Charles Williams
Last updated: 13/8/20

IDIOT'S GUIDE TO SETTING UP GC3.1 EOCENE SUITE ON NEXCS

C. Williams, A. Sellar & W. Roberts

INTRODUCTION

The purpose of this dummy or idiot's guide is to bring together all correspondence/notes on setting up the fully-coupled Eocene suite on NEXCS, including creating/modifying all relevant Eocene boundary & initial conditions, then using Rose/Cylc to run suites. This was written by and is aimed at the ultimate dummy, Charlie Williams, with major input from Alistair Sellar and Will Roberts, who actually know what they are doing. Very many thanks to NCAS-CMS (especially Ros H, Grenville L, Simon W and Jeff C) for providing invaluable additional advice.

Nomenclature of this guide

- Text (black) in Times New Roman = instructions
- Text (blue/underlined) in Times New Roman = websites
- Text (grey) in Times New Roman = pathnames/filenames
- Text (red) in Times New Roman = error messages
- Text (black) in Arial = commands (to use at command prompt) or code (to be written in scripts)
- <TEXT_IN_CAPITALS_INSIDE_ARROWS> = needs to be replaced with own individual username, ID, etc

Useful websites

- Logging into Monsoon: http://collab.metoffice.gov.uk/twiki/bin/view/Support/MONSooNIntroPack
- Getting Rose jobs running on Monsoon: https://collab.metoffice.gov.uk/twiki/bin/view/Support/MONSooNRose
- Setting up on NEXCS: http://collab.metoffice.gov.uk/twiki/bin/view/Support/NEXCS (and links therein)
- Checking changes from suite: https://code.metoffice.gov.uk/trac/roses-u/log/

Charles Williams
Last updated: 13/8/20

INITIAL SETUP AND LOGIN DETAILS

The following assumes that all usernames have been created and registered (e.g. access to NEXCS. If not, need to email monsoon@metoffice.gov.uk to request an application form - and specify access to nexcs-n02. Won't need a Met Office collaborator for this, but will need one if access to MASS is wanted (either write access for output, or read access from JASMIN - see below). Once registered, electronic key fob will be sent in post and will be needed to initially logging into landing stage. Follow instructions on website above to log into Monsoon for first time.

If no access to the repository - MOSRS (as well as appropriate groups e.g. UM) - see http://cms.ncas.ac.uk/wiki/UmTraining.

The following also assumes that space is available elsewhere e.g. local server or, preferably, on JASMIN. This space, probably at least 50G, will be needed to create/modify ancillaries. If not using JASMIN, need to make sure all appropriate software is installed i.e. IDL v8.5 (or later), Python v2.7 (or later), CDO, NCO, xconv/convsh, xancil etc. All of these are available on JASMIN.

Special word about xconv: On NEXCS, standard version of xconv (opened with simply xconv) lives at /home/d00/jecole/bin/xconv. However, this cannot open HDF5 files (or if it does, the image will not be displayed correctly) so need to use the most up-to-date version, which lives at /projects/um1/bin/xconv.

If wanting to send output directly to JASMIN, need to request an account on NEXCS group workspace on JASMIN: https://accounts.jasmin.ac.uk/services/group_workspaces/. Likewise, if wanting to access data/programs discussed throughout this guide, need to request access to NCAS-Climate group workspace on JASMIN, also at above link. This will enable you to see all scripts/data discussed in this guide, in group workspace (at /gws/nopw/j04/ncas_climate_vol2/users/cwilliams2011/data1_for_use.d/sweet.d/gc31.d) or home directory (at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/ocean.d/bathy.d). In order to transfer data from NEXCS to JASMIN, need to request an account for high performance data transfer, at http://www.jasmin.ac.uk/services/high-

performance-data-transfer/. Don't need to worry about supplying an IP address because NEXCS is whitelisted.

Once everything has been requested and granted, need to create directories (for archiving of output) on both NEXCS (at /projects/nexcs-n02/<USERNAME>) and the JASMIN NEXCS group workspace (at /group_workspaces/jasmin2/nexcs/<USERNAME>). Directories should begin with <USERNAME> then whatever.

Notes on differences between Monsoon & NEXCS

NEXCS is exactly the same as Monsoon - they are same machine. Science node of NEXCS (xcs-s) is both Monsoon and NEXCS (NEXCS is implemented as a Monsoon project). When running on NEXCS, just specify a different project code to run under (e.g. nexcs-n02 rather than clpredic), everything else is the same. So effectively NEXCS is just a project on Monsoon. Therefore, when running suites, simply use the nexcs-n02 code within the Rose editor.

When initially logging in (using key fob), won't be able to do anything here because landing stage only, so need to straightaway login to either XCS0 or 1: ssh -Y -A xcslc0 or 1

Setting up shortcuts on NEXCS

All alias commands need to be in .bash_profile, because system will source the .bash_profile first (.profile is sourced only if this doesn't exist). These shortcuts will only work on XCS, not landing stage. Recommended one: alias xconv2=/projects/um1/bin/xconv.

Setting up password-less access from NEXCS to JASMIN

1) Copy jasmin ssh key, usually id_rsa_jasmin, from existing .ssh (in home directory) to ~/.ssh directory on NEXCS
2) On NEXCS do:

```
cd ~/.ssh
<JASMIN-KEY> <JASMIN-KEY>.orig
openssl rsa -in <JASMIN-KEY> -out new.key
mv new.key <JASMIN-KEY>
```

Might need to close down permissions, so that it is only read/write for user i.e.

chmod 600 <JASMIN-KEY>

3) Modify (or create if it doesn't exist) ~/.ssh/config on XCS to add:

```
Host jasmin-xfer1 jasmin-xfer1.ceda.ac.uk
    Hostname jasmin-xfer1.ceda.ac.uk
    User <JASMIN-USERNAME>
    IdentityFile ~/.ssh/<JASMIN-KEY>

Host jasmin-xfer2 jasmin-xfer2.ceda.ac.uk
    Hostname jasmin-xfer2.ceda.ac.uk
    User <JASMIN-USERNAME>
    IdentityFile ~/.ssh/<JASMIN-KEY>

Host hpxfer1 hpxfer1.jasmin.ac.uk
    Hostname hpxfer1.jasmin.ac.uk
    User <JASMIN-USERNAME>
    IdentityFile ~/.ssh/<JASMIN-KEY>
```

Note: Since writing version 1, jasmin-xfer1 has been retired (outside Met Office) and jasmin-xfer2 will be retired imminently. Therefore, only really need to include last section.

