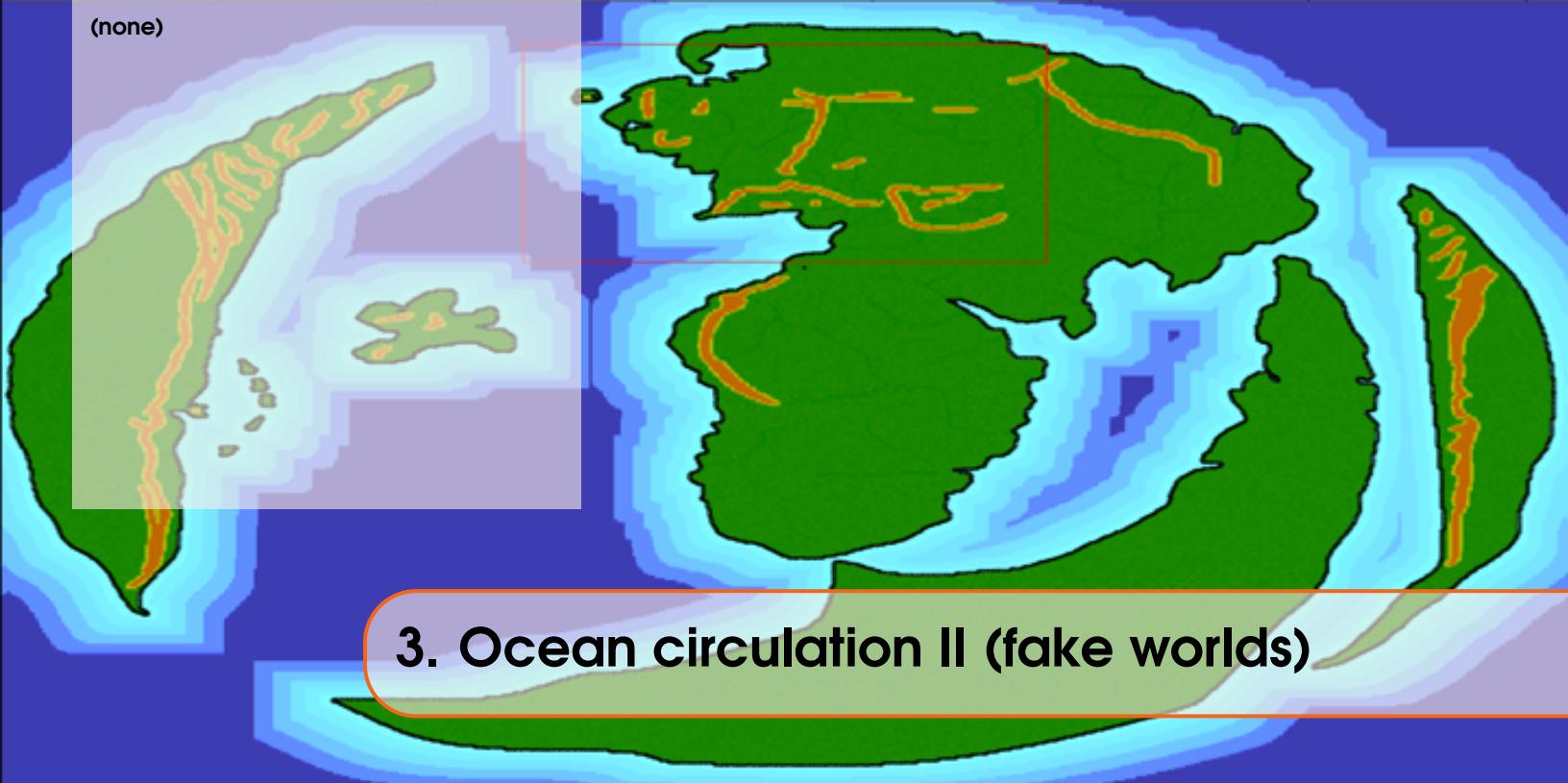
1.1 (none)

(none)



2.1 (none)

