

# cGENIE HOW-TO

Andy Ridgwell

January 2, 2012

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>HOW-TOs: Getting started</b>	<b>4</b>
2.1	Do some thing dumb . . . . .	4
2.2	Install cGENIE . . . . .	4
2.3	Find configurations for cGENIE . . . . .	4
<b>3</b>	<b>HOW-TOs: General</b>	<b>5</b>
<b>4</b>	<b>HOW-TOs: Model output</b>	<b>6</b>
4.1	Set the frequency of time-series and time-slice output . . . . .	6
4.2	Diagnose orbital (insolation) changes . . . . .	6
<b>5</b>	<b>HOW-TOs: Climate</b>	<b>7</b>
5.1	Adjust solar forcing in a time-dependent manner . . . . .	7
<b>6</b>	<b>HOW-TOs: Ocean biology and biogeochemical cycling</b>	<b>9</b>
6.1	Configure an abiotic ocean . . . . .	9
6.2	Specify the CaCO <sub>3</sub> :POC export ratio . . . . .	9
6.3	Prescribe biological export production . . . . .	10
6.4	Include a R-DOM cycle in the ocean . . . . .	11
6.4.1	R-DOM degradation . . . . .	11
6.5	Include a Cd cycle in the ocean . . . . .	11
6.6	Determine the CH <sub>4</sub> flux required to achieve a particular atmospheric pCH <sub>4</sub> value . . . . .	11
<b>7</b>	<b>HOW-TOs: Forcings of the system</b>	<b>13</b>
7.1	Prescribe a an injection of radiocarbon-dead DIC . . . . .	13
7.2	Prescribe a spatial map of benthic tracer release . . . . .	14
<b>8</b>	<b>HOW-TOs: Sediments and weathering</b>	<b>15</b>
8.1	Spin-up the full marine carbon cycle including sediments . . . . .	15
8.2	Run the sediments at higher resolution . . . . .	16
8.3	Include shallow water depositional systems . . . . .	16
8.4	Accelerate the weathering-sedimentation mass balance ('GEMlite') . . . . .	16
<b>9</b>	<b>HOW-TOs: Visualization</b>	<b>20</b>
9.1	Generate ensembles and visualize their output using Mathematica . . . . .	20

<b>10 HOW-TOs: Miscellaneous</b>	<b>21</b>
10.1 Speed up the model . . . . .	21
10.2 Chop genie runs into manageable pieces (useful for very long runs) . . . . .	22
<b>11 Contact Information</b>	<b>23</b>

# **1 Introduction**

This document provides HOW-TO's for cGENIE users; where obvious changes are made; now items include instructions for speeding up the model and incorporation of (ocean) cadmium.

## 2 HOW-TOs: Getting started

### 2.1 Do some thing dumb

Easy! Just close your eyes and change some parameter values at random. Better still, start using the model without reading the manual first ...

### 2.2 Install cGENIE

See: *cGENIE Quick-start Guide*. (Also refer to the **READ-ME** file for e.g., details of changes in configuring and running cGENIE compared to GENIE.)

### 2.3 Find configurations for cGENIE

A series of (example) cGENIE configurations are provided, many of which are detailed in full in the *cGENIE Tutorial* document. Example configurations comprise *base-config* and *user-config* files, plus any *forcings* needed.

- cGENIE *base-configs* are stored in:  
    /cgenie/genie-main/configs  
    and all start with 'cgenie\_', for example:  
    cgenie\_eb\_go\_gs\_ac\_bg\_hadcm3l\_eocene\_36x36x16\_2i\_080928\_BASE.config
- cGENIE user configs are stored in:  
    /cgenie/genie-userconfigs
- cGENIE forcings are stored in:  
    /cgenie/genie-forcings

### **3 HOW-TOs: General**

## 4 HOW-TOs: Model output

### 4.1 Set the frequency of time-series and time-slice output

See: *cGENIE user-manual* (section 5).

### 4.2 Diagnose orbital (insolation) changes

Two new misc category time-series files have been provided:

`biogem_series_misc_ocn_insol.res`

and

`biogem_series_misc_ocn_swflux.res`

with the the SW (shortwave) flux (`swflux`) being equivalent to the incident strength of solar radiation at the surface (`insol`) but accounting for the prescribed planetary albedo. Both variables are calculated and saved on a global mean ocean grid basis (2nd data column) and have units of W m<sup>-2</sup>.

In addition, to help diagnose orbital variability, `biogem_series_misc_ocn_insol.res` includes two further insolation variables (3rd and 4th columns). These reflect the strength of insolation at a single point in the annual cycle and at discrete latitudes (i.e. `j` grid indices). (The insolation at 2 different latitudes are saved so that both .e.g. N and S hemisphere insolation signals can be simultaneously recorded.)

Three new namelist parameters are provided to configure this:

1. `bg_par_t_sig_count` – which sets the BIOGEM 'time-step' in the annual cycle at which the insolation value will be saved. e.g. for 96 time-steps in the ocean physics, and a 2:1 GOLDSTEIN:BIOGEM gearing (the default for the 16 level configuration), there are 48 BIOGEM time-steps. (It is left to the user to work out which part of the annual cycle *cGENIE* starts at (i.e. time-step #1) – I haven't a clue ...)
2. `bg_par_sig_j_N` – sets the '`j`' value for a northern hemisphere (but could be southern) snap-shot.
3. `bg_par_sig_j_S` – sets the '`j`' value for a southern hemisphere snap-shot.

## 5 HOW-TOs: Climate

### 5.1 Adjust solar forcing in a time-dependent manner

The value of the solar constant in cGENIE is set by the *namelist parameter*:

`ma_genie_solar_constant` and by default is set to 1368 W m<sup>-2</sup>, i.e.:

```
ma_genie_solar_constant=1368.0
```

Specifying a different value for `ma_genie_solar_constant` in the user config file allows the solar forcing of the EMBM to be altered. For example, to induce a 'snowball Earth' like state under a solar constant applicable to the late Neoproterozoic (some 6% less than modern) you would set:

```
ma_genie_solar_constant=1330.56
```

Modification of `ma_genie_solar_constant` can be turned into a time-dependent forcing of solar forcing, but only by frequent re-starting using a sequence of short model integrations.