4) Test that it works i.e. that it's possible login to JASMIN from NEXCS with no prompting for passphrase/password: ssh jasmin-xfer1.ceda.ac.uk

Running rose

If, when trying to run rosie go, error message appears, might be missing following lines within .bash_profile so insert if so:

```
# Provide access to FCM, Rose and Cylc
PATH=$PATH:~fcm/bin
```

If unable to access MOSRS

If, when logging in to XCS, it does not ask for MOSRS password (needed to run rosie go and access suites), quite likely that there is still an ssh process still running - if there is already a connection, albeit hidden, it won't ask for password. So firstly check if another terminal is open and already connected. If not, check to see what ssh processes are running,

using ps -flu <USERNAME> | grep ssh, then kill -9 <JOBID> for any ssh connections. If link has broken (e.g. if it asks for password when trying to load Rose) and don't want to logout completely, run mosrs-cache-password and retype password.

Charles Williams
Last updated: 13/8/20

NOTES

Note: every single suite is different, with different things being defined in different windows/files. This is particularly the case for early suites, which were set up differently as design hadn't settled at that time. Newer suites resolve this and are being set so that they are site portable (e.g. flicking a switch to get all the platform specific settings). To do this, on NEXCS the /home/d05/<USERNAME>/roses/<SUITEID>/suite.rc file is now split to include the correct settings for the platform being used. Therefore instructions below will probably be somewhat irrelevant for another, older, suite.

Part 1: Obtaining & setting up a Rose suite to run GC3.1 on NEXUS: PI control

Firstly, need to obtain a suite that is set-up to run on this machine. Given that suite needs to be as close as possible to kosher CMIP6 PI control suite (other than known modifications), best to find this: u-ar766@86126

On XCS, run rosie go. Search for above suite and revision, then right click and select Copy. This both copies and checks out the suite, making it editable under home directory. Example now: u-be194 (hereafter "PI suite").

Probably good idea to get this working i.e. to make sure suite will run with all standard boundary/initial conditions, before making it Eocene.

Things to change within Rose (either open GUI using rose edit or do everything in individual files[1]):

1) Project code (at suite conf > Project accounting): set Use default account = True, Sub-project name = Other, Other subproject name = nexus-n02
2) Host (at suite conf > Machine options): don't want any of these, instead set HOST_XC40='xcs-c' in /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf

[1] Note: if things are being changed in files themselves e.g. rose-suite conf, need to make sure that Rose GUI is closed. Otherwise, if it is already open, changes made in files will revert back to previous version. Likewise, if using GUI, need to make sure file is closed.

3) Queue (at suite conf > Machine options): if wanting to use long24 queue (which allows up to 24 hours of wall clock time), HPC queue can't be set here. Instead, set this in HPC_QUEUE in /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf. NEXCS has 3 queues: normal (4 hours = 14000 seconds), long12 (12 hours = 43200 seconds) & long24 (24 hours = 86400 seconds), see https://collab.metoffice.gov.uk/twiki/bin/view/Support/NEXCS.
4) Resources for fcm building: these are currently set to Met Office settings, so need to change from 'slurm' to 'background' in /home/d05/<USERNAME>/roses/<SUITEID>/site/meto_cray.rc under [[EXTRACT_RESOURCE]] section, so:

```
{{{
  [[EXTRACT_RESOURCE]]
    [[[job]]]
      batch system = background
      execution time limit = PT5M
    [[[directives]]]
      --mem=1G
      --ntasks=1
}}}
```

5) Processes/nodes (at suite conf > Domain decomposition > Atmosphere): for testing purposes, can leave this as is - but will need to change this when converting to Eocene (see below)
6) Building (at suite conf > Build and Run): all need to be turned on. Can turn off Build UM/Drivers/Ocean when it has been run once, but will need to turn back on when suite is copied and/or if suite is cleaned using rose suite-clean. Only turn on Post Processing if archiving has been included. If running for first time, turn on Run Reconfiguration to reconfigure atmosphere start dump
7) Run length & cycling (at suite conf > Run initialisation and cycling): again, can leave as is for testing purposes e.g. 1 year run, 6 months cycling, 5 hour wallclock time. Also specified in /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf. But will need to change when converting to Eocene (see below)
8) Restart dumps: need to set atmospheric reconfigured start dump (at um > namelist > Reconfiguration and ancillary control > General technical options, or just search for "ainitial" within um app) and sea ice start dump (at nemo_cice > Restart files > CICE restart file) to appropriate locations, because default options may not exist. These are

currently at /home/d05/cwilliams/gc31/restarts/aq853a.da25000101_00 and /home/d05/cwilliams/gc31/restarts/u-aj572i.restart.2000-01-01-00000_cjrw.nc respectively. Can leave non-reconfigured start dump (at um > namelist > Model input and output > Dumping and meaning, or just search for "astart" within um app) as is, because this is not used (overwritten by ainitial). Also need to remove both NEMO restart files (at nemo_cice > Restart files), so that they begin from climatology

9) Run! 2 ways of doing this: either press Play button within GUI or, at command line within working directory (i.e. /home/d05/<USERNAME>/roses/<SUITEID> do rose suite-run

It's possible that, since time of writing, certain boundary/initial conditions may have been moved. Error messages are usually fairly intuitive about this. One particular culprit = sea ice (set in Eocene suite GUI at nemo_cice > Restart files > CICE restart file), which was removed after PI control was run. Currently set to /projects/ukesm/jwalton/startdumps/cice/u-aj572i.restart.2000-01-01-00000.nc so needs be copied to appropriate directory. Also possible that various branches or their revisions (set in fcm_make_um > Sources) have become outdated - if no revision is specified, then latest revision is used (bad practice, because suite could be run at two different times and get different code if the branch was updated in between!). Again, error messages are usually intuitive if branch is missing. If this happens, need to find original revision of this branch within repository, and add this to the end (see FIXING PROBLEMS).

