[simple carbon project] model
SCP-M user instructions
Australian National University, 2018

$$\frac{d(\mathbf{C})}{dt} = \left[\frac{d\mathbf{C}}{dt}\right]_{phys} + \left[\frac{d\mathbf{C}}{dt}\right]_{bio} + \left[\frac{d\mathbf{C}}{dt}\right]_{carb} + \left[\frac{d\mathbf{C}}{dt}\right]_{gas} + \left[\frac{d\mathbf{C}}{dt}\right]_{weath} \tag{22}$$

The calculation of atmospheric CO_2 is:

$$\frac{dAtCO_2}{dt} = \left[\frac{dAtCO_2}{dt}\right]_{gas} + \left[\frac{dAtCO_2}{dt}\right]_{NPP} + \left[\frac{dAtCO_2}{dt}\right]_{volcs} + \left[\frac{dAtCO_2}{dt}\right]_{weath} + \left[\frac{dAtCO_2}{dt}\right]_{anth} \tag{23}$$

[simple carbon project] model user guide

This document describes how to set up and run the [simple carbon project] model ("SCP-M"). It's called that because the model is pretty simple compared to many oceanographic or carbon cycle models. However, as shown in the model documentation (submitted for publication and will be deposited here soon), it does a pretty good job of reproducing modern and paleo observations from the carbon cycle, and perhaps even as well as the 'big models'. This (rather informal) user guide should be read in conjunction with the (formal) model documentation. Please use that document for the references for data quoted in this guide.

The model was set up to run against paleo data and try to elucidate the model parameter values that correspond to different paleo settings, in the model, thereby perhaps helping to solve a problem or two. We were most interested in the LGM-Holocene carbon cycle transition, "The LGM-Holocene dilemma" we called it, and undertook a series of model-data experiments with ocean and atmosphere d13C, D14C, CO_2 and oceanic carbonate ion proxy data. You might simply want to run the model manually, and check it against your own data.

The first section of this document provides an overview of installation, model files and process. The following section goes over the model and data files in a bit more detail. I hope this is useful. If you have any questions or are having problems getting the model to run, email: cameron.oneill@anu.edu.au. In due course, we may set up a SCP-M dedicated email address that isn't dependant on me being around – however, that may not be needed.

The standard "disclaimer" with these models applies. It has been compiled faithfully and with the best of intentions, but comes with no guarantees/warranties, you use at your own risk. Sorry, but apparently, I have to write that here. I'm/we're here to help if needed.
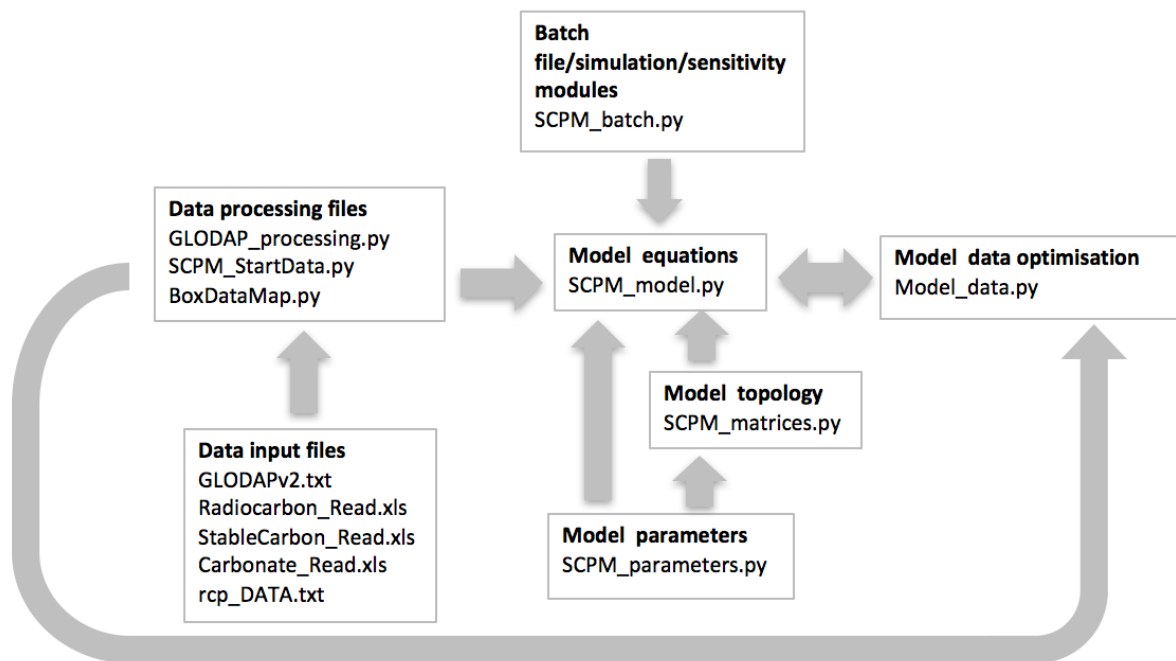
Model architecture and process

Installing SCP-M

Simply download the model directory "SCPM" from Zenodo and put it in a sensible place e.g. hard drive/Computer/home. SCP-M has been set up to access the constituent modules/files in their current locations within the present directory and folder locations, so it is best to leave them where they are, unless you want to change the directory paths in the modules. SCP-M has been compiled in Python 3.6, and was mainly operated using the Spyder interface from the Anaconda suite. I would suggest installing the latest version of Anaconda, as it will contain the dependencies for SCP-M.