Alternatively, a crude (temporary) hack is provided to allow a semi-continual adjustment of solar forcing. Whether you wish to vary the solar constant in a time-dependent manner is determined by the *parameter*:

`bg_ctrl_force_solconst`. By default this is set to `.false..` By adding to the user config file:

```
bg_ctrl_force_solconst=.true.
```

a time-varying change in the value of the solar constant will be imposed. For this, BIOGEM will expect the presence of a file called `biogem_force_solconst_sig.dat` in the forcing directory<sup>1</sup>.

`biogem_force_solconst_sig.dat` must contain two columns of information: the first is a time marker (year) and the second is a paired value for the solar constant. In the current crude incarnation of this feature, the time markers (1st column) **must** correspond exactly to the time markers in the time-series specification file<sup>2</sup>. GENIE will exit with an appropriate error message if this is not the case.

Seasonal solar insolation is re-calculated each year with a call to:

```
radfor(genie_solar_constant)
```

(EMBM file: `radfor.F`) at the start of the time-stepping loop in `genie.F`. At each time-marker, BIOGEM sets the value of `genie_solar_constant` equal to the corresponding value specified in:

`biogem_force_solconst_sig.dat`. Thus, regardless of how closely-spaced the time-marker years are, (seasonal) solar insolation is only adjusted every year. For a longer time-marker interval than yearly, no interpolation is performed on the series of solar constant values, and in this way time-dependint solar forcing currently differs from the calculation of other forcings.

EXAMPLE: A simple file might look something like:

```
-START-OF-DATA-
0.5  1367.0
1.5  1366.0
2.5  1365.0
3.5  1364.0
4.5  1363.0
5.5  1362.0
6.5  1361.0
7.5  1360.0
8.5  1359.0
9.5  1358.0
10.5 1357.0
-END-OF-DATA-
```

---

<sup>1</sup>REMEMBER: The location of which is specified by the namelist parameter `bg_par_fordir_name`.

<sup>2</sup>REMEMBER: The filename of which is specified by the namelist parameter `bg_par_infile_sig_name`.

which will decrease the value of the solar constant by 1 W m<sup>-2</sup> each year. Note that because the solar forcing is only updated each year (with the call to **radfor.F**), the first year will be characterized by climate with a solar constant of 1368 W m<sup>-2</sup>, the default. Although BIOGEM sets a new value of **genie\_solar\_constant** (1367 W m<sup>-2</sup>) mid way through the first year, it is only at the start of the second year that solar insolation is recalculated according to the reduction in solar constant.



## 6 HOW-TOs: Ocean biology and biogeochemical cycling

### 6.1 Configure an abiotic ocean

Biological productivity in the ocean can be completely turned off to create an abiotic ocean (why you would want to do this is another matter ... perhaps analyzing the solubility pump or a 'deep-time' and prior to significant marine life study ... (?)). The biological option is set by the *parameter* `bg_par_bio_prodopt` which by default takes a value of "1N1T\_P04MM" which selects the scheme described in *Ridgwell et al.* [2007a]. To have no biological production in the ocean, add the following line to the end of the *user-config* file (or edit the existing line in the section '--- BIOLOGICAL NEW PRODUCTION ---'):

```
bg_par_bio_prodopt="NONE"
```

With this set, you do not have to specify any biological production or remineralization namelist parameter values in the *user-config* file.

### 6.2 Specify the CaCO<sub>3</sub>:POC export ratio

In the default<sup>3</sup> 'biological' scheme in GENIE the CaCO<sub>3</sub>:POC export ratio from the surface ocean in BIOGEM is parameterized as a power law function of the degree of ambient oversaturation w.r.t. calcite [Ridgwell et al., 2007a,b]. The calculated CaCO<sub>3</sub>:POC ratio will vary therefore both spatially, particularly w.r.t. latitude (and temperature), as well as in time, if the surface ocean saturation state changes. The latter can arise from climatic (temperature) or circulation changes, or through a change in the DIC and/or ALK inventory of the ocean (such as resulting from emissions of fossil fuel CO<sub>2</sub>) or the re-partitioning of these species vertically within the ocean (e.g., as a result of any change in the strength of the biological pump).

There may be situations in which it is advisable to hold the CaCO<sub>3</sub>:POC export ratio invariant. For instance, considering the current very considerable uncertainties in the impacts of ocean acidification on marine calcifiers [Ridgwell et al., 2007a] the safest assumption is arguably to exclude any acidification impact on calcification and carbonate export. Specifying a spatially uniform value of the CaCO<sub>3</sub>:POC ratio (e.g. 0.25 or 0.3) also allows comparison with the results of early carbon cycle model studies. For deeper-time geological studies where little about marine carbonate production may be known *a priori*, a spatially uniform value represents the simplest possible assumption (e.g., Panchuk et al. [2008]).

BIOGEM can be told to use a prescribed (spatially and temporally invariant) 2D field of CaCO<sub>3</sub>:POC export rain ratios (instead of calculating these internally as a function of ocean chemistry) by setting the 'Replace internal CaCO<sub>3</sub>:POC export rain ratio?' namelist flag to `.true.`:

```
bg_ctrl_force_CaCO3toPOCrainratio=.true.
```

You must also then provide a 2D data field that specifies the value of the rain ratio at each and every surface ocean grid point. The filename of this field is set by default to:

```
bg_par_CaCO3toPOCrainratio_file="CaCO3toPOCrainratio.dat"
```

and the file must be located in the 'BIOGEM data input directory'<sup>4</sup>, which by default is:

```
bg_par_indir_name="$RUNTIME_ROOT/genie-biogem/data/input"
```

This 2-D field must be in the form of an ASCII file with space (or tab) separated values arranged in rows and columns of latitude and longitude. The format of the file must follow the GOLDSTEIN ocean grid with the first line being the most Northerly row, and the last line the most Southerly row of grid points. Data along a row is from West to East. The latitude of the first column of values must be consistent with the defined starting latitude of the model grid, which is specified by the namelist parameter `gm_par_grid_lon_offset`<sup>5</sup>. Examples are given in the code repository<sup>6</sup>.