Assuming it works, need to commit suite to repository so it is permanently saved with a revision number: within working directory, do fcm commit.

Once committed, shut down Rosie go, then reopen: rosie go. Find above suite (containing its new revision), then right click and select Copy. Again, this both copies and checks out the suite, making it editable under home directory. Example now: u-bh301 (hereafter "Eocene suite"). Begin Eocene process.

Part 2a: Setting up suite to become Eocene - Ancillary/boundary condition creation

Firstly need to create ancillary files using land sea mask from bathymetry (because otherwise ocean land sea-mask will be different to atmosphere land-sea mask). Multistep process:

1) Create bathymetry ancillary
2) Use this to run ocean-only suite (hereafter "NEMO suite"), to create mesh mask files
3) Send these to A. Sellar, who will use these to create coupling weights (currently this can only be done with Met Office firewall)
4) Create rivers ancillary
5) Create topography ancillary
6) Use these (rivers & topography) to run Rose ancillary suite (hereafter "ancil suite"), to create all other ancillaries using correct ocean or atmosphere land-sea mask

In detail…

1) Create bathymetry ancillary - use proxy bathymetry created by Herold *et al*. (2014). All Herold data are on NEXCS at /home/d05/cwilliams/gc31/final_ancils/herold.d/ or JASMIN at /gws/nopw/j04/ncas_climate_vol2/users/cwilliams2011/data1_for_use.d/sweet.d/gc31.d/build.d/ancils.d/herold.d. For bathymetry, need to use: herold_etal_eocene_topo_1x1.nc (which includes both topography and bathymetry). Bathymetry will need to: be extracted from file, have positive (i.e. topography) zeroed, have negative values (i.e. bathymetry) reversed in sign (because NEMO expects positive values as bathymetry), be regridded to NEMO resolution, be chiselled to open certain gateways, have various attributes added, and be written out to new file in same format that NEMO expects. Full instructions, including code to do all of above, can be seen on JASMIN at /gws/nopw/j04/ncas_climate_vol2/users/cwilliams2011/data1_for_use.d/sweet.d/gc31.d/build.d/ancils.d/bathy.d/readme.txt. Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/nemoimput.d and copied to appropriate name e.g. cp orca_all_chiselled_new_round2.nc eORCA_R1_bathy_meter_v2.0_eocene_chiselled.nc
   Note: this final step isn't actually necessary, because file can be called whatever as is pointed to within ocean ancillary versions file.
2) Use this to run NEMO suite, to create mesh mask files - once bathymetry is ready (final version is on NEXCS at /home/d05/cwilliams/gc31/nemoimput.d/eORCA_R1_bathy_meter_v2.0_eocene_chiselled.nc), find and copy NEMO suite: u-ba251@87020 (using rosie go, as above) to generate a new suite ID. Need to copy ocean ancillary versions file (currently on NEXCS at /home/d05/

cwilliams/gc31/ancil_versions/ocean_ancil_GO6_GSI8_CORE_forcing_eORCA1_u-ba251_cjrw) to new version, then change BATHY_METER= to point to new version of bathymetry. Then make sure new NEMO suite points to new ancillary versions file (in suite conf > Input data). Once saved, run the suite - it will build, but then fail at first timestep. This doesn't matter, as long as it has created mesh mask files on NEXCS at ~/cylc-run/<SUITEID>/work/19760101T0000Z/nemo_cice/mesh_mask*. There will be 192 of these. Need to rebuild these, by doing ~hadtq/bin/local_rebuild/rebuild_nemo_login mesh_mask 192, which will combine all of them into mesh_mask.nc

Send this to A. Sellar, who will use these to generate coupling weights (currently this can only be done with Met Office firewall). Once created, transfer the entire weights directory from A. Sellar to a suitable place on NEXCS e.g. /home/d05/cwilliams/gc31/coupling_weights_round2, then point the Eocene suite to this directory using RMP_DIR= in ~/roses/<SUITEID>/app/coupled/rose-app.conf. In addition, in the same directory will be the area fractions file, atmo_mask_fracarea_anc_ns, which needs to be included in the ancillary suite (see below)

3) Create rivers ancillary - again use proxy data created by Herold *et al*. (2014), at above directories. Need to use: herold_etal_eocene_runoff_1x1.nc (which contains direction field). Given that model requires 3 fields (direction, sequence and storage), and Herold only has direction, need to create new sequence field from this. Don't need to worry about storage, because that can be initialised to zero and will spin-up quickly
    - To do this, within Eocene suite GUI go to um > namelist > Reconfiguration and ancillary control > Configure ancills, then find stash_req 153 (easiest way of finding this is to order all items by stash request), then open up this item and change Source to "Set to 0" as well as setting ancilfilename= to blank (because there is now no longer a file)

So just need to create river sequence. Full instructions, including code, can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/rivers.d/readme.txt. Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31.d/ancilinput.d/rivers_build. When writing out new river sequence, be very careful to make sure that it is not upsidedown, the longitudes are correct and it is not displaced

Note: This step does not need to be repeated each time new bathymetry is created, because this is a filled field, when mask only being imposed onto the output of the ancillary suite (see below).