Figure 1 shows the model's files structure. All of the Python files are located in the master "SCPM" directory, with so-named subdirectories for data files and model output files. The SCP-M suite is a mix of model and data processing files.
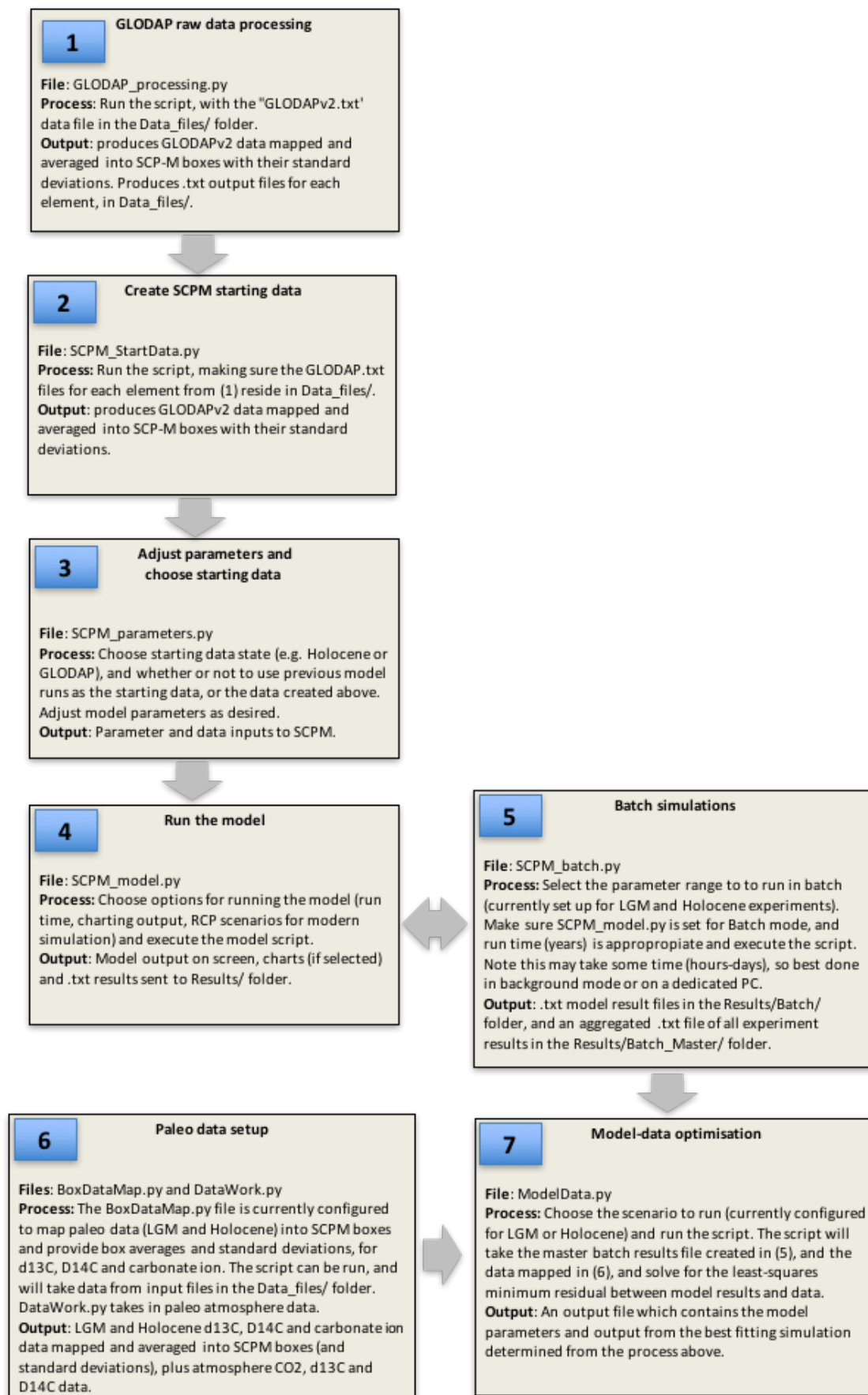
Figure 1: SCP-M files

SCP-M run process

The SCP-M model run process is combination of data processing and model runs (Figure 2) The initial data processing and setup described below, from the raw GLODAPv2 data file to the model inputs, have already been completed in the version of SCP-M which is located here. If you plan to use this model a lot, it's probably a good idea to go through the process for yourself.

The run process beings with processing raw GLODAPv2 data into useful box model data, in terms of geo-located box average element concentrations for those of interest: DIC, Alk, d13C, D14C, carbonate ion. This is achieved by running the "GLODAP_processing.py" script, (1) in Figure 2 To make this data more useful for the purposes we have used it for, it undergoes a second round of processing in (2) in

Figure 2. In "SCPM_StartData.py", a number of corrections are made to the GLODAPv2 data for the Suess effect on ocean DIC, d13C and D14C, and bomb radiocarbon for ocean D14C. The file also sets up the option for a number of different starting data states: Holocene, preindustrial (1751), or the GLODAPv2 start year (circa 1990).

In step (3) in the "SCPM_parameters.py" module, the model parameters are set, and you can choose whether you want to start the model with Holocene or preindustrial era (adjusted) data from (2) above. In this module, you also set the parameters for the flux processes in the model, such as ocean circulation, mixing, biological pump, rain ratio, piston velocity and scenario adjustments, such as for an LGM simulation. This module provides the conduit for data and parameter inputs to SCP-M.

Figure 2: SCPM model run procedure

1 **GLODAP raw data processing**

File: GLODAP_processing.py
Process: Run the script, with the "GLODAPv2.txt" data file in the Data_files/ folder.
Output: produces GLODAPv2 data mapped and averaged into SCP-M boxes with their standard deviations. Produces .txt output files for each element, in Data_files/.