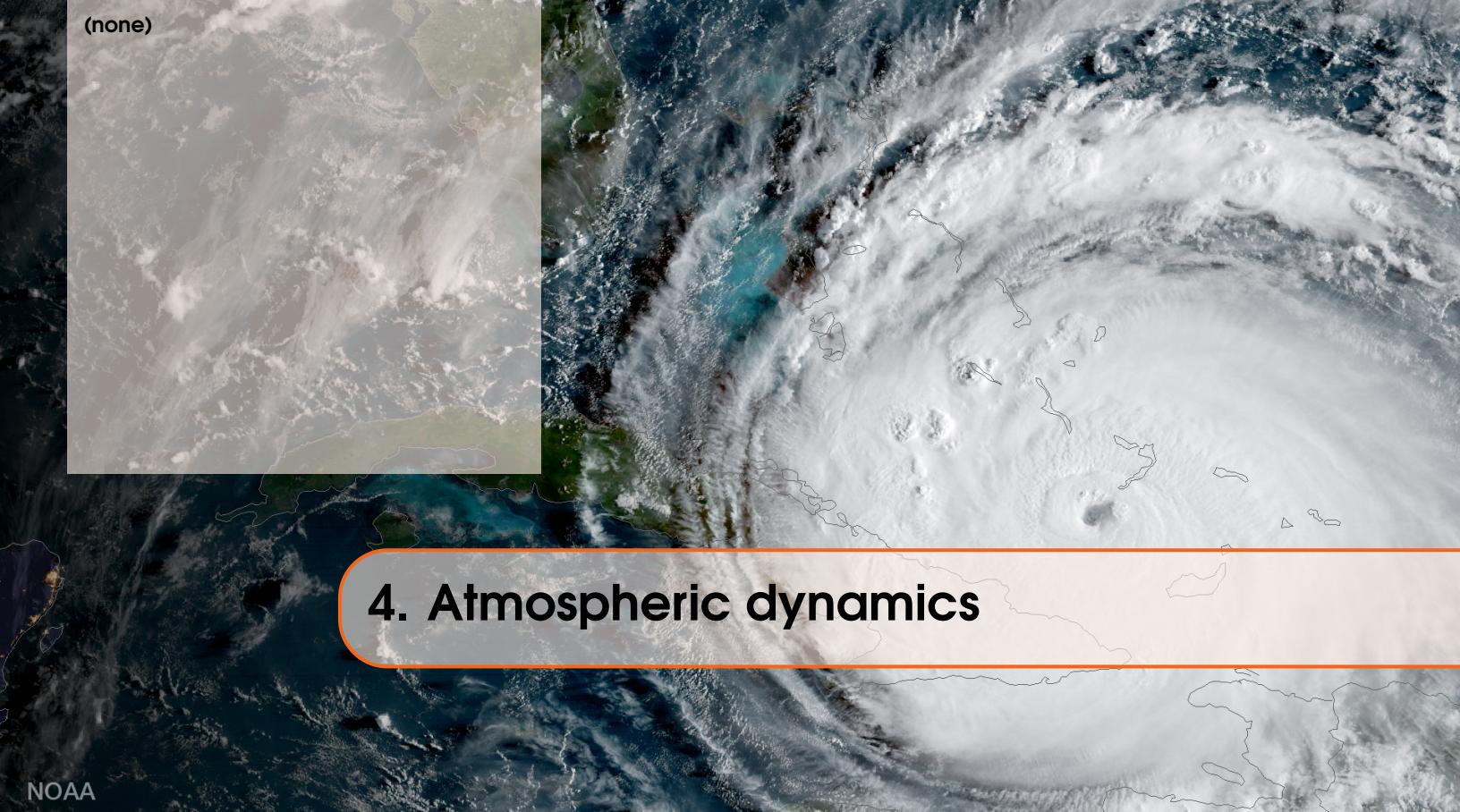
(none)



3. Ocean circulation II (fake worlds)

3.1 (none)

(none)