4) Create topography ancillary using Will's code, then transfer to NEXCS, e.g. at /home/d05/cwilliams/gc31.d/ancilinput.d/orog
5) Use 3 files (rivers, topography & area fractions file) to run ancil suite, to create all other ancillaries using correct land-sea mask - find and copy ancil suite: u-ar826@65219 (using rosie go, as above) to generate a new suite ID. Within GUI of new suite, set:
    - ancil_river_routing_preproc > env > source so that it points to new rivers file
    - ancilOrog > env > Orography sources > orog_source_data so that it points to new topography file
    - ancilMask > env > FRACIN so that it points to new area fractions file

   Run ancil suite, which will generate all required ancillary files (containing modern data but Eocene mask) at /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data/n96e_orca1_go6/ so copy these to another appropriate directory (don't need .xml files). Suite will fail when it tries to make river storage, and this won't be written out to above directory, but doesn't matter because don't need it

Part 2b: Setting up suite to become Eocene - Ancillary/boundary condition paleaotising

Once ancillaries have been created, need to go through many of them (not all) and modify i.e. paleaotise. Doesn't matter in which order these are done. 2 types of paleaotising:

- Global paleaotising = take global mean of modern field and apply to every grid box of new field, using Eocene mask
- Zonal paleaotising = take zonal mean of modern field, averaged separately over land and sea, and apply to every grid box of new field, using Eocene mask

Firstly, need to make a copy of existing atmosphere and ocean ancillary versions file e.g. at /home/d05/cwilliams/gc31/ancil_versions/GC31_UM107_deepmip_atmos_cjrw_round2 and GC31_eORCA1v22_deepmip_ocean_cjrw_round2 respectively. All atmosphere and ocean boundary conditions (except aerosol emissions/oxidants) need to be pointed to in this file.

Aerosol emissions and oxidants need to be pointed to in /home/d05/<USERNAME>/roses/<SUITEID>/app/um/rose-app.conf or within Eocene suite GUI: um > namelist > UM Science settings > Section 34: UKCA: UK aerosols and chemistry. For both aerosol emissions and oxidants, either need to set directory pathname and individual filenames, or leave directory pathname blank but then include full path with filenames e.g.

```
ukca_offline_dir=''
ukca_offline_files='/home/d05/<USERNAME>/ukca_oxid_clim_H2O2.nc'
```

or

```
ukca_offline_dir=' /home/d05/<USERNAME>'
ukca_offline_files='ukca_oxid_clim_H2O2.nc'
```

Once ancil versions files are ready, point to them in Eocene suite GUI: install_ancil > file.

Once ancillaries have been created/modified and transferred to NEXCS, some (e.g. aerosols emissions) can remain as netcdf whereas others (e.g. vegetation) need to be converted to .pp using xancil. Some of these can be converted using xancil's built-in options, whereas others need to use "Generalised ancillary fields". See below for individual ancillaries.
Note: If using previous xancil jobs (e.g. one that has been saved from previously), be VERY careful about pathnames. For unknown reasons, even if pathname is changed to most recent version of file, xancil reverts back to previous job when it is run. Therefore, unless absolutely certain that pathname is correct, better not to use previously saved jobs.

Atmospheric boundary conditions (aerosol emissions)

Natural emissions (SO2, DMS & monoterpene), biomass burning & anthropogenic emissions

- Need to take modern (for natural emissions) & PI control (for biomass burning) monthly climatologies (single level), and zonally palaeotise at every month. For anthropogenic emissions, need to be zeroed. Full instructions, including code, can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/aerosols.d/readme.txt and NEXCS at /home/d05/cwilliams/gc31/final_ancils_round2/aerosols/readme.txt (multistep process using both machines). Final versions need to be transferred to NEXCS, e.g. at above directory. Don't need to use xancil here because model expects netcdf (unusually!) - Exception

to this is the seawater concentration (qrclim.sulpdms.nc) which does need to be converted into UM format, using xancil's "Generalised ancillary option"

Note: There are two DMS[2] ancils: ukca_emiss_DMS.nc is a direct emission of DMS gas from land, while qrclim.sulpdms provides a seawater concentration which the model takes as input to an emission parameterisation which uses the model wind speed (means that changes in wind due to climate change can affect DMS emissions even though the ancil is a fixed climatology). Both of these need to be zonally palaeotised.

Atmospheric boundary conditions (other emissions)

Ozone & offline oxidants

For ozone, need to take UKESM1 4xCO_2 monthly climatology (85 levels), and zonally palaeotise at every month/level[3]. Code can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/ozone.d/make1a_ozone.pro and make1b_ozone.py which need to be run in that order (IDL does palaeotising, then python inserts data into original file thus maintaining all attributes/meta data). Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/ozone, then use xancil to convert to .pp

Note: Here, input file won't contain correct attributes and so will give an error in xancil - to get round this, use separate file containing vertical values, currently at /home/d05/cwilliams/gc31/final_ancils_round2/ozone/sparc/1994-2005/vertlevs_L85_50t_35s_85km. Then, in xancil, before importing ozone, go to Grid configuration > Specify atmosphere vertical levels and read in this file, setting both model and ozone levels to 85 and "Is level data stored upwards" to yes. Then create ozone as usual. More details at:

http://cms.ncas.ac.uk/documents/xancil/index.html#document-config

Standard CMIP6 suite uses 3D ozone, but here it's a zonal mean, so need to set zon_av_ozone=.true. (in /home/d05/<USERNAME>/roses/<SUITEID>/app/um/rose-app.conf)

[2] Causal pathway of DMS: DMS water concentration > DMS emissions > DMS air concentration > oxidises to SO2 > oxidises to sulphate aerosol

[3] If there is zonal structure aloft, it is likely to arise from an orographic footprint on the dynamics, e.g. through gravity wave drag. It is cleaner to take the mean at all levels, and also more consistent with the chemical oxidants.

- For offline oxidants[4], need to take UKESM1 PI control monthly climatology (85 levels), and zonally palaeotise at every month/level. Code can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/oxidants.d. Again, run IDL scripts before python. Final versions need to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils/ozone or oxidants, then (for ozone only) use xancil to convert to .pp.