2 **Create SCPM starting data**

File: SCPM_StartData.py
Process: Run the script, making sure the GLODAP.txt files for each element from (1) reside in Data_files/.
Output: produces GLODAPv2 data mapped and averaged into SCP-M boxes with their standard deviations.

3 **Adjust parameters and choose starting data**

File: SCPM_parameters.py
Process: Choose starting data state (e.g. Holocene or GLODAP), and whether or not to use previous model runs as the starting data, or the data created above. Adjust model parameters as desired.
Output: Parameter and data inputs to SCPM.

4 **Run the model**

File: SCPM_model.py
Process: Choose options for running the model (run time, charting output, RCP scenarios for modern simulation) and execute the model script.
Output: Model output on screen, charts (if selected) and .txt results sent to Results/ folder.

5 **Batch simulations**

File: SCPM_batch.py
Process: Select the parameter range to to run in batch (currently set up for LGM and Holocene experiments). Make sure SCPM_model.py is set for Batch mode, and run time (years) is appropriate and execute the script. Note this may take some time (hours-days), so best done in background mode or on a dedicated PC.
Output: .txt model result files in the Results/Batch/ folder, and an aggregated .txt file of all experiment results in the Results/Batch_Master/ folder.

6 **Paleo data setup**

Files: BoxDataMap.py and DataWork.py
Process: The BoxDataMap.py file is currently configured to map paleo data (LGM and Holocene) into SCPM boxes and provide box averages and standard deviations, for d13C, D14C and carbonate ion. The script can be run, and will take data from input files in the Data_files/ folder. DataWork.py takes in paleo atmosphere data.
Output: LGM and Holocene d13C, D14C and carbonate ion data mapped and averaged into SCPM boxes (and standard deviations), plus atmosphere CO2, d13C and D14C data.

7 **Model-data optimisation**

File: ModelData.py
Process: Choose the scenario to run (currently configured for LGM or Holocene) and run the script. The script will take the master batch results file created in (5), and the data mapped in (6), and solve for the least-squares minimum residual between model results and data.
Output: An output file which contains the model parameters and output from the best fitting simulation determined from the process above.

Step (4) is to run the model, taking the data and parameter inputs from steps (1-3). There are a number of user controls at the top of the module. This is where you choose how long you want to run the model for (years), whether you want charts, whether you want to analyse the Anthropocene emissions, and whether you want to store the model's output at the end of the run, for use as starting data in future model runs. If you select "Store results", you'll get a nice set of .txt files with the final year output from the model run.

Running the model once is good, but what if you want to explore the parameter space and see which of your model results best fit the data, and what parameter inputs led to those results? Using the batch module (5), you can schedule SCP-M to run as many times and across as many parameter values as you want. This file is currently set up to run LGM and Holocene-style experiments. You choose the type of experiment you want to run, and the range of input parameters to vary and hit go. Make sure that "SCPM_model.py" is set to "BatchRun=on" and "Charting=off" otherwise you'll end up with charts and printed model output for each run which will slow things down considerably (think, over a 1,000 simulations). At the end of the experiment, the script will aggregate model output into a master output file which will be deposited in Results/Batch_Master/. Note the output from our LGM and Holocene experiments are provided in the version on Zenodo.

If you want to optimise your model experiments against some data, there is an ancillary data file that maps and averages data into SCP-M boxes for this purpose, step (6). The "BoxDataMap.py" script is set up to read in data from MS Excel data sheets with following column structure and headings: Core_Ref_2, Lat, Long, Sample_Depth, Age, element, uncertainty. This is currently set up for LGM and Holocene data for ocean d13C, D14C and carbonate ion proxy. The script will map this data into SCP-M box coordinates, bin into LGM and late Holocene time periods, and then compute averages and standard deviation.

At step (7), you should have a batch master file that contains the results of your batch simulations, in Results/Batch_Master/ e.g. the "Output_all_Hol.txt" and "Output_all_LGM.txt" batch master output files from our experiments are provided, and a set of data (that you wish to compare the model with) of average element compositions (and standard deviation) in SCP-M boxes, as you will generate from the "BoxDataMap.py" module using the data spreadsheets provided for LGM and Holocene proxy data. Step (7) is about finding the best model run for your data. In our case, this was for LGM and late Holocene data. The "ModelData.py" script takes the output from (5) and (6) and solves for the parameters of best fit for the data, via a least-squares model results-data residual minimisation calculation. Basically, it searches through the model results that you generated from your batch runs and then finds the simulation that gave results closest to the data values, and tells you what parameters were inputted to deliver those results. Can be quite useful. This module is currently set up to solve for parameters of best fit for the LGM and late Holocene ocean and atmosphere proxy data. The output file is deposited in Results/ModelData_Outputs/.

SCP-M model files and procedures

GLODAPv2 data processing

The model uses GLODAPv2 bottle data as its starting base. The GLODAPv2 data is already downloaded and currently resides at /Data_files/GLODAPv2.txt.

To implement the mapping of GLODAPv2 data into SCP-M, simply run the "GLODAP_processing.py" file.

This will average the GLODAPv2 data into the SCPM boxes using the box dimensions specified (minimum and maximum depth, longitude, latitude), and create starting average box concentrations for phosphate, DIC, Alk, temperature, salinity, pressure, d13C, D14C, oxygen, silicate and pCFC12. The script calculates carbonate (CO_3^{2-}) from the GLODAPv2 data for DIC, Alk, salinity, phosphate and silicate. The script saves the box averages into .txt in Data_files/, with naming convention for GLODAP and the element. Each .txt file is an average concentration of the named element (e.g. DIC) in each SCP-M box. The standard deviation of values in each box is also produced in .txt files in the same directory. There are no routine user controls required to operate the GLODAP_processing module, unless the user wants to change the dimension or number of boxes (Figure 3).