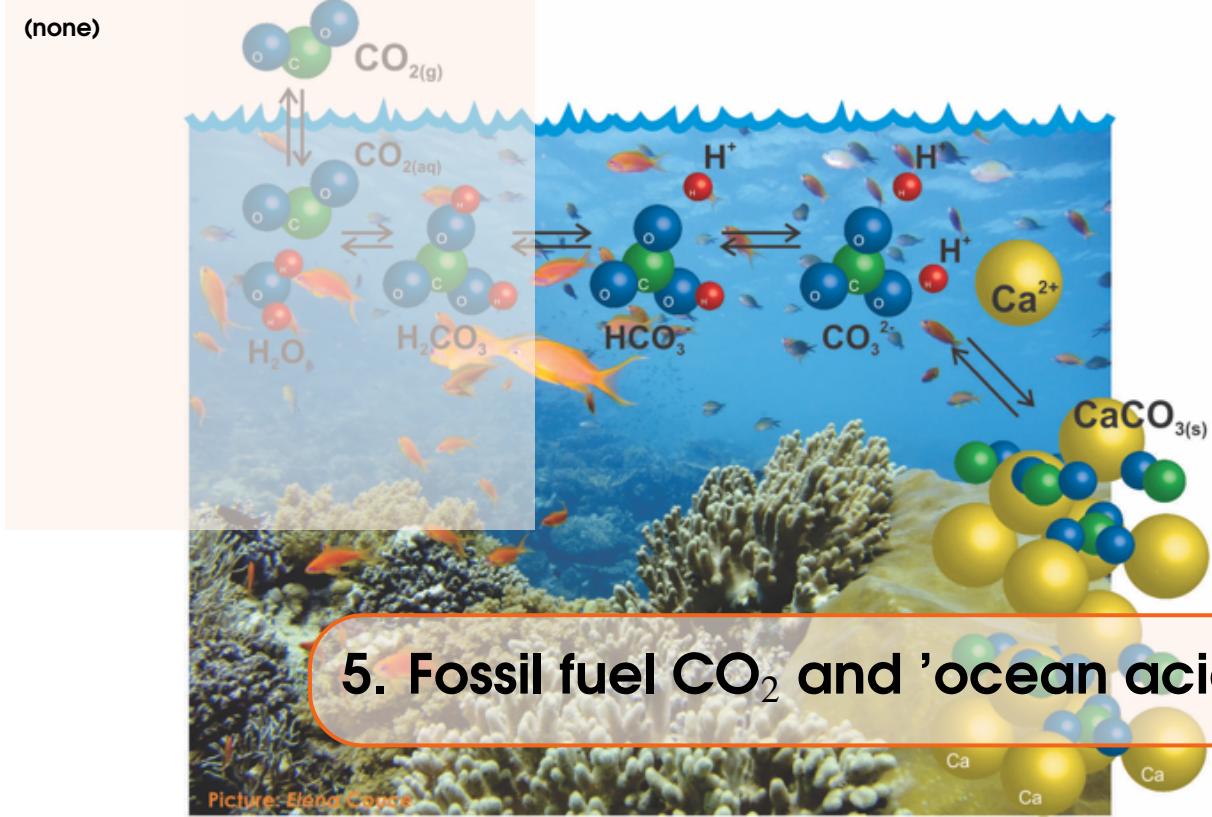


4. Atmospheric dynamics

NOAA

4.1 (none)

(none)



Picture: Elana Cenca

5.1 (none)

(none)