Land surface boundary conditions

Vegetation fraction & function types

- Need to create new fractions (9 PFTs, single timeslice) and functional types (2 fields, 5 PFTs, monthly climatologies) based on Herold *et al*. (2014).
    - Fractions: This is done by taking the Herold PI biomes and the model's PI vegetation and calculating a typical set of vegetation types for each biome (must be done at N96, rather than the native 1° Herold resolution). Then create the Eocene vegetation fields by saying that for each grid point we find what its Herold biome type is and then choose the average set of vegetations for that biome. Lastly, zero out urban, lakes and ice (tiles 6, 7 and 9, respectively) and redistribute values equally amongst other tiles to maintain sum of 1 at each grid point

      *Note from W. Roberts: When regridding from 1° to N96, need to deal with fractional land sea masks. In previous work (using atmosphere-only), there was an issue with using a fractional LSM when doing the initial runs. This meant that an integer mask was used - the problem with this was that it needed to create some areas of land that aren't there at 10° but must be there to give a full unit of land at N96. This is fine, but as these points have no defined biome in Herold. The simplest workaround: define them as having vegetation parameters that are defined by the zonal mean. To clarify: the areas that are needed to be filled have zonal mean *vegetation parameters*, not zonal mean biome.*
    - Functions: this method is identical method used for Pliocene i.e. creating a seasonally & latitudinally varying function, based on PI zonal means for each

[4] Oxidants have a strong impact on the creation of sulphate and secondary organic aerosol, by governing oxidation rates of sulphate and monoterpene.

month, then creating a filled field with this, before finally imposing Eocene mask. This is true for LAI, whereas for canopy height it just uses the global mean of the PI, everywhere (this is because canopy height in the PI does not vary monthly, or even much latitudinally)

Full instructions, including code, can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/veg.d/round2/frac.d/readme.txt and func.d/new_way.d/readme.txt. Final versions need to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/vegetation/*, then use xancil to convert to .pp

Note: When creating vegetation fractions/functions, and indeed any field using the land sea mask (e.g. soils), ocean needs to be model's own missing data flag (2.0000e+20) instead of NaN, as otherwise model will fail. Although xancil will work either way, it will only convert to model flag (-1.0737e+09) if ocean = 2.0000e+20

Note: Can ignore "dist" file - it's not used at all. This is agricultural fraction (disturbance) for the old land use scheme in HadGEM2-ES. Neither HadGEM3 nor UKESM1 use it.

Topography/orography

- This has already been created by ancil suite, so don't need to do anything more with this other than transferring it to appropriate directory, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/orography

  Note: For unknown reasons, ancil suite will name this file qrparm.orog_1 whereas model will be looking for qrparm.orog so, rather than changing model code, simply change this filename (using `cp` to create a copy).

Soil parameters & soil dust

- Need to take modern soil parameters and soil dust fields (single timeslice), and globally palaeotise each[5]. Code can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/soil.d. Again, run IDL scripts before python. Final versions need to be transferred to

[5] Soil dust ancils prescribe the soil properties relevant to dust emissions: soil type (silt/clay/sand) and size distribution. Taking a global mean for these is sensible since these properties are affected by erosion / deposition processes which are likely to have a very different pattern in the Eocene.

NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/soil_parameters and soil_dust, then use xancil to convert to .pp

Note: Need to be careful about exactly where global mean is taken from. Modern soil parameters/dust will be zero over ice regions (i.e. Antarctica & Greenland) so, to avoid distorting global mean, need to average over non-ice regions i.e. actual values. But then, in new file, need to apply global mean everywhere i.e. including ice regions. Might be tempting to leave ice regions at zero, to be consistent with modern, but this won't work and will cause model blow up! This is because modern vegetation fraction has zeros over ice regions, so it is fine for soil to also be zeros (because no vegetation = no soil fields). But newly-created Eocene vegetation fraction DOES have values over ice regions (as it should), so soil fields can't be zero here. So need to apply global mean to these regions as well.

Note: already included in the code, but important to remember to reimpose land-sea mask onto global mean - original version of this, used in atmosphere only suite, used filled versions of soil parameters. However, for unknown reasons, needs to have mask here. Also remember to doublecheck multi-field rules still apply e.g. sum(sand/silt/clay) = 1.

Note from J Cole: need to use Generalised ancillary file option in xancil for both parameters and dust, because otherwise one of the parameters is missed out. For parameters, current version of xancil has a bug and doesn't pick up the fields, so use previous version (xancil0.57). For dust, need to update the actual stash codes because otherwise it puts all of the divisions 1-6 into one field, rather than 6 separate fields. Need to change the STASH code to the correct number, so click on each name and change this (i.e. Dust soil mass fraction div n has STASH code 420+n). Each time that STASH code is changed, click on the Update button, this should change the STASH name to the correct name. In current version of xancil, this is no longer necessary. However, what is necessary is to make sure that for each field, the mask is set to "Land mask" and is taken from the missing data as otherwise the large values remain. Also need to make sure each field has the correct variable name - click the select button and double click on the correct STASH name, this should change the variable name. There seems to be a bug in xancil where the Ok button in the select window resets the variable name back to its original value every second time. To work around this problem, click on the top right X button to close the window. Don't get upset if, when you click on Update, it gets rid of the _n-1 from the variable name

and, therefore, when looking at the resulting file, all dust divisions have the same field name i.e. field1633 (as opposed to, in the netcdf version, field1633_1 etc). This doesn't matter! This is only a feature of the netcdf versions, and even the PI version of this only has field1633 for all.