---

<sup>3</sup>The default biological scheme is given by: `bg_par_bio_prodopt='1N1T_P04MM'`.

<sup>4</sup>`$RUNTIME_ROOT` being equal to `~/genie`.

<sup>5</sup>-260E by default

<sup>6</sup>e.g., `~/genie/genie-biogem/data/input/CaCO3toPOCrainratio.worbe2_preindustrial.dat`

If you are using a uniform value, it is an easy enough job to create a  $36 \times 36$  array of the value you so desire<sup>7</sup>.

If you want to hold a previously-calculated (spatially variable)  $\text{CaCO}_3$ :POC field constant, then the easiest way to achieve this is to copy the information contained in the *time-slice* results field:

`misc_sur_rCaCO3toPOC` in the results netCDF file `fields.biogem_2d.nc`<sup>8</sup>. Because this is a 3D data field ( $36 \times 36 \times 8$ ), carefully highlight just the surface ocean (2D) distribution (e.g., from the Panoply viewer) or extract from the netCDF file by some other means, and then copy and paste into:

`CaCO3toPOCrainratio.dat` (or whatever you have specified the filename as). When copying Panoply data, 'NaN's should be replaced by values of zero. Take care that the final (steady-state) time-slice is being copied and not the first (un-spunup) one ...

**TIP:** In order to quantify the importance of calcification feedbacks with  $\text{CO}_2$  and climate, two model integrations are required: one with the  $\text{CaCO}_3$ :POC ratio held constant and the other with it allowed to vary, thereby allowing the effect of a changing  $\text{CaCO}_3$ :POC ratio on the system to be elucidated.

## 6.3 Prescribe biological export production

Two possibilities:

### 1. Via a full prescription of all particulate fluxes in surface ocean

Create a full set of particulate (sediment tracer) flux forcings fields for the surface ocean, one for each biologically-related sediment tracer selected in the model, including isotopes (and trace metals). Everything except for the surface layer can be left as a zero (0.0) in the two 3D spatial fields required for each tracer.

You must also create a set of dissolved (ocean) tracer flux forcings fields for the surface ocean, one for each dissolved tracer associated with the particulates and selected in the model (including isotopes etc). The dissolved tracer flux fields must be created so as to exactly cancel out the particulate fields to conserve mass. For most tracers this is trivial, i.e., the fields for P in particulate organic matter (`sed_POP`) need be associated with fields for dissolved  $\text{PO}_4$  (`ocn_PO4`) which will simply be equal in magnitude but opposite in sign to POP. Complications start to arise for  $\text{CaCO}_3$  (2 alkalinities) and there is also the question of alkalinity changes associated with organic matter creation/destruction (via changes in  $\text{NO}_3$ ).

### 2. By just prescribing just the POC flux

An alternative has been provided enabling a full biological productivity in the surface ocean, but controlled by prescribing just the particulate organic carbon export flux. This 'biological' scheme is selected with::

```
bg_par_bio_prodopt="bio_POCflux"
```

What happens in practice is that the POC flux is used to calculate the equivalent  $\text{PO}_4$  change in the surface ocean, and then this is passed to the biological scheme and export production calculated 'as usual'. (The POC flux forcing is set to zero once the associated  $\text{PO}_4$  (uptake) flux has been calculated.)

A particulate (sediment tracer) flux forcing for POC in the surface ocean still has to be defined and selected, but no other forcings (including dissolved) are required. An example forcing configuration is given in `EXAMPLE_bio_POCflux` (and which can be obtained from [mygenie.seao2.org](http://mygenie.seao2.org)) and would naturally be selected by:

```
bg_par_forcing_name="TEST_bio_POCflux"
```

**NOTE:** Take care with dissolved organic matter (DOM) production, as a fraction of the specified POC flux will be converted into POC (and similarly for the other components of POM). Simplest is to set no DOM production if you are uncertain:

```
bg_par_bio_red_DOMfrac=0.0
```

---

<sup>7</sup>It doesn't matter if you specify a value over land because only values associated with wet cells will be acted on.

<sup>8</sup>You must have the 'miscellaneous properties' time-slice save flag set to:

`bg_ctrl_data_save_slice_misc=.true.` (the default) for this field to be saved.

**NOTE:** Also take care with the units of the flux forcing to the surface layer in the ocean (mol yr<sup>-1</sup>). Since GENIE-1 is often run on a equal area grid, it is not difficult to convert export production densities to mol yr<sup>-1</sup>. However, with an equal area grid, whatever the units are of the desired POC export field are do no matter – the global export can be set and the spatial distribution will then be appropriately scaled (as per in the `EXAMPLE_bio_POCflux` example). Also be aware that if there is insufficient PO<sub>4</sub> to support the require POC flux, you will not get your entire POC flux. Global total export may thus be somewhat less than that specified.

## 6.4 Include a R-DOM cycle in the ocean

### 6.4.1 R-DOM degradation

The parameter: `bg_ctrl.bio_remin.RDOM_photolysis` determines whether RDOM degradation is restricted to the surface layer and occurs only by/associated with photolysis. It can be `.true.` or `.false.` and by default is set to:

```
bg_ctrl.bio_remin.RDOM_photolysis=.false.
```

When set `.true.`, RDOM degradation is set to zero everywhere in the ocean except the surface layer. Here, the lifetime (parameter: `bg_par.bio_remin.RDOMlifetime`) is modified in *\*inverse\** proportion to the solar insolation integrated over the surface layer. (There is a field in the 2D netCDF of solar insolation at the ocean surface, and the average over the surface layer is approx 1/4 of this.). i.e., in lower latitude and higher insolation regions, the lifetime is shorter than specified by `bg_par.bio_remin.RDOMlifetime` (and by approx a factor of 1/4 of the solar insolation in W m<sup>-2</sup>).

## 6.5 Include a Cd cycle in the ocean

In order to run cGENIE with ocean cadmium cycle, the following *base config*:

`cgenie_eb_go_gs_ac_bg_itfclsd_161_JH_BASEFeCd` is provided (under SVN).