6. Ocean biogeochemical cycles

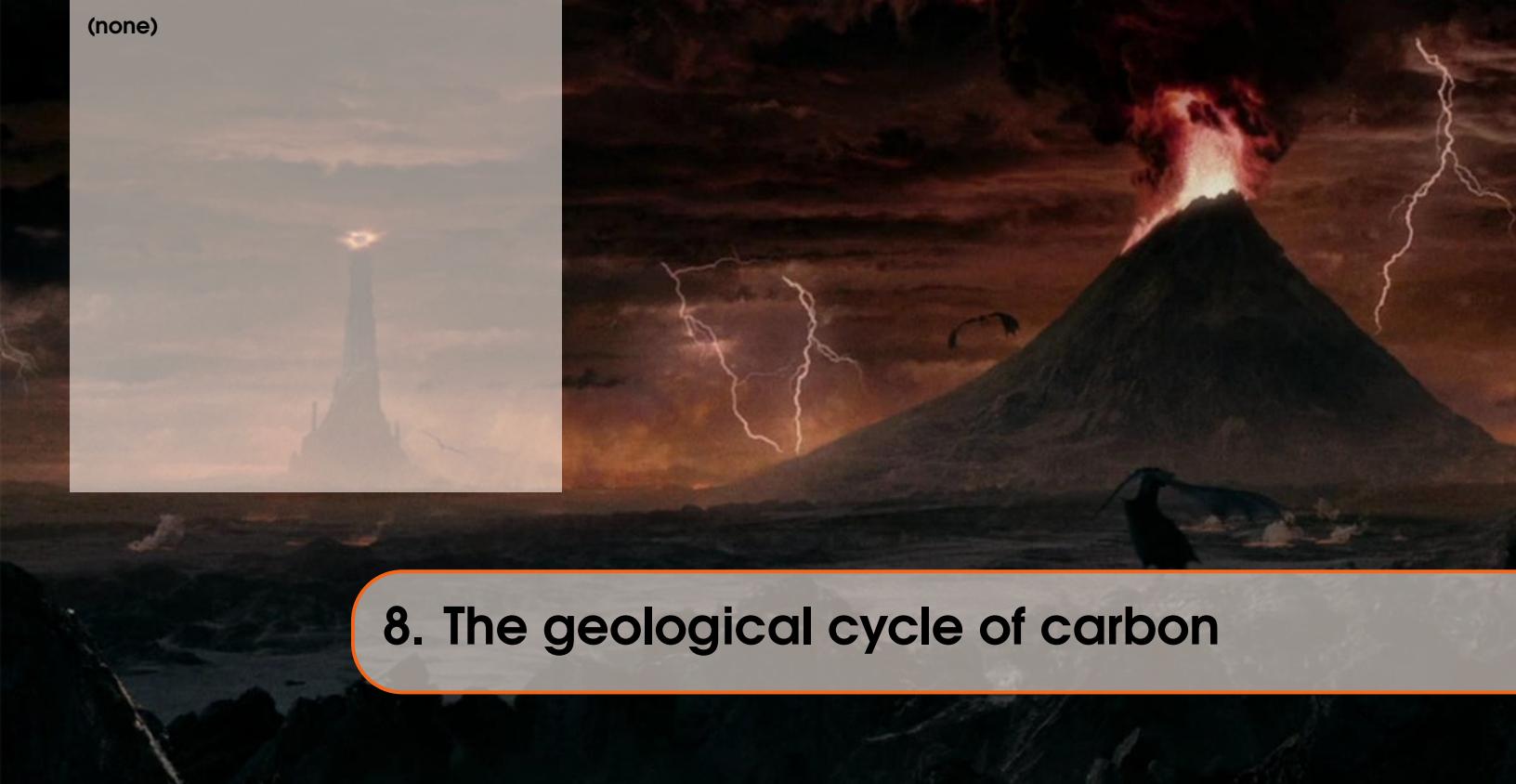
6.1 (none)

(none)

7. Marine ecosystems and dynamics