Others soil-related fields: soil moisture, deep soil temperature and snow depth

- Need to take modern fields (either single or monthly timeslices) and globally palaeotise soil moisture and deep soil temperature at each level/month, and set snow depth to 0 everywhere at each month. Code can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/smowslt.d. Again, run IDL scripts before python. Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/smc_snow and soil_temp, then use xancil to convert to .pp
- Neither of these files (qrclim.smow (containing soil moisture and snow depth) and qrclim.slt (containing soil temperature)) are currently reading via the reconfiguration, because in the PI these are brought in via the restart dump from previous parents. So once these files have been created, need to insert them into suite. In Eocene suite GUI, go to: um > namelist > Reconfiguration and ancillary control > Configure ancils… then right click and create a new section. Open this, then select the relevant fields e.g. for smow there should be 2 primary fields: soil moisture over land (23) and soil moisture content in a layer (9). Select both of these, then point the relevant file in the pathname. Make sure the domain is the whole grid, update_anc is false, and source = Initialise from ancillary file. Save, then repeat this process for slt (primary field: deep soil temperature (20)

Hydrology (topographic index)

- Need to take modern hydrology fields (single timeslice), and globally palaeotise each. Code can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/hyd.d. Again, run IDL scripts before python. Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/hydrol_lsh, then use xancil (again use “Generalised Ancillary File” option, and note that it will insert field names, hoping this won’t be a problem) to convert to .pp

River sequence & direction

- This has already been created by ancil suite using input from /home/d05/cwilliams/rivers_build, so don't need to do anything more with this other than transferring rivers output to appropriate directory, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/rivers_trip

Ocean boundary conditions

Sea ice

- Need to take modern sea ice used in PI suite, and zero certain fields: aicen, vicen, vsnon, uvel, vvel. Code can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/cice.d. Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/restarts. Don't need to use xancil here because model expects netcdf (unusually!)
  Note: Because sea ice will melt away so rapidly with increased temperatures due to CO_2 (see below), this step is possibly unnecessary and so probably fine to just use straight PI control sea ice.

Bathymetry

- This has already been created by process above, so don't need to do anything more with this other than checking it is in appropriate directory e.g. /home/d05/cwilliams/gc31/final_ancils_round2/ocean/bathy. Don't need to use xancil here because model expects netcdf (unusually!)

Ocean switches/ancillaries

- Ocean ancils used by the standard model are listed below - first 4 are based on information about present-day bathymetry which can't be replicated for the Eocene, so just switch off spatially varying parts of these schemes:
  - Tidal mixing (M2, K1 files): Switch it off tidal mixing scheme[6] (i.e. remove key_zdftmx in ~/roses/<SUITEID>/app/fcm_make_ocean/rose-app.conf)
  - Geothermal heating: Switch it off (i.e. set ln_trabbc= .false. in /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf)

[6] Prescribes increased vertical mixing in tidal regions, and location of these regions is prescribed in an ancillary file - locations are specific to the modern day so will not make sense for the Eocene, so switch off

- Indonesian throughflow mixing: Switch it off (i.e. set ln_tmx_itf=.false. in /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf)
- 2D bottom friction (bfr_coef.nc): Switch it off (i.e. set ln_bfr2d=.false. in /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf)
- Viscosity (ahmcoef.nc) - need to make new file, but using Eocene mask. Purpose of this file is to reduce viscosity between 20° N/S to increase equatorial current speed and vertical shear. This can't be applied next to western boundaries for stability reasons, or to ensure the WBCs remained sensible. So it needs to be 0 between +/- 20°, except for a 10° buffer to the east of any land mass, followed by a 5° ramp from 100 to 0. To do this, need to take modern version (at /projects/ocean/hadgem3/ancil/ocean/eORCA1v2.2x) then, using this, create new version using code on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/ocean.d/ahmcoef.py. Final version needs to be transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/ocean
- Iceberg calving & ice sheet basal melt flux need to be switched off (as no ice sheets). To do this, set in /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf:
    - ln_icebergs=.false.
    - nn_isf=0
    - rn_antarctica_total_fw_flux=0
    - rn_greenland_total_fw_flux=0

  Note: Last two may not be necessary as first two should switch off fluxes and hence their magnitude is irrelevant, but just in case…
    - Postprocessing/archiving may fall over if it tries to archive iceberg files but none exist (because iceberg advection scheme is off), so set in /home/d05/<USERNAME>/roses/<SUITEID>/app/postproc/rose-app.conf:
        - archive_iceberg_trajectory=true
        - nemo_icebergs_rst=true

Ocean initial conditions

Temperature & salinity

- Need to take modern control monthly climatology (75 levels), filled version (currently at /home/d05/cwilliams/gc31/final_ancils/ocean/ts_init/EN4_v1.1.1995_2014.monthlymean_eORCA1T_NEMO_L75_orig.nc) and modify temperature according to Lunt *et al*. (2017), but with 15°C reduced to 10°C (or even 5°C). Need to make salinity constant at 34.7 psu. Full instructions, including code, can be seen on JASMIN at /home/users/cwilliams2011/analysis.d/projects.d/sweet.d/gc31n96orca1_eo.d/build.d/ocean.d/ocean_init.d/readme.txt. Final version needs to be renamed to something appropriate e.g. EN4_v1.1.1995_2014.monthlymean_eORCA1T_NEMO_L75_gmdmod.nc and transferred to NEXCS, e.g. at /home/d05/cwilliams/gc31/final_ancils_round2/ocean/ts_init/. Don't need to use xancil here because model expects netcdf (unusually!)

As mentioned above, once all of these have been created and transferred to appropriate places on NEXCS, point all of these in new atmosphere and ocean ancillary versions file e.g. at /home/d05/cwilliams/gc31/ancil_versions/.

Part 2c: Setting up suite to become Eocene - Other scientific modifications

Once all ancillary files have been created/modified, a few more scientific modifications are needed to turn PI suite into Eocene suite. Again, it doesn't matter in which order these are done.