A typical experiment command line, using the *user config* file: `EXAMPLE_worjh2_PO4Fe_Cd_SPIN` (also provided under SVN), would look like:

```
./runCCgenie.sh cgenie_eb_go_gs_ac_bg_itfclsd_161_JH_FeCdBASE /  
EXAMPLE_worjh2_FeCd_SPIN 11
```

To submit this job to the cluster (from \$HOME):

```
qsub -q kitten.q -j y -o cgenie_log -S  
/bin/bash subcgenie.sh cgenie_eb_go_gs_ac_bg_itfclsd_161_JH_BASEFeCd /  
EXAMPLE_worjh2_PO4Fe_Cd_SPIN 10001
```

## 6.6 Determine the CH<sub>4</sub> flux required to achieve a particular atmospheric pCH<sub>4</sub> value

Unlike the concentration of CO<sub>2</sub> in the atmosphere, which if restored to a chosen value during a *spin-up* experiment, will remain at that value in a *continuation* experiment (if no other perturbation of the carbon cycle or CO<sub>2</sub> emissions have been prescribed), CH<sub>4</sub> in the atmosphere decays with a lifetime of ca. 8 years (with a small fraction dissolving in ocean surface waters and being oxidized in the ocean). Hence, atmospheric CH<sub>4</sub> *restored* to a particular value in a spin-up, requires that restoring to be maintained in any *continuation* experiment or CH<sub>4</sub> will quickly decay to zero. However, doing this (*restoring* CH<sub>4</sub> concentrations), prevents the effect of CH<sub>4</sub> emissions on being assessed (as the atmospheric composition is being help constant).

An alternative would be carry out the *spin-up* experiment with no *restoring* of atmospheric CH<sub>4</sub> (or *restoring* to zero), and then run the *continuation* experiment with no CH<sub>4</sub> *restoring*. This would enable e.g. CH<sub>4</sub> emissions experiments to be carried out and the change in atmospheric CH<sub>4</sub> in response to be simulated. The problem here is that the lifetime of CH<sub>4</sub> in the atmosphere scales with the CH<sub>4</sub> concentration. So in starting with no CH<sub>4</sub> in the atmosphere, the CH<sub>4</sub> lifetime is relatively short, and

the response to CH<sub>4</sub> emissions will be underestimated.

What is in effect 'missing' are the (natural) sources of CH<sub>4</sub> to the atmosphere such as wetlands, which at steady state, provide a CH<sub>4</sub> flux that balances the oxidation rate of CH<sub>4</sub> in the atmosphere (and ocean). cGENIE has a *parameter* for this: `ac_par_atm_wetlands_FCH4` (mol yr<sup>-1</sup>) (with the isotopic composition of this source set by: `ac_par_atm_wetlands_FCH4_d13C`). All that then remains is to determine the flux of CH<sub>4</sub> that balances the rate of oxidative loss for the desired atmospheric CH<sub>4</sub> concentration. TO do this:

1. Carry out a *spin-up* with atmospheric CH<sub>4</sub> *restored* to the desired concentration.<sup>9</sup>
2. Determine the total loss rate of CH<sub>4</sub> (including both atmospheric oxidation and invasion (and subsequent oxidation) into the ocean) – this is recorded in the *time-series* results file: `biogem_series_focnatm_pCH4.res`<sup>10</sup>.
3. Set the *parameter* `ac_par_atm_wetlands_FCH4` equal to this value.

An example of a spin-up in which a prescribed ('wetland') flux of CH<sub>4</sub> to the atmosphere is set, is described in:

*cGENIE.Examples* – *spin-up* example `EXAMPLE_p0055c_P04_CH4_SPIN2`

---

<sup>9</sup>For an example: see experiment `EXAMPLE_p0055c_P04_CH4_SPIN` described in *cGENIE.Examples*.

<sup>10</sup>Second column (the value in units of mol yr<sup>-1</sup>)

## 7 HOW-TOs: Forcings of the system

Note: All the *forcings* described here assume the 'new' (simplified) methodology for prescribing forcings. This methodology is enabled by including the parameter setting:

```
bg_ctrl_force_oldformat=.false.
```

Taking the example of the ocean (dissolved tracers): flux and restoring *forcings* are defined in the *forcings* specification file: `configure_forcings_ocn.dat`. As detailed in the notes to this file, there is a flag (COLUMN #6) which sets the spatial attributes of the *forcing* as follows:

```
3 == 3D
2 == 2D
0 == point
-1 == SURFACE
-2 == BENTHIC
```

the default (33) being that the forcing is applied uniformly to the entire (3D) ocean volume.

Options 3, 2, and 0, as: uniform 3D<sup>11</sup> (volume), uniform 2D (surface), and point forcing, respectively, require no additional (spatial) information. Hence, with options 3, 2, or 0 set, only an additional file specifying the time-dependent information for each forcing need be provided, in files of the form:

```
biogem_force_flux_ocn_xxx_sig.dat
```

for flux forcings, and

```
biogem_force_restore_ocn_xxx_sig.dat
```

where: xxx represents the mnemonic of the tracer (e.g., DIC is dissolved inorganic carbon. CH4 is methane, etc.).

Options -1 (SURFACE) and -2 (BENTHIC) require a 2D field to be provided in addition to the time-dependent information for each forcing. The grids for both are the same – i.e., all 'wet' grid points (non dry land) in the model. The filename for these 2D files is of the form:

```
biogem_force_flux_ocn_xxx_SUR.dat
```

for flux forcings, and

```
biogem_force_restore_ocn_xxx_SUR.dat
```

for restoring *forcings*.

Examples of point and 2D (benthic) ocean tracer *forcing* are given below.

For details of the 'old', fully 3D spatially-explicit forcing methodology, refer to the *cGENIE user-manual*.

### 7.1 Prescribe a an injection of radiocarbon-dead DIC

First ... a *base-config* with 14C tracers is needed, e.g.,:

```
cgenie_eb_go_gs_ac_bg_itfclsd_16l_JH_ANTH
```

```
cgenie_eb_go_gs_ac_bg_itfclsd_16l_JH_ANTHFe
```

(with Fe co-limitation of marine biological productivity).