Figure 3: Box latitude and depth coordinates for GLODAPv2 data mapping, in GLODAP_processing.py

```
59    return list(scen_twice)
60
61 # Set up SCP-M model coordinates
62
63 #Latitude SCP-M boxes (no longitude as this is zonally averaged)
64 Box1LatMin=-40.0
65 Box1LatMax=40.0
66 Box2LatMin=40.0
67 Box2LatMax=60.0
68 Box3LatMin=-40.0
69 Box3LatMax=40.0
70 Box4LatMin=-61.0
71 Box4LatMax=60.0
72 Box5LatMin=-80.0
73 Box5LatMax=-61.0
74 Box6LatMin=-80.0
75 Box6LatMax=80.0
76 Box7LatMin=-61.0
77 Box7LatMax=-40.0
78
79
80 # Depth SCP-M boxes
81 Box1DepthMin=0.0
82 Box1DepthMax=100.0
83 Box2DepthMin=0.0
84 Box2DepthMax=250.0
85 Box3DepthMin=100.0
86 Box3DepthMax=1000.0
87 Box4DepthMin=1000.0
88 Box4DepthMax=2500.0
89 Box5DepthMin=0.0
90 Box5DepthMax=2500.0
91 Box6DepthMin=2500.0
92 Box6DepthMax=6000.0
93 Box7DepthMin=0.0
94 Box7DepthMax=250.0
95
```

SCPM starting data

Once the GLODAPv2 data is mapped into SCP-M boxes as per above, and stored in the .txt files in Data_files/ (automatically generated by "GLODAP_processing.py", the data can

undergo a second round of processing to make it more useful for SCP-M model runs. The "SCPM_StartData.py" file offers three starting data states for the SCPM spin-up. GLODAP data year (assumed 1990), pre-industrial (1751) and the late Holocene average (6.6-0.2 kyr BP).

The GLODAPv2 dataset contains data sourced during the period 1972-2013, a non-steady period under the heavy influence of anthropogenic industrial and deforestation emissions. In order to take the model back to the late Holocene or pre-industrial periods, a number of adjustments must be made. These include for anthropogenic emissions ("The Suess effect") and atmospheric nuclear bomb testing. These events have changed the global inventory of CO_2, DIC, 13C and 14C. There are a number of estimates from the published literature of the magnitude of these perturbations, which can be used to correct the GLODAP data for a Late Holocene setting. SCPM_StartData incorporates estimates of Suess DIC from Sabine et al (2004), and d13C from Eide et al (2017). These basically subtract (in the case of DIC) or add back in (in the case of d13C and D14C) to cancel out the effects of anthropogenic emissions. Estimates of atmospheric bomb testing 14C are also subtracted out of the GLODAPv2 ocean data, according to Key (2001).

There are no user controls needed to execute the script.

SCPM_parameters.py

This file is the home for the SCPM dimensions, box set-up, conversion factors, parameters and is the main conduit for starting data feeding into the model. This module provides a complete listing of the model parameters and variables, with their descriptions, in the code.

There are a number of user controls at the top of the script, to be checked, prior to executing the script. You can determine whether the model simulation uses the starting data from the processes described above, or the output from the previous run as the start data. Note you need to go through the process above (NB its already set up if you don't do it) to generate a set of model output, before you can choose the previous model output as the starting data for subsequent runs.

For the first simulation of SCP-M, select "UseLastRunStart='off'" so that it uses the SCPM_StartData.py as the first run. For "StartData", choose from "HolAv" (Holocene average start data), "PreInd" (circa 1751 start data) or "GLODAP" (circa 1990 start data). If you want to implement "LGM" style settings for sea surface temperature, salinity, ocean volume, piston velocity and radiocarbon production (as per "LGM state" in the model documentation), select 1 for the "LGM" switch, otherwise ignore.

```
19 # Import data
20 from SCPM_StartData import SuessDIC
21 import SCPM_StartData as SCPM_Start
22
23 LastDir='Last_run/'
24 DataDir='Data_files/'
25
26
27 ##-------------------CONTROLS-------------------------------------------
28 # Switch to use the results from last run for starting data (must have selected option
29 # to store model results in the SCP-M model)
30 UseLastRunStart='off' # Choose 'on' or 'off'
31 StartData='HolAv' # Choose between 'HolAv', 'PreInd', 'GLODAP'
32 LGM=0 # 0 or 1 to switch on 'glacial' background settings
33
34 ##-------------------------------------------------------------------
35
36 #Time step setup
37 secsyr = 60.0 * 60.0 * 24.0 * 365.0 # 31,536,000
38 secskyr=secsyr*1e3
39 secsday=60*60*24
40
```

SCPM_matrices.py

This file contains the matrices that implement the ocean physical and biological fluxes in SCP-M. Each of the flux parameters has a matrix, which activates the parameter in SCP-M box space. You generally don't need to make any changes to this file in the course of running the model, unless you want to add more boxes or change the trajectory of the fluxes. Each row and column of the matrix corresponds to a box in SCP-M.