7.1 (none)

(none)



8. The geological cycle of carbon

8.1 (none)

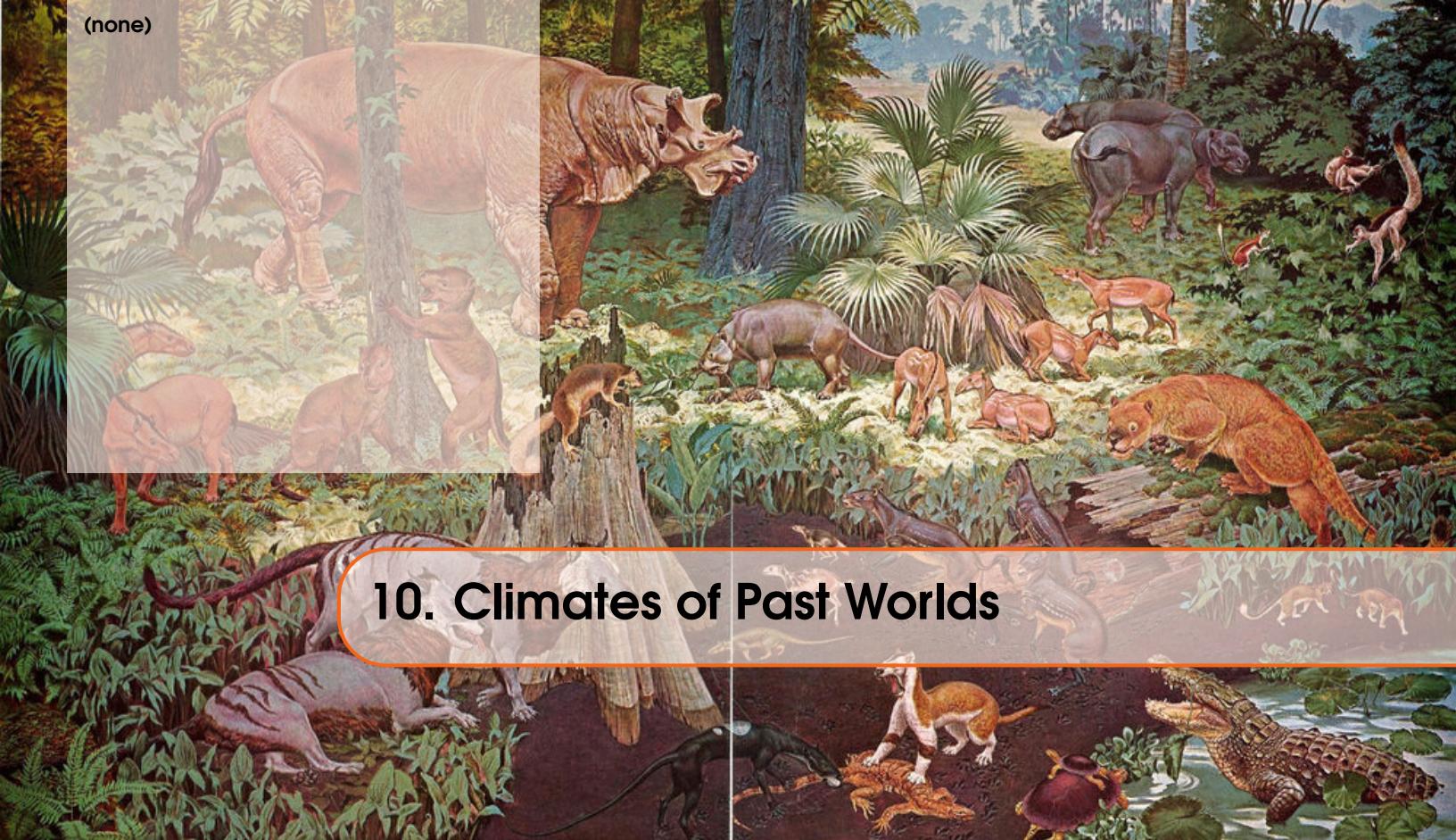
(none)