Then, in the *user-config*, an appropriate *forcing* needs to be specified, e.g.:

```
pyyyx_FDIC_F13DIC_F14DIC
```

and under the heading --- FORCINGS ---, might look something like:

- `bg_ctrl_force_oldformat=.false.`  
`bg_par_forcing_name="worjh2_FDIC_F13DIC_F14DIC"`  
which prescribes a forcing of DIC plus its (13C and 14C) isotopes to the ocean (somewhere or everywhere).
- `bg_par_ocn_force_scale_val_03=0.0833e15`  
sets the flux (mol yr<sup>-1</sup>), which is equivalent to 1 PgC yr<sup>-1</sup>.
- To scale the isotopic composition:

---

<sup>11</sup>Note that here: '3D' does not mean a spatially explicit 3D pattern and hence the original ('old') way of specifying *forcings*, but instead: that the forcing is applied uniformly in 3D space (i.e., is in effect a volume *forcing*).

```
bg_par_ocn_force_scale_val_04=-60.0
bg_par_ocn_force_scale_val_05=-999.0
```

for example gives -60 per mil for 13C like methane and 14C that is pretty isotopically dead<sup>12</sup>.

- By default in the *forcing*, the duration of the emission 1 year, and can be re-scaled (e.g., to 1000 years duration) by:

```
bg_par_ocn_force_scale_time_03=1000.0
bg_par_ocn_force_scale_time_04=1000.0
bg_par_ocn_force_scale_time_05=1000.0
```

- Finally – the emission location is specified by, e.g.:

```
bg_par_force_point_i=18
bg_par_force_point_j=26
bg_par_force_point_k=7
```

Alternatively, the DIC release can be made over the entire ocean floor or simply sections (or depth intervals) of the ocean floor instead of a point source.

## 7.2 Prescribe a spatial map of benthic tracer release

Flux *forcings* to the ocean with a variety of different spatial attributes can be specified in the *forcings* specification file: `configure_forcings_ocn.dat`. As detailed in the notes to this file, there is a flag (COLUMN #6) which sets the spatial attributes of the *forcing*:

```
3 == 3D
2 == 2D
0 == point
-1 == SURFACE
-2 == BENTHIC
```

with the default requiring a (3D) spatially explicit field to be provided. Options -1 (SURFACE) and -2 (BENTHIC) require a 2D field to be provided. The grids for both are the same – i.e., all 'wet' grid points (non dry land) in the model. Templates for either can be created as follows:

1. Open up the *BIOGEM* results file: `fields_biogem_2d.nc` (any experiment).
2. Display the variable: `grid_mask`.
3. Select the **Array 1** tab (to display the actual gridded values rather than the color-coded map); highlight the grid of values and then copy-and-paste to a text editor.
4. You should have a grid of values, with a '1.0' representing ocean, and 'NaN' land. The NaNs can then be search-and-replaced to '0.0' and you have a grid valid for either the entire surface ocean or entire benthic surface.

From here: 1s can be replaced by 0s to remove unwanted locations.<sup>13</sup> In the forcing configuration file, if the COLUMN #5 flag ('scale flux forcing of tracer?') is set to 't', then the flux applied at each selected location is scaled such that the total applied flux is equal to that given in the *forcing* time-signal file.<sup>14</sup>

---

<sup>12</sup>Note this is on the scale of d14C not D14C

<sup>13</sup>This can be quite time-consuming and tedious and there is no particular short-cut :(

<sup>14</sup>The values in the forcing map need not be all 1.0 of course.

## 8 HOW-TOs: Sediments and weathering

### 8.1 Spin-up the full marine carbon cycle including sediments

By a 2-step process<sup>15</sup>:

#### 1. First-guess 'closed system' spin-up

As of **r4211** this it is possible to carry out the initial spin-up, **with** a solute input to the ocean via rivers, but also with the system configured 'closed', i.e.,:

```
bg_ctrl_force_sed_closedsystem=.true.
```

The weathering flux is subtracted from ocean cells overlying the sediments to balance the global budget and ensure a closed system. This subtraction involves partitioning the total global weathering flux between each ocean floor cell with a subtraction in proportion to the estimated CaCO<sub>3</sub> preservation and burial rate. To utilize this methodology now requires that the ROKGEM module is used, i.e., a *base config* such as:

```
cgenie_eb_go_gs_ac_bg_sg_rg_itfclsd_161_JH_BASE
```

A first guess for the weathering flux must now be prescribed. This could be derived from a previous closed system model experiment with no weathering flux specified (diagnosing weathering from total global CaCO<sub>3</sub> burial as described earlier), or from the literature, e.g., *Ridgwell* [2007] cites 20 Tmol HCO<sub>3</sub><sup>-</sup> yr<sup>-1</sup>, an equivalent CaCO<sub>3</sub> weathering rate of 10 Tmol yr<sup>-1</sup>:

```
rg_par_weather_CaCO3=10.00E+12
```

The following *user config* file

```
EXAMPLE_worjh2_P04_S36x36_SPIN
```

can be used for the closed system spin-up.

To launch an experiment, type (all in one line; notes space separators between line items in this document format):

```
./runCCSgenie.sh cgenie_eb_go_gs_ac_bg_sg_rg_itfclsd_161_JH_BASE /  
EXAMPLE_worjh2_P04_S36x36_SPIN 20001
```

To submit to the cluster type:

```
qsub -q kitten.q -j y -o cgenie_log -S /bin/bash subcgenie.sh  
cgenie_eb_go_gs_ac_bg_sg_rg_itfclsd_161_JH_BASE /  
EXAMPLE_worjh2_P04_S36x36_SPIN 20001
```

20000 years (20001 if using the default *time-series* saving points in order to record the last (annual averaged) year of the experiment) is probably about the minimum practical *spin-up* time. Primarily – you are looking for convergence in the mean wt% CaCO<sub>3</sub> value (averaged sediment composition), which is recorded in the *BIOGEM time-series* file:

```
EXAMPLE_worjh2_P04_S36x36_SPIN
```

#### 2. Open system spin-up

The last stage is an open system spin-up as described previously. The prescribed weathering flux (`rg_par_weather_CaCO3`) is revised and set equal to the diagnosed global CaCO<sub>3</sub> burial rate ('`Total CaCO3 pres (sediment grid)`') as reported in the SEDGEM module results file:

`seddiag_misc_DATA_GLOBAL.res`. In addition, an *open system* must now be specified in the *user config*:

```
bg_ctrl_force_sed_closedsystem=.false.
```

50000 years (50001 if using the default *time-series* saving points in order to record the last (annual averaged) year of the experiment) is probably about the minimum practical *spin-up* time. Again – you are looking for convergence in the mean wt% CaCO<sub>3</sub> value.