Figure 5: SCPM_matrices.py module

```
22
23 # Ocean physical------------------------------------------------
24
25 def makeT1mat(Psi1, Varr, Fract):
26     makeT1mat = np.zeros([7,7])
27     makeT1mat[3,3]=-1
28     makeT1mat[3,5]=1
29     makeT1mat[4,4]=-1
30     makeT1mat[4,6]=Fract
31     makeT1mat[4,3]=1-Fract
32     makeT1mat[5,4]=1
33     makeT1mat[5,5]=-1
34     makeT1mat[6,3]=Fract
35     makeT1mat[6,6]=-Fract
36     return makeT1mat
37
38 def makeT2mat(Psi2, Varr):
39     makeT2mat = np.zeros([7,7])
40     makeT2mat[1,1]=-1
41     makeT2mat[1,2]=1
42     makeT2mat[2,2]=-1
43     makeT2mat[2,6]=1
44     makeT2mat[3,1]=1
45     makeT2mat[3,3]=-1
46     makeT2mat[6,3]=1
47     makeT2mat[6,6]=-1
48     return makeT2mat
49
50 def makeE1mat(gamma1,Varr):
51     makeE1mat = np.zeros([7,7])
52     makeE1mat[3,3]=-1
53     makeE1mat[3,5]=1
54     makeE1mat[5,3]=1
55     makeE1mat[5,5]=-1
56     return makeE1mat
57
58 def makeE2mat(gamma2,Varr):
```

This module implements SCP-M. It contains the setup and operation of the model's equations. There are a number of controls you can vary before running the model.

- Charting: you can choose whether you want charts of the results for the model run. The model has a few charts pre-loaded (e.g. CO_2, d13C, d14C, DIC, Alk, carbonate, phosphate). If you are running the model in batch mode, best to turn this off as it will generate a chart at each model run (its going to slow it down if you're running many hundreds or thousands of model runs).
- BatchRun: if you're running the model manually, i.e. executing the SCP-M_model.py code, turn the BatchRun switch off. If you are going to run the model in batch model turn it on – this will enable the SCPM_batch.py module to execute the model script as part of the batch process.
- StoreResults: if you want to store the model run equilibrium output (basically, the model variable values at the end of the simulation) and keep it as the starting data for future simulations, i.e. a baseline that you are happy with (e.g. late Holocene or pre-industrial baseline with which to explore LGM or Anthropocene perturbations), switch this on. You must also switch on "UseLastRunStart" in the SCPM_parameters.py module.
- TerrestrialGeo, TerrestrialBios, Rivers, Carbonate: these controls switch on these functions in the carbon cycle. TerrestrialGeo turns on volcanism and weathering sources and sinks of CO_2. TerrestrialBios switches on the terrestrial biosphere sink and respiration fluxes of CO_2 carbon isotopes. Rivers switches on the river fluxes of carbon and alkalinity into the low latitude surface box from continental weathering. Carbonate switches on the ocean carbonate flux, water column and marine sediment dissolution. For most model runs, these are left switched on.
- AnthEmits: this switches on Anthropocene fossil fuel and deforestation emissions. If you switch this on, you must also choose a selection of a RCP scenario for these emissions sources (RCP 2.6, RCP 4.5, RCP 6.0 and RCP 8.5) under the heading "RCP settings, for modern scenarios" a few lines down. You must also select "Anthrop" for the timing of the model run. This will set the model to run for the period 1751-2100 (or any period you choose as long as it starts in 1751 and ends before 2100 in "AnthEmitStart" and "AnthEmitEnd".
- If you enable "Bomb Radiocarbon", you must also select "Anthrop" as the time interval for the model run. This will release 14C into the atmosphere over the period 1954-1963.
- Interval: this is where you set the time duration for the model run (years). If you are running the Anthropocene simulation, enter "Anthrop" here, or otherwise just enter the timing in years. Typically, the model takes 5-15,000 years to reach equilibrium in spin-up or after a perturbation, due to the slower carbon cycle processes such as carbonate compensation.

```
45
46
47 ##---------CONTROLS-------------------------------------------------
48 ## User adjusted options, use these to edit model settings
49
50 Charting='off'# switches charting on or off (turn off for batch runs)
51 BatchRun='on' # turn off for manual model runs (ie those runs operated from this module)
52 StoreResults='off' # turn on to store results for starting data for future runs (turn off fo
53
54 TerrestrialGeo=1 #1 is on, 0 is off – switch on volcanic emissions and weathering sink
55 TerrestrialBios=1 #1 is on, 0 is off – switch on terrestrial biosphere
56 Rivers=1 #1 is on, 0 is off – switch on river fluxes of carbon, alkalinity and phosphorus
57 Carbonate=1 #1 is on, 0 is off – switch on carbonate chemistry
58 AnthEmits=0 #1 is on, 0 is off – switch on anthropogenic emissions
59 BombRadiocarbon=0 #1 is on, 0 is off – switch on bomb radiocarbon
60 ## Remember – turn off 'AnthEmits' and 'BombRadiocarbon' when doing paleo runs
61 # with a time horizon that is not 'Anthrop' i.e. user-specified in years
62
63
64 # RCP settings, for modern scenarios
65 rcpname='rcp60' # Also need to adjust FFems, DeforestC, SST scenarios immediately below
66 rcp_FFems=rcp60_FFems
67 rcp_DeforestC=rcp60_LUCems
68 rcpSST=rcp60_SST
69 # rcp names ='rcp26','rcp45','rcp60','rcp85'
70
71 #Time calculations for modern day scenarios
72 AnthEmitStart=1751
73 AnthEmitEnd=2100
74 Anthrop=AnthEmitEnd-AnthEmitStart
75 BombStart=1954-AnthEmitStart
76 BombEnd=1963-AnthEmitStart
77
78 # Model run time (years) – important setting
79 interval = 10000 #Anthrop # Model run time in years. Enter a time period in years,
80 # or if running the modern day emissions scenario, enter "Anthrop"
81
```

SCPM_batch.py

If you are bored of running the model manually for individual runs, there is a batch file which can set the model to run automatically over multiple simulations. The batch file is currently set up to explore the parameter space in Holocene and LGM experiments. To run the model in this way, first turn "BatchRun" on and "Charting" off in the SCPM_model.py model, as described above. Then, in "SCPM_batch.py" select the experiment you wish to run – currently configured for "LGM" or "Hol". If you are running "Hol", this will run the model over the parameter values is shown in the Holocene code from line 40. Once you have selected the above options, execute the "SCPM_batch.py" scrip and you are away. Note, depending on how long you have SCPM_model.py configured to run (in years), the experiment could take quite a long time – hours, days. Best to run this on a dedicated PC or when you don't need Python.