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

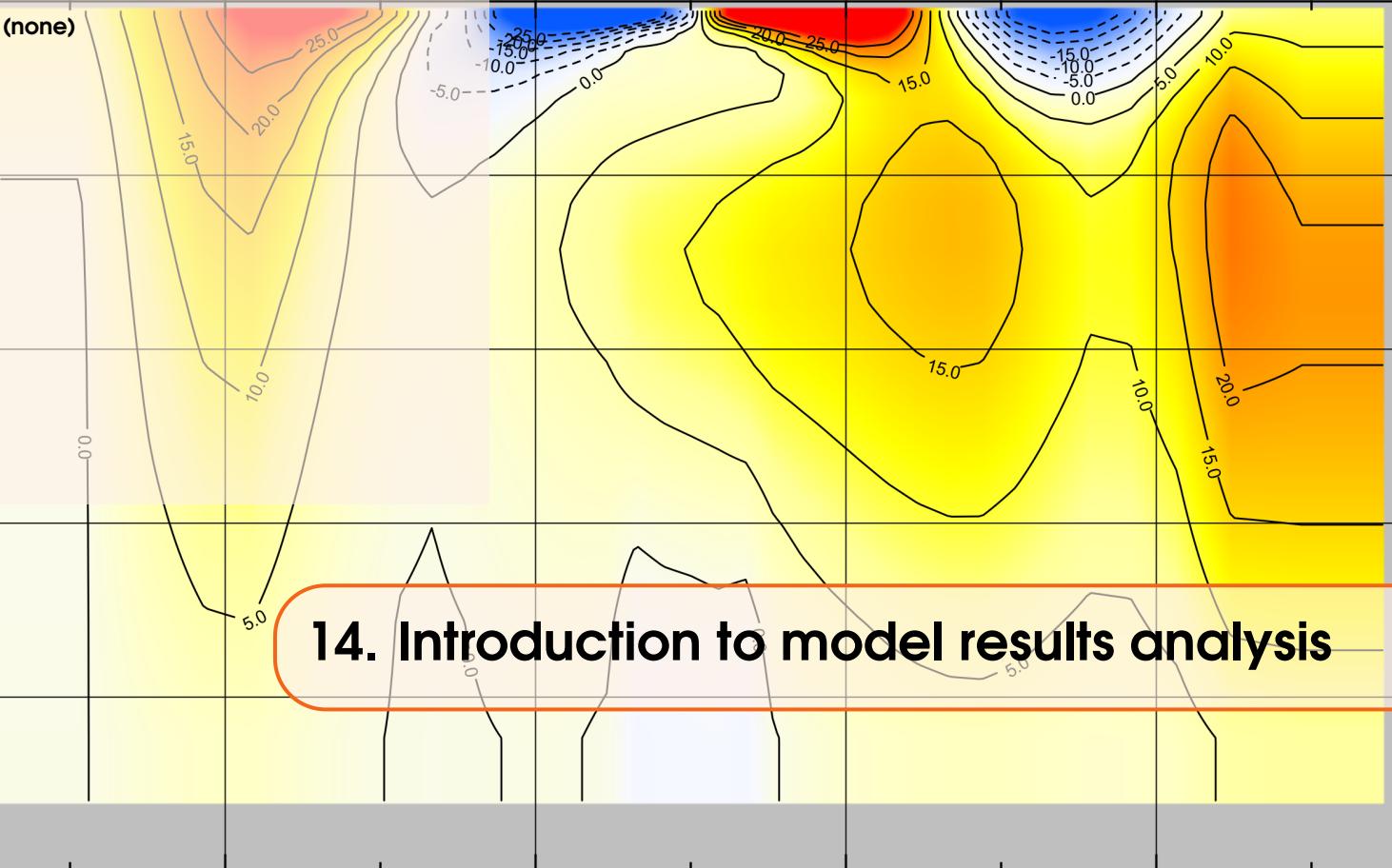
12.1 (none)