There is still some departure of ocean Ca and ALK inventories during the revised multi-stage spin-up compared to observed (and the initialized values), but this is substantially reduced compared to the original 2-part spin-up methodology as well as to a single spin-up methodology.

---

<sup>15</sup>This is a revised methodology compared to that described in the GENIE-1 HOW-TO.

**TIP:** Having completed the full marine carbon cycle spin-up, it is recommended that the CaCO<sub>3</sub>:rain ratio is set invariant – see earlier HOWTO. If the default CaCO<sub>3</sub> parameterization setting is retained, CO<sub>2</sub>-calcification feedback as described in *Ridgwell et al.* [2007b] is enabled.

**NOTE:** There is no climate feedback by default. To run experiments with feedback between CO<sub>2</sub> and climate, add:

`ea_36=y`

at the end of the *user config*.

## 8.2 Run the sediments at higher resolution

By default (as set in the *base config* file in `~/genie/genie-main/configs`) the SEDGEM sediment grid is configured at a resolution of 36x36 (and on an equal area grid), by:

```
SEDGEMNLONSOPTS='$(DEFINE)SEDGEMNLONS=36'
```

```
SEDGEMNLATSOPTS='$(DEFINE)SEDGEMNLATS=36'
```

Several data input files are required by SEDGEM consistent with the specified grid:

- A mask, which specifies the sediment grid locations (if any!) at which 'sediment cores' (see: *Ridgwell* [2007]) are to be generated at:  
`sg_par_sedcore_save_mask_name="sedgem_save_mask.36x36"`  
 The example provided on SVN contains some illustrative locations set (by a '1') for cores to be generated in.
- The required sediment grid topography (bathymetry):  
`sg_par_sed_topo_D="sedgem_topo_D.36x36"`  
 This particular grid is derived from observed bathymetry and excludes sediment locations shallower than the surface ocean layer (of the 8-level model) as described in *Ridgwell and Hargreaves* [2007].

The directory location of the required files is set by input directory namelist setting, which by default is:

```
sg_par_indir_name=" $RUNTIME_ROOT/genie-sedgem/data/input"
```

As described in *Ridgwell and Hargreaves* [2007], SEDGEM can be sub-gridded to a resolution of 72x72 (equal area). The following namelist additions are necessary to the *user config* file:

```
SEDGEMNLONSOPTS='$(DEFINE)SEDGEMNLONS=72'
```

```
SEDGEMNLATSOPTS='$(DEFINE)SEDGEMNLATS=72'
```

```
sg_par_sed_topo_D="sedgem_topo_D.72x72"
```

```
sg_par_sedcore_save_mask_name="sedgem_save_mask.72x72"
```

**NOTE:** Carbonate chemistry stability problems (= model crash) may occur in the 16-level configuration in conjunction with 72x72 resolution sub-gridded sediments. Who knows why?! :(

## 8.3 Include shallow water depositional systems

Relevant **EXAMPLES**:

## 8.4 Accelerate the weathering-sedimentation mass balance ('GEMlite')

Relevant **EXAMPLES**:

**Also see:** *cGENIE* user-manual 'FAQ' (further comments on **GEMlite** applicability).

A (pseudo) module is provided: '**GEMlite**' which provides a means of much more rapidly solving the weathering-sedimentation mass balance – i.e. the long-term (>10 kyr) carbon cycle processes and feedbacks. The motivation behind **GEMlite** is the stark disparity between the time-scales of ocean circulation and biological pump (ca. 0.1-1000 years) and those of sedimentation and weathering (2-20 kyr) and particularly the silicate weathering feedback (>100 kyr). This makes running *cGENIE* to an open system



steady state (with or without the silicate weathering feedback) challenging. Is there any way of 'accelerating' the calculation of the 'long tail' [Archer et al., 2009] of the CO<sub>2</sub> curve (e.g. in response to fossil fuel CO<sub>2</sub> emissions)?

The philosophy is as follows: the long-term weathering-sedimentation processes are effectively just an imbalance between the supply of solutes via weathering and preservation and burial of esp. carbonates in deep-sea (and shallow) marine sediments. For a small imbalance between weathering and sedimentation, atmospheric pCO<sub>2</sub> and climate (and hence the solute flux when including weathering feedbacks) will only change very slightly. For long intervals characterized by only a small imbalance in weathering-sedimentation the key assumption is made: Ocean circulation and the biological pump, and hence the \*gradients\* of dissolved species in the ocean can be considered \*invariant\*. Hence, for the purpose of solving weathering-sedimentation over an intervals of time: The ocean can be treated as a \*single box\*. It further assumes that: The ocean is initially in equilibrium with the atmosphere (w.r.t. CO<sub>2</sub>). (This latter assumption does place important limitations on under what circumstances GEMlite can be employed to accelerate experiments.)

This is what GEMlite does – it solves for weathering-sedimentation and applies the mass difference \*uniformly\* throughout the ocean (as if it were a single box), hence preserving the tracer gradients in the ocean. It also (optionally) calculates and re-partitioning of carbon between ocean and atmosphere. Because ocean circulation and the biological pump etc. do not have to be re-calculated, the accelerated quasi box-model phase can be calculated very considerably faster than the 'full' model. Obviously, if atmospheric pCO<sub>2</sub> and hence climate are changing at an appreciable rate then the assumption of invariance in ocean tracer gradients breaks down and it is not 'safe' to apply the accelerated calculation. Similarly, appreciable changes in nutrient inventories will affect the biological pump and hence also change tracer gradients.