If you are choosing the "LGM" scenario, it's a good idea to enable "LGM" in SCPM_parameters.py, as discussed above. This will enable a host of 'LGM-style' changes in the model's settings for sea surface temperature, salinity, piston velocity, atmospheric radiocarbon production, and ocean surface area/volume.

At the end of the experiment, the script will aggregate model output into a master output file which will be deposited in Results/Batch_Master/. Note the output from our LGM and Holocene experiments are provided in the version on Zenodo.

Figure 7: SCPM_batch.py module for running batch simulations with SCP-M

```python
28
29 BatchResultsDir='Results/Batch/' # Results from batch run
30 MasterOutDir='Results/Batch_Master/' # Master batch file output
31
32
33 ##---CONTROLS-------------------------------------------------------
34
35 Experiment='None'
36 # Options ['Hol','LGM','None' (turns batch module off- leave it like this when not using,
37 # because you don't want to trigger a batch run when using other modules!)]
38
39 ##-----------------------------------------------------------------
40
41 # Holocene only experiment - 4 parameters
42 # Really 5 parameters in the experiment but gamma2 is not varied
43
44 if Experiment=='Hol':
45
46     # Holocene experiment - parameter ranges
47     Psi1ex=[20.0e6,21.0e6,22.0e6,23.0e6,24.0e6,25.0e6,26.0e6,27.0e6,28.0e6,29.0e6,30.0e6,31.0
48     Psi2ex=[15.0e6,16.0e6,17.0e6,18.0e6,19.0e6,20.0e6,21.0e6,22.0e6,23.0e6,24.0e6,25.0e6]
49     gamma1ex=[15.0e6,16.0e6,17.0e6,18.0e6,19.0e6,20.0e6,21.0e6,22.0e6,23.0e6,24.0e6,25.0e6,26
50     gamma2ex=[40.0e6]
51     Zedex=[2.0,3.0,4.0,5.0,6.0,7.0]
52
53     master='_Hol_4'
54     nh = 0 # Filename counter
55     for Psi1 in Psi1ex:
56         for Psi2 in Psi2ex:
57             for gamma1 in gamma1ex:
58                 for gamma2 in gamma2ex:
```

BoxDataMap.py

If you are running the model in batch mode as per above, there's a good chance that you will want to compare the model results with some data. In the SCP-M model documentation, we performed a model-data experiment on the LGM and late Holocene paleo proxy data to tell us what the right model inputs were each period. That needed some data. This module maps and averages (plus standard deviation) the ocean paleo proxy data for d13C, D14C and carbonate ion, based on the SCP-M box latitude and depth coordinates and for LGM and late Holocene time slices. The script reads in data in three spreadsheets (one for each of d13C, D14C and carbonate). Figure 8 shows the d13C core data from Peterson et al (2014) used in the experiment. In this case, the data are already averaged into LGM and late Holocene time slices, therefore the script only has to perform average and standard deviation calculations based on latitude and depth. To run the script, execute it with the data files in the Data_files/ location. Basically, its already set up to run on the LGM and Holocene proxy data. If you want to add more data to the workbooks, simply add data according to the column headings in the spreadsheet. As shown in Figure 9, there are no user controls needed to run the script, it should work as is on the data as located.

Figure 8: StableCarbon_Read.xls data input file for LGM and Holocene ocean d13C data