(none)

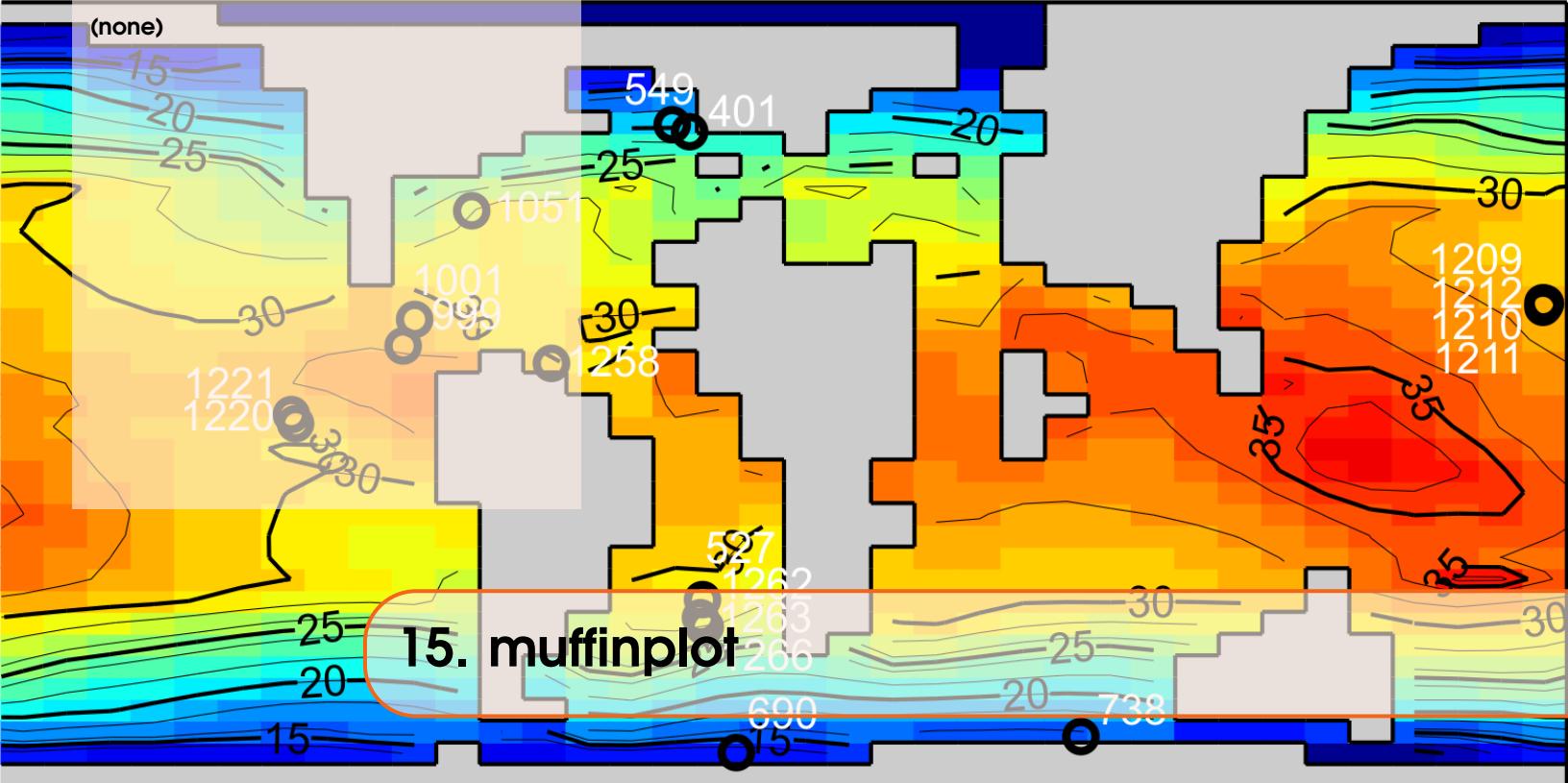


13. cookie model output

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**¹ 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 proceed² (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).

¹As 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 matlabananas 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.

²Note 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.³

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.⁴
- 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.⁵

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.

³In 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).

⁴The file example provided is: wordrake.txt.

⁵You 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.

5. 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 cookiegen generated, delete the lines in-between:

and the paste in the contents of TRACERCONFIG.AGE.txt that you have just copied.⁶

LAB.2.1.EXAMPLE would also work.

6. In place of the automatically-generated *user-config* in OUTPUT, use USERCONFIG.AGETRACER.SPIN, which you can find in the sub-directory: CGENIE.userconfigs.

+ carbon If you want a basic (abiotic) ocean carbon cycle as per Chapter 5:

5. 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.

⁶If 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 filename⁷ 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**.⁸ 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 editor⁹. 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.
```

⁷i.e. the single parameter passed to **makecookie** when it is invoked at the command line.

⁸Note 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.

⁹But make sure they retain a .m rather than .txt filename extension

¹⁰If 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.
```

¹¹This need not be changed.

¹²This need not be changed.

¹³Note 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!^{14,15}
- 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).

¹⁴Because it assumes that you are going to go on and draw something ...

¹⁵Edit 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 anything¹⁶:

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.¹⁷

¹⁶Note 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.

¹⁷If 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 ($5^{\circ}C$)¹⁸ 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.¹⁹

- The sea-ice diffusivity parameters are adjusted for improved stability²⁰:

```
# 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 seasonal²¹:

```
# 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.

¹⁸These parameters could potentially be increased further (to say ... $10^{\circ}C$).

¹⁹Instead, a salt forcing in the desired deep-water formation location, balanced by a freshwater flux elsewhere, can be defined.

²⁰The original parameter setting was:

gs_11=6200.000

²¹The 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.²² This includes the same changes as per made for the 8-level ocean circulation model based configuration:

- The ocean starts warmer (5°C).
- There is no (salt / freshwater) flux adjustment.
- The sea-ice diffusivity parameters are adjusted for improved stability²³.
- 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.²⁴ 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
```

²²If 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.

²³The original parameter setting was:
`gs_11=3573.718`

²⁴The 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 file²⁵: 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
#####
# ****
```

²⁵In 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⁻¹*.

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 properties²⁶ – 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].

²⁶See: '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 ocean²⁷ (see HOW-TO

²⁷If 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 *PO₄* 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 m d^{-1})²⁸. 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’²⁹ value for the sinking rate. At the same

²⁸It 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.)

²⁹Sufficient 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`³⁰, 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*³¹.

- 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,

³⁰By default it is set to zero which is interpreted as allowing complete reactant consumption as per previously.

³¹There 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**³² 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-userconfigs³³ 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 there³⁴ 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!

³²Which by default will appear in a directory: OUTPUT

³³The directory parameter passed to runcookie.sh is then just /

³⁴If 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 selected³⁵.
- `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 selected³⁶.
- `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.)

³⁵But 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.)

³⁶But 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`)³⁷. 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).³⁸

³⁷(rather than the maximum number of ocean levels as usually applied)

³⁸Note 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 GCMs³⁹. 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.

³⁹Example 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.⁴⁰

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 down⁴¹, 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 *CaCO₃* stability calculation) is slightly restricted in #3 (the upper limit increased to 1000m).

⁴⁰To change this, simply search-and-replace the 3 in the .k1 file with e.g. 1.

⁴¹(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 36×36 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 36×36 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 ...