The key to employing GEMlite, in addition to knowing when it is appropriate/not appropriate to employ it, is to decide what balance of accelerated (GEMlite) time-stepping vs. normal (full system update of ocean circulation, biological pump, etc.) time-stepping to employ. This division is implemented by creating a sequence of accelerated vs. non-accelerated time-stepping. This can be done in one of two ways:

1. Fixed sequence.

By default, GEMlite will employed a fixed, pre-determined sequence of accelerated vs. non-accelerated time-stepping. The parameters to specify this sequencing are:

**ma\_gem\_notyr** – which sets the number of years (the assumed time-step of GEMlite) for 'normal' time-stepping.

**ma\_gem\_yr** – which sets the number of years for accelerated time-stepping.

For instance: if **ma\_gem\_notyr**=50 and **ma\_gem\_yr**=50, you would have a sequence with 50 years of full updating, followed by 50 years of accelerated.

For instance: if **ma\_gem\_notyr**=10 and **ma\_gem\_yr**=90, you would have a sequence with 10 years of full updating, followed by 90 years of accelerated.

etc.

Note that the GEMlite cycle phase of 'normal' time-stepping is \*always\* done first.

Also note that choosing e.g. **ma\_gem\_notyr**=10 and **ma\_gem\_yr**=100, while appearing a desirably simple ratio, would result in the change-over point in cycle phase (to accelerated) occurring at the end of year 10, 120, 230, 240, etc. – something that might affect/influence your choice of data saving pattern (i.e., the sequence of time-points for time-series and time-slice data saving).

By default, the parameter values are: **ma\_gem\_notyr**=999999 and **ma\_gem\_yr**=1 meaning that in practice you will never get to the end of the 'normal' time-stepping phase. Note that these parameters are **integers** (setting real numbers, e.g. 1.0E6 will not work ...).

2. Adaptive sequencing.

Here, GEMlite attempts to be clever and optimizes the ratio between the duration of each phase of the cycle.

The motivation for this is that often in model experiments, environmental parameters will tend to change faster at the beginning of an experiment compared to towards the end. Fossil fuel CO<sub>2</sub>

release and its long tail of declining pCO<sub>2</sub> is a good example of this. Obviously this complicates the choice of a (fixed) ratio of cycle phases – 100:100 (or more likely: 1000:1000) might not lead to too much degradation of the simulation, but you would only gain a speed advantage of x2 for the experiment as a whole, which if 100-1000 kyr in total duration, is still going to be l o n g. On the other hand: 10:90 would give you a factor almost x10 increase in overall speed, but would seriously degrade the simulation during the initial, rapidly changing environment following CO<sub>2</sub> release. Adaptive sequencing adjust the time-stepping via 2 criteria:

- In the normal time-stepping phase, if the rate of change of pCO<sub>2</sub> is \*more than\* a specified threshold over any one year, then the total duration of this phase is extended by one year.
- In the accelerated time-stepping phase, if the total change in pCO<sub>2</sub> since the last normal phase is \*less than\* a specified threshold, then the total duration of this phase is extended by one year.

The result is that the phase durations are always a minimum of the values set by `ma_gem_notyr` and `ma_gem_yr`. If it is 'unsafe' to switch to accelerated mode, because pCO<sub>2</sub> is changing rapidly, then the model stays in normal mode. If it is safe to stay in the accelerated mode, because pCO<sub>2</sub> has not changed much in total during the phase, then the model stays in the accelerated phase.

The parameter names are default values for the two thresholds are:

- `ma_gem_adapt_dpCO2dt=0.1` (ppm yr<sup>-1</sup>)
- `ma_gem_adapt_DpCO2=1.0` (ppm)

but these will not necessarily be the ideal of any particular experiment (and some trial-and-error may be called for).

Adaptive time-stepping is enabled by setting:

`ma_gem_adapt_auto=.true.`

(by default it is `.false.`).

The switching between normal (non accelerated) and accelerated phases is saved in a time-series file:

`biogem_series_misc_gemlite.res`

Finally ... you will need a *base-config* that has GEMlite enabled. This actually requires nothing more than the addition of a couple of lines (to a *base-config* file):

`ma_flag_gemlite=.TRUE.`

which can go e.g. near the start of the file under # GENIE COMPONENT SELECTION. Plus:

`ma_kgemlite=xx`

which can go e.g. under # TIME CONTROL AND TIME-STEPPING.

Here, `xx` will depend on the time-step assumed in the base-config. This is likely to be either 96: the standard for most *base-configs*, or 48: for low resolution and faster model configurations, which typically have `.t48` in their filename. By convention, I name *base-configs* including GEMlite with `_gl`,

e.g. `cgenie-eb-go-gs-ac-bg-sg-rg-gl.p0000c.BASESLi.t48.config`

but you can name it BobTheLeglessPony for all I care.

The **most important** thing is to ensure you are not seriously degrading model fidelity (of carbon cycle simulation) by your adoption and configuration of GEMlite.

**Test** different assumptions of how the time-stepping phases are scheduled and compare (of possible) against a full experiment in which GEMlite is not used.

It is important to recognize that when the model switches into the GEM phase, it assumes all ocean tracer gradients are fixed, and updates only ocean composition as a whole according to weathering vs sedimentation imbalance (and also tries to re-equilibrium ocean and atm). As part of this, the flux to the sediments is taken from the average of the last year of the preceding normal phase, and fixed. This also means that the d13C of the CaCO<sub>3</sub> deposited to the sediments is fixed ... even if the ocean d13C is being updated and changing ... So, basically you lose the feedback that leads to d13C converging as sinks balance (weathering and volcanic) inputs<sup>16</sup>.

---

<sup>16</sup>Adjusting the fluxes themselves during the GEM intervals would break the underlying assumption inherent in the

The solution is to not run in the GEM phase for such long intervals – instead giving the normal phase a chance to make a brief update of ocean gradients and also d13C of export flux. BUT, if pCO2 hardly changes, *c*GENIE runs the risk of staying in the GEM phase for ever (ish)!

A further option:

`gem_adapt_auto_unlimitedGEM`

sets whether GEM is allowed an unlimited phase duration or not. By default it is `.false..` This means that the maximum GEM duration is limited to the normal `gem_yr` parameter. Also, if excessive (pCO2) drift occurs, the model will immediately switch to the normal phase.