	A	B	C	D	E	F	G
1	Core_Ref_2	Lat	Long	Sample_Depth	Hol_d13C	LGM_d13C	References
2	KNR159-10GGC	-26.48	-45.93	630	1.6000	1.0000	Curry & Oppo, 2005
3	OCE205-106GGC	25.98	-78.18	654	1.3000	1.9000	Curry & Oppo, 2005; Slowey & Curry, 1995
4	OCE205-108GGC	25.98	-78.18	743	1.2000	1.8000	Curry & Oppo, 2005; Slowey & Curry, 1995
5	GEOB6728-1_ISO	52.15	-12.77	749	0.986	1.3171	Dorschel et al., 2005
6	GEOB6719-1	52.15	-12.77	758	NaN	1.3650	Rüggeberg et al., 2005
7	GeoB3104	-3.67	-37.72	767	1.4300	0.9000	Bickert & Mackensen, 2004; Arz et al., 1999
8	KNR140-68GGC	32.94	-76.55	775	0.7000	NaN	Curry & Oppo, 2005; Keigwin, 2004
9	OCE205-33GGC	26.22	-77.67	783	1.2000	1.7000	Curry & Oppo, 2005; Slowey & Curry, 1995
10	KNR159-99GGC	-27.37	-46.84	790	1.5000	0.9000	Curry & Oppo, 2005
11	GIK16006-1	29.25	-11.50	796	1.2000	1.4800	Sarnthein et al., 1994; Hoogakker et al., 2010
12	GIK15666-6	34.96	-7.12	803	1.2700	1.7300	Sarnthein et al., 1994; Hoogakker et al., 2010
13	GIK16017	21.25	-17.80	812	1.0018	1.2715	Sarnthein et al., 1994
14	V16-51	-33.47	-17.03	898	0.6100	0.9000	Lynch-Stieglitz et al., 2006
15	KNR159-153JPC	-26.39	-45.69	898	NaN	0.6000	Curry & Oppo, 2005
16	GeoB6718-2	52.16	-12.75	900	0.9911	1.3293	Rüggeberg et al., 2007
17	KNR140-63JPC	32.99	-76.41	900	1.0000	1.5000	Curry & Oppo, 2005; Keigwin, 2004
18	KNR159-38JPC	-27.27	-46.63	936	1.4000	0.7000	Curry & Oppo, 2005
19	KNR159-37GGC	-27.25	-46.63	938	1.4000	0.8000	Curry & Oppo, 2005
20	GIK12345-5	15.48	-17.36	945	0.7500	0.9200	Sarnthein et al., 1994
21	V19-258	-20.40	11.62	965	0.5540	0.3300	Lynch-Stieglitz et al., 2006
22	OCE205-103GGC	26.07	-78.07	965	1.4215	1.4645	Curry & Oppo, 2005
23	MG-237	-5.20	11.33	1000	-0.1100	-0.1000	Sarnthein et al., 1994; Lynch-Stieglitz et al., 2006
24	BT4	-4.00	10.00	1000	NaN	0.1167	Sarnthein et al., 1994
25	KNR140-61GGC	33.00	-76.36	1005	1.2000	NaN	Curry & Oppo, 2005; Keigwin, 2004
26	GIK15627-3	29.17	-12.09	1024	NaN	1.3700	Sarnthein et al., 1994
27	OCE205-100GGC	26.06	-78.03	1057	1.5000	1.5000	Curry & Oppo, 2005; Slowey & Curry, 1995
28	KNR159-90GGC	-27.35	-46.63	1105	1.3000	0.6000	Curry & Oppo, 2005
29	KNR159-105JPC	-27.35	-46.63	1108	1.1000	0.4000	Curry & Oppo, 2005
30	KNR159-73GGC	-27.89	-46.04	1128	1.3000	0.5000	Curry & Oppo, 2005
31	ODP982	57.51	-15.85	1134	0.9611	1.4350	Jansen et al., 1996; Venz et al., 1999, 2002
32	V29-198	58.73	-15.57	1139	1.0700	1.5400	Oppo & Lehman, 1993

Figure 9: BoxDataMap.py module for mapping paleo data into SCP-M boxes for the model-data experiments

```python
20
21
22 # Read in data
23 # For data sources see model documentation
24 Radiocarbon=pd.read_excel(DataDir+'Radiocarbon_Read.xlsx')
25 StableCarbon=pd.read_excel(DataDir+'StableCarbon_Read.xlsx')
26 Carbonate=pd.read_excel(DataDir+'Carbonate_Read.xlsx')
27
28 # Age slices
29 LGMAge_Rad=(Radiocarbon.Age<23000) & (Radiocarbon.Age>19000)
30 HolAge_Rad=(Radiocarbon.Age<6000) & (Radiocarbon.Age>200)
31
32 LGMAge_Carb=(Carbonate.Age<23000) & (Carbonate.Age>19000)
33 HolAge_Carb=(Carbonate.Age<6000) & (Carbonate.Age>200)
34 # Note Peterson et al (2014 d13C data already binned into LGM and late Holocene
35 # slices)
36
37
38 #Latitude of boxes Talley (2013) interpretation (no longitude as this is zonally averaged)
39 Box1LatMin=-40.0
40 Box1LatMax=40.0
41 Box2LatMin=40.0
42 Box2LatMax=80.0 # extended for data coverage
43 Box3LatMin=-40.0
44 Box3LatMax=40.0
45 Box4LatMin=-61.0
46 Box4LatMax=60.0
47 Box5LatMin=-80.0
48 Box5LatMax=-61.0
49 Box6LatMin=-80.0
50 Box6LatMax=80.0
51 Box7LatMin=-61.0
52 Box7LatMax=-40.0
53
54 # Depth of boxes
55 Box1DepthMin=0.0 #Box0
56 Box1DepthMax=100.0 #Box0
57 Box2DepthMin=0.0 #Box1
58 Box2DepthMax=250.0 #Box1
```

ModelData.py

After the previous two steps (SCPM_batch.py and BoxDataMap.py), you should have a batch master file that contains the results of your batch simulations, in Results/Batch_Master/ e.g. the "Output_all_Hol.txt" and "Output_all_LGM.txt" batch master output files from our experiments are provided, and a set of data (that you wish to compare the model with) of average element compositions (and standard deviation) in SCP-M boxes, as you will generate from the "BoxDataMap.py" module using the data spreadsheets provided for LGM and Holocene proxy data. Step (7) is about finding the best model run for your data. In our case, this was for LGM and late Holocene data. The "ModelData.py" script takes the output from (5) and (6) and solves for the parameters of best fit for the data, via a least-squares model results-data residual minimisation calculation. Basically, it searches through the model results that you generated from your batch runs and then finds the simulation that gave results closest to the data values, and tells you what parameters were inputted to deliver those results. Can be quite useful. This module is currently set up to solve for parameters of best fit for the LGM and late Holocene ocean and atmosphere proxy data. The output file is deposited in Results/ModelData_Outputs/.