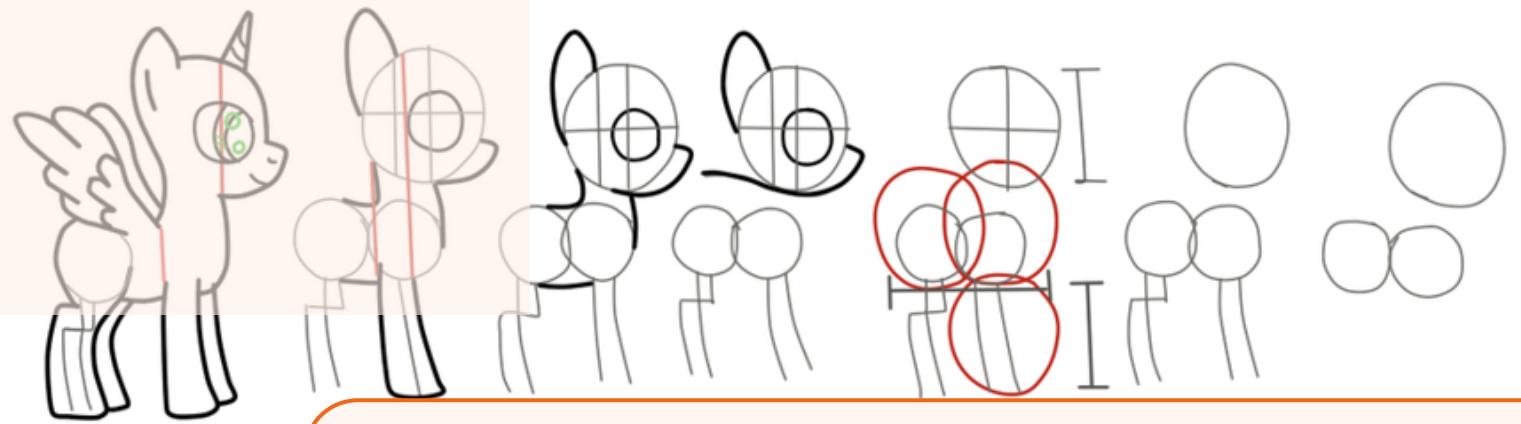
To make modified (e.g. wind fields) versions of either of the 2 modern 'paleo' format configurations:

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).



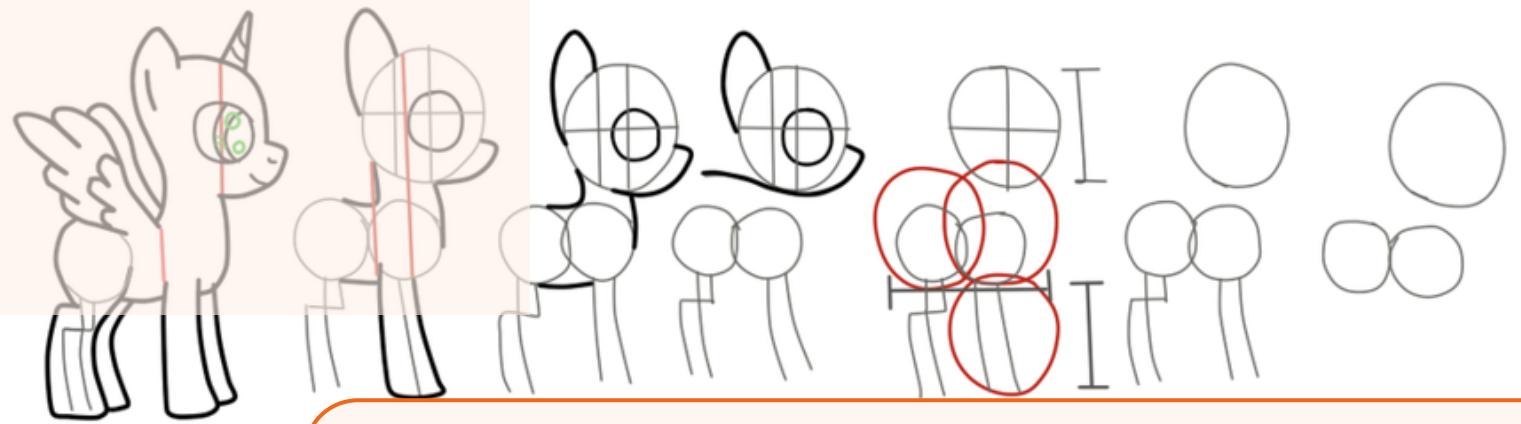
18. FAQ

18.1 **(none)**



19. HOW-TO (technical)

19.1 (none)



20. HOW-TO (experiment)

20.1 (none)

(none)

21. cookie Development

21.1 (none)