By default then:

- `gem_notyr` specifies a MINIMUM duration for a normal phase..
- `gem_yr` specifies a MAXIMUM duration for a GEM phase.

Values of `gem_yr` much less than 100 are not advisable as you will not reestablish a new equilibrium gradient of tracers in the ocean in that time.

---

acceleration approximation.

## 9 HOW-TOs: Visualization

### 9.1 Generate ensembles and visualize their output using Mathematica

Use the scripts in `genie-tools/runscripts`. Read the `README` file in this directory for instructions. The Mathematica notebooks allow for easy ensemble generation across multiple variables (still using non-XML version though!), and also for interactive plotting of time-series and netcdf data with GUIs - including comparisons and movies etc. (note: you can get a free trial of Mathematica if you don't have access to it).

## 10 HOW-TOs: Miscellaneous

### 10.1 Speed up the model

\*sign\* You speed freak. Is this all you care about? What about the 'quality' of the simulation - does that mean absolutely nothing to you? Oh well ... There is a bunch of stuff that slows GENIE down that may not be absolutely essential to a particular model experiment. These include:

1. The number of tracers - if you don't need 'em, then don't select 'em! Selected tracers are automatically passed to GOLDSTEIN and advected/convected/diffused with ocean circulation. Similarly, BIOGEM does a whole bunch of stuff with tracers, particularly those which can be biologically transformed. All this is numerically wasteful if you aren't interested in them. Equally importantly, the more tracers you have selected the more careful you have to be in configuring the model. Superfluous tracers therefore cost more configuration time and/or increase the change of a model crash.
2. 'Tracer auditing' - the continuous updating and checking global tracer inventories to ensure that there is no spurious loss or gain of any tracer (i.e., a bug) has computational overheads associated with it. Whether this checking is carried out or not is set by the value of the flag `bg_ctrl_audit`<sup>17</sup>.
3. Time-series results saving. Model tracer (plus some 'physical') properties are brought continuously averaged in constructing time-series results files. Cutting down on time-series that you don't need will help minimize model run-time. The various categories of time-series that will be saved are specified by a series of namelist parameter flags. However, within each category (such as `ocn` tracers - `bg_ctrl_data_save_sig_ocn`) all properties will be saved - you are not given the option to save a defined sub-set (for example, DIC and PO4 in the ocean but not ALK). Note that time-series saving of data that is a 2-D average, such as atmospheric composition at the ocean-atmosphere interface, sediment composition at the ocean-sediment interface, or just ocean surface conditions, is less numerically demanding than mean values that have to be derived from a 3-D data field.
4. Time-slice results saving. If you have relatively few requested time-slices over the course of the model integration then this is unlikely to significantly impact the overall run-time (even with all possible data category save namelist flags set to `.true.`). However, note that if you have accidentally triggered the default time-slice saving interval (by having no data items in the time-slice specification file (`bg_par_infile_slice_name`)) you may end up with the model running about as fast as a 2-legged dog super-glued to a 10-tonne biscuit.
5. Alter the degree of asynchronicity between climate and biogeochemistry (see later HOW-TO).

As a very rough guide, the impact on total run-time of making various changes to the model configuration are listed as follows. Numbers are given as a percentage increase in total model run-time (using the `/usr/bin/time` linux command). Tracers selected in the ocean are DIC, ALK, PO4, O2, DOM\_C, DOM\_P, DOM\_O2, as well as <sup>13</sup>C isotopic components (DIC\_13C and DOM\_C\_13C) (+ T and S). The corresponding tracers are present in the atmosphere and as particulates. The model is run for 10 years as a new run (i.e., not loading in a restart file):

- ADD auditing  $\Rightarrow$  +15%
- ADD time-slice saving  $\Rightarrow$  +20%<sup>18</sup>
- ADD time-series saving  $\Rightarrow$  +15%
- REMOVE <sup>13</sup>C isotopic species (= DIC and DOC ocean tracers)  $\Rightarrow$  -10%<sup>19</sup>

---

<sup>17</sup>It is `.false.` by default.

<sup>18</sup>Because only a 10 year integration has been carried out with a time-slice saved at 10 years, the computational cost of time-slice saving is disproportionately high as displayed. With a longer integration, the relative cost of saving a time-slice will fall. In contrast, the computational cost as a fraction of total run-time of time-series saving and auditing is likely to remain the same.

<sup>19</sup>The speed gained by removing two tracers is not proportional to the fractional decrease in number of tracers (in this example reducing from 11 to 9 the number of tracers in the ocean gives only a ca. 10% improvement in overall speed).

The basic configuration for a faster 'lego box' cGENIE configuration consists of a 18x18 model grid and an 8 level ocean. The continents are in a zonally-averaged configuration and there is no topography in the oceans.

The model is accelerated by:

- (a) it's low resolution
- (b) taking 48 instead of 96 ocean timesteps per year
- (c) BIOGEM is only being updated every 4 rather than every 2 ocean time-steps.

Note: I am still in the process of carrying out numerical stability tests; in the end it may have to slow down a bit.

Because the time-stepping is different a new *rungenie* (make executable) script has to be used:

`runCCSgenie_t48.sh`  
(an equivalent ocean-only full carbon cycle including sediments has yet to be created!)

Using the following example *user config*:

`EXAMPLE_p0000b_SPIN_x1C02_S18x18`

and base config:

`cgenie_eb_go_gs_ac_bg_sg_rg_modern_18x18x8_0i_BASE_t48.config`

A typical run would look like something like:

```
./runCCSgenie_t48.sh
cgenie_eb_go_gs_ac_bg_sg_rg_modern_18x18x8_0i_BASE_t48 /
EXAMPLE_p0000b_SPIN_x1C02_S18x18 101
```

(Don't forget to do a make cleanall if you were using a different ocean configuration!)

In this configuration 100 years take about 40 seconds, 10 kyr would just take over an hour, and 100 kyr could be run overnight!

## 10.2 Chop genie runs into manageable pieces (useful for very long runs)

Use the scripts in `genie-tools/runscripts`. Read the `README` file in this directory for instructions.

## 11 Contact Information

- Andy Ridgwell: [bandy@seao2.org](mailto:bandy@seao2.org)