Data files provided with the model

GLODAPv2 data

\Data_files\GLODAPv2.txt
This is the GLODAPv2 database in full, incorporating ~1m samples taken between 1972-2013. The data is accessed from:
https://www.nodc.noaa.gov/ocads/oceans/GLODAPv2/
Click on "Data products"
https://www.nodc.noaa.gov/archive/arc0107/0162565/2.2/data/0-data/
And then: data_product/
https://www.nodc.noaa.gov/archive/arc0107/0162565/2.2/data/0-data/data_product/
We used the GLODAPv2 Merged Master File, which is a combination of Pacific, Atlantic and Indian Ocean databases, located at:
GLODAPv2 Merged Master File.csv.zip

Late Holocene and LGM ocean data files

Radiocarbon_Read.xlsx
This file incorporates the ocean radiocarbon time series data gathered for the LGM and Holocene analysis. Information includes: core reference, sample location, depth, absolute calendar year age, D14C, age uncertainty. Data is sourced from the following:

Ahagon et al., 2003, Barker et al (2010), Broecker et al., 1990, Broecker et al., 2004, Broecker et al., 2007, Bryan et al 2010, Burke and Robinson (2012), Chen et al (2015), Cook and Keigwin (2015), Davies-Walczak et al (2014), Duplessey et al., 1989, Fairbanks et al (2005), Freeman et al (2015), Galbraith et al., 2007, Gorbarenko et al., 2005, Hines et al (2015), Ikehara et al., 2011, Ingram and Kennett, 1995, Keigwin and Jones, 1990, Keigwin, 2002, Keigwin and Swift (2017), Lindsay et al (2016), Lund et al., 2011, Marchitto et al (2007), McKay et al., 2005, Minoshima et al., 2007, Mix et al., 1999, Murayama et al., 1992,

Rae et al., 2014, Ronge et al (2016), Sagawa and Ikehara, 2008, Sarnthein et al., 2007,, Schlung et al., 2012, Shackleton et al., 1988, Sikes et al (2015), Skinner et al (2010), Skinner et al (2014), Skinner et al (2015), Thornalley et al (2011), Thornalley et al (2015), van Geen et al., 2003, VanGeen et al., 1996

StableCarbon_Read.xlsx

This file includes d13C data used for the LGM and Holocene analysis. The data file includes core reference, location, depth, LGM and late Holocene d13C values. The data is sourced from Peterson et al (2014), which is previously averaged into LGM and late Holocene values, hence there is no time series data.

Carbonate_Read.xlsx

This file contains carbonate proxy time series data used for the LGM and Holocene analysis. The data file includes core reference, location, depth and carbonate proxy values. The data is sourced from Yu et al (2013, 2014a, 2014b).

Modern data files

Modern_AtCO2Dat.txt

This file contains historical atmospheric CO2 data (ppm) from 1752 until 2016, and the IPCC's representative concentration pathways (RCP) from 2016 until 2100.

Historical atmospheric CO_2 data is sourced from the Scripps CO_2 Program data website. At the time of writing this data is still available, but understood to being transferred to another site following the closure of the Scripps $_{CO2}$ Program.
Data was sourced here:
http://scrippsco2.ucsd.edu/data/atmospheric_co2/icecore_merged_products
csv file download:
http://scrippsco2.ucsd.edu/assets/data/atmospheric/merged_ice_core_mlo_spo/merged_ice_core_yearly.csv

rcp_AtCO2.txt
The RCP atmospheric CO_2 projections are sourced from the IPCC Assessment Report 5 (AR5), Climate Change 2013, the Physical Science Basis, Working Group I, Annex II – data downloads:
http://www.climatechange2013.org/report/full-report/
A spreadsheet containing data tables from the Physical Science Basis report, was downloaded. RCP atmospheric CO_2 projections are sourced from the "All4-1" worksheet. These are provided at ten year intervals. These were converted to an annual time series using a linear interpolation between 10-yearly data points, and transferred to the .txt file.

Hist_AtCO2Ems.txt
This file contains historical estimates/data for human fossil fuel and land use change emissions for the years 1751-2014. The fossil fuel emissions historical data 1751-2014 is sourced from (Boden et al, 2017):
http://cdiac.ess-dive.lbl.gov/trends/emis/tre_glob_2014.html

With csv file downloaded from:
http://cdiac.ess-dive.lbl.gov/ftp/ndp030/CSV-FILES/global.1751_2014.csv

Land use change emissions data from 1850-2005 is sourced from Houghton (2010). Data for 1751-1850 is a linear interpolation from a nominal value in 1751 to the Houghton (2010) estimate for 1850. Input values between 2005 and 2010 are a linear interpolation between the Houghton (2010) value for 2005 and the IPCC RCP estimates for 2010.

rcp_FFems.txt
This file contains projections of human fossil fuel combustion emissions from 2020 to 2100 under the IPCC's RCP scenarios. Projections are sourced from the IPCC's RCP projections from the IPCC Assessment Report 5 data tables as described above ("All-2-1abc" worksheet)

rcp_LUCems.txt
This file contains projections of human land use change emissions from 2010 to 2100. Values are sourced from the IPCC's RCP projections from the IPCC Assessment Report 5 data tables as described above ("All-2-1abc" worksheet).

rcp_SST.txt
This file contains estimates of sea surface temperature from the IPCC Assessment Report 5 data tables as described above ("All-7-6" worksheet). These estimates are provided as a delta from 1990 measurements. The data includes an estimate of the preindustrial delta. We have provided a linear interpolation between the preindustrial estimate for the period 1751-1990, and then apply the IPCC projections out to 2100 under the RCP scenarios.

Modern_Atd13CDat.txt
This file contains atmospheric d13C data for the modern era spanning from 1751 to 2016. The time series is a combination of measurements from two sources. Data for 1751-1979 is sourced from Rubino et al (2013), and data for 1979-2016 sourced from Keeling/Scripps CO2 Program

Modern_AtD14CDat.txt
This file contains atmospheric D14C data for the modern era spanning from 1751 to 2015. The time series is a combination of measurements from two sources.
1954-2015 Turnbull (2017) Wellington NZ data
https://www.niwa.co.nz/atmosphere/our-data/trace-gas-plots/carbon-dioxide
1751-1953: Stuiver et al (1998)
http://www.radiocarbon.org/IntCal04.htm
Dataset 1: University of Washington (QL) tree-ring