Switches to prevent buildup of snow on broadleaf trees, and to help prevent oceanic grid point storms

- Significant trial and error suggests that one possible reason for land blowup is the building up of too much snow on broadleaf trees at high altitudes, because in the PI there aren't any but here there are. Rather than fiddling with fraction ancil even further, easier way is small code change. None of the changes below are large enough to affect climate state (this has been tested). These are all found in /home/d05/<USERNAME>/roses/<SUITEID>/app/um/rose-app.conf
    - puns: change this to 1.0

    - canopy snow changes: change can_clump to 4.0,4.0,1.0,1.0,1.0, cansnowpft to .true.,.true.,.false.,.false.,.false. and unload_rate_u=2.31e-06,2.31e-06,0.0,0.0,0.0
- These can be seen in https://code.metoffice.gov.uk/trac/roses-u/changeset/160818/b/v/0/2/3/trunk (ignore reinit_end setting)

Coupling

OASIS coupler

- To avoid crash with OASIS coupler and newly-created coupling weights, need to replace existing namcouple file (in /home/d05/<USERNAME>/roses/<SUITEID>/app/coupled/file) with new one written by R. Hill (an example of which is currently in /home/d05/cwilliams/roses/u-bq324/app/coupled/file). Not suite specific, so don't need to edit it. If pasting text into existing file, don't use emacs as this leaves a residual backup file (e.g. "~namcouple"), which will break everything. Likewise don't rename existing file to e.g. namcouple_orig as this will also break everything. Instead, copy entire new file to appropriate directory, move existing one elsewhere, and rename new one
  Note from R. Hill: It [error with OASIS] won't be caused directly by the ancils. Looking at your namcouple file you seem to have a number of fields which want to use 2nd order conservative remapping, but none of your rmp files contain second order terms for the gradients, as far as OASIS is concerned. So the UM will be calculating gradients for these fields and trying to pass them to OASIS which doesn't want them. It looks like your rmp files have been generated by ESMF rather than SCRIP. ESMF doesn't generate 2nd order conservative weights in a way that's consistent with OASIS or the UM's understanding of what OASIS is expecting. So assuming that youere happy to use 1st order regridding for heatflux, sublimation and emp: if you replace the contents of the existing namcouple with the contents of a new one, then that should set things up to expect 1st order regridding for all fields.

Reconfiguration

- By default, UM reconfiguration is off switched off for most ancils, because the CMIP jobs deliberately do not reconfigure these fields. This doesn't matter for fields where update_anc=true, but for the rest (including land sea mask), the reconfiguration is

where they get into the model. Need to go to the reconfiguration panel (in um > namelist > Reconfiguration and ancillary control > Configure ancils…) and switch on (right-click, then "Enable") any row which uses a file that specified in ancillary versions file - best to sort by ancilfilename, then just do this for files beginning $UM_ANCIL (excluding the $MODE_INIT ones) plus ozone which begins $CMIP6_ANCILS

Other switches

- Need to set BITCOMP_NRUN=false (in /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf), at least for first run[7]
- Also best to reconfigure the coupling prognostics to default values - in um > namelist > Reconfiguration and ancillary control > Configure ancils…, sort by stash_req, and enable stash codes in the range 153-193

Start dump

- Absolutely fine to use same atmosphere start dump as standard CMIP6 PI control (currently at /home/d05/cwilliams/gc31/restarts/aq853a.da25000101_00), even though it uses a modern mask. This is because the reconfiguration "re-masks" the land and ocean fields in the dump, populating any new land points using the ancils which have been provided and filling others using a mix of nearest-neighbour filling and its own rules. This can be shown by looking at the reconfigured dump (e.g. /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data/<SUITEID>.astart) - it has the new mask for land fields

Atmosphere

Greenhouse gases

- Decide whether to set CO_2 at 2x, 3x or 4x PI control. Set within Eocene suite GUI: um > namelist > UM Science settings > Section 01 - 02 Radiation > Gas MMRs
- All others e.g. N_2O, CH_4, O_2 should stay as PI suite and CFCs should be 0

Ocean

[7] This enforces some computations which help to stabilise the model, particularly when getting started from ocean and atmosphere dumps which may not be consistent with each other.

Add in NEMO branch

- NEMO currently has hard-wired indices for modern straits, so need to remove these. A. Sellar has created a branch to do this, so to pick it up need to make 2 changes[8]:

  https://code.metoffice.gov.uk/trac/roses-u/changeset/108589

  https://code.metoffice.gov.uk/trac/roses-u/changeset/108610

  Need to modify rose-app.conf files so that they match above changes, in /home/d05/<USERNAME>/roses/<SUITEID>/app/fcm_make_ocean/rose-app.conf and /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf, respectively

  Note: Best to fcm commit *before and, most importantly, after making the changes. Changes will only be picked up after a recompile, so need to do this before running suite (or, if it is already running, will need to restart from the beginning, after a rebuild).*

Ocean forcing diagnostics

- In order to (theoretically) spin-up the model faster, need to run for ~50 years then run another NEMO suite. To do this, need to output ocean forcing diagnostics, to use in NEMO suite - these are standard in UKESM1 runs, but not in HadGEM3. A. Sellar has created a branch to add them, so need to do:

```
cd ~/roses/<SUITEID>
fcm commit   # merge is simpler if with a clean working copy
svn merge -r114463:HEAD https://code.metoffice.gov.uk/svn/roses-u/b/h/3/0/1/wind_diags          # merges in changes
fcm commit   # commit again
```

Processes/nodes and ocean timestep

Atmosphere processes

- Current setup of domain decomposition is in Eocene suite GUI (at suite conf > Domain decomposition > Atmosphere) or ATM_PROCX and ATM_PROCY in top-level conf file at /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf

[8] The second one is a namelist change which actually makes most of the difference by skipping all the bits of code which test 'if cp_cfg == "orca"'. In fact branch is only needed because there is one bit of code which only operates for global configurations (it changes viscosity near the equator) which is wanted because it makes no assumptions about bathymetry

and is: EW = 36, NS = 28, IO server processes: = 6, OpenMP threads = 2. This is the same as ATM_PROCX = 36, ATM_PROCY = 28, IOS_NPROC = 6, OMPTHR_ATM = 2 in rose-suite.conf. This will almost certainly fail after a couple of years runtime, giving halo error (see Fixing errors). So will need to reduce this. Could try: ATM_PROCX=24, ATM_PROCY=12 which will reduce the number of atmosphere nodes from 28 to 8[9]. Alternatively, try ATM_PROCX=12, ATM_PROCY=15, OMPTHR_ATM=1, IOS_NPROC=0 which will reduce the number of atmosphere nodes to 5 - however experience shows that this is VERY slow (e.g. it takes ~8 hours to do 6 months). Best to start with current decomposition but with no IO server processes[10] and reduce OpenMP threads to 1, which will halve the number of PE used by atmosphere (i.e. ATM_PROCX = 36, ATM_PROCY = 28, IOS_NPROC = 0, OMPTHR_ATM = 1) and then reduce East-West or North-South processes according to error message from halo error. This should give 36*28 = 1008 processes, which /36 (because on NEXCS there are 36 cores per node) = 28 nodes
*Note: Experience shows that above will need to be reduced at least once or twice due to halo error e.g. current Eocene suite (u-bh301) is running at ATM_PROCX = 32, ATM_PROCY = 20, IOS_NPROC = 0, OMPTHR_ATM = 1 which gives 32*20 = 640 processes, which /36 = 18 nodes.*

Ocean processes

- Current setup of domain decomposition is in Eocene suite GUI (at suite conf > Domain decomposition > Ocean) is: EW = 9, NS = 8. This appears stable, and works with ocean timestep, so leave as is

Ocean timestep

- Current setup is using a 45 minute ocean timestep. This is defined as:

  OCEAN_SEAICE_TIMESTEPS_PER_DAY=32
  nn_stock=960
  nn_write=320
  diagfreq=960

[9] This will slow the atmosphere down, but this doesn't matter because ocean timestep needs to be slowed down as well (see below) and will probably reduce queue time a lot because smaller jobs get scheduled faster.
[10] This step is important, because otherwise IO error will occur (see Fixing errors)

First of these is in /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf and the rest are in /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf. Quite likely that this will blow-up, giving various errors in /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/ocean.output (see Fixing Errors for how to diagnose). Possibly due to irregular/uneven bathymetry, however smoothing does not seem to resolve this. What does resolve it, however, is reducing the ocean timestep. Possible options: Multiply by 9, to get 5 minute timestep i.e.:

OCEAN_SEAICE_TIMESTEPS_PER_DAY=288

nn_stock=8640

nn_write=2880

diagfreq=8640

This works, however it is VERY slow. Might be tempting to try speeding this up by giving the ocean more processers to speed it up relative to the atmosphere, by setting:

NEMO_IPROC=15

NEMO_JPROC=12

in /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf. This will increase the ocean nodes from 2 to 5, but requires a rebuild, as NEMO PE decompositions are baked into the compilation. However, this fails at building stage, giving:
ice_distribution: Number of active blocks not equal to number of processors
which suggests that CICE is not happy with the PE decomposition. So need to revert back to original 2 nodes (09 and 08, respectively). Testing shows little difference in run time between 20 and 30 minute ocean timestep, so change to 20 minutes i.e.:

OCEAN_SEAICE_TIMESTEPS_PER_DAY=72

nn_stock=2160

nn_write=720

diagfreq=2160

Note: Might be tempting to revert back to 45 minute timestep once a year has run, however this is likely to blow-up straightaway as well.

STASH Table

- Current setup is for full, CMIP6 PI control STASH table, containing lots of high-frequency (e.g. hourly) diagnostics. This is about an hour of extra runtime per

coupled stage, and results in ~0.5 Tb of output data per year. For Eocene suite, and especially during spin-up phase, don't need much of this. Many items, especially high-frequency ones, can be turned off in STASH table (in Eocene suite GUI at um > namelist > Model input and output > STASH Requests and Profiles > STASH Requests) - efficient way to do this is to group by "tim_name" (menu in the top-left of the panel), expand profiles T1H, T3H or T6H, select the items and right-click, then "Ignore". If, however, want to import minimal STASH table from official CMIP6 PI control spin-up phase, need to:

1) Copy CMIP6 PI control spin-up phase suite (final version, u-aq853)
2) If the STASH macro to generate the hashes for the STASH namelist items (ie the new usage profiles) has not been run, need to do this first - run macro by going to um > namelist > Model input and output > STASH Requests and Profiles > STASH Requests and run stash_indices.TidyStashTransform.transform. If don't do this, might get error saying it can't find one of the profiles e.g. UP7. This means that although there are UP7 entries in the rose-app.conf, these are not the specification of what the usage profile UP7 is. These entries are the individual STASH rrequests listed in the stash table, but these need to be defined i.e. what the time, domain and usage profiles actually are.
3) Export this stash by going into /home/d05/<USERID>/roses/<SUITEID>/app/um and running rose macro stash_copy.STASHExport. Choose default values. This will generate a file called STASHexport.ini
4) Move this file into target suite (i.e. /home/d05/<USERID>/roses/<SUITEID>app/um (where now <SUITEID> = suite wanting to be changed)
5) Rename this file: mv STASHexport.ini STASHImport.ini
6) The import macro only adds to an existing suite's STASH ttable, it doesn't overwrite it completely. So, before importing, need to delete existing table, contained in /home/d05/<USERID>/roses/<SUITEID>app/um/rose-app.conf. Need to delete all STASH rrequests (*streq), as well as time, usage, & domain items (probably don't need to remove these, but no harm in doing so). Need to remove entire thing, e.g. [namelist:umstash_streq(00002_28c09)
7) Once saved, check back within Eocene suite GUI - the STASH table should now be empty

8) Import the previously exported stash table by (still within /app/um) running rose macro stash_copy.STASHImport, again choosing default values
9) Check it has worked. Within Rose, the STASH table should now contain new items. Still within this window, run 2 more macros: stashtestmask.stashTstmskValidate (which will check that the STASH cconfiguration is suitable for this UM configuration) and stash_indices.TidyStashTransform (which will tidy up any out of place items). It's possible that some items will have red warning triangles, which is likely because they are not suitable for this particular configuration - these should be changed so that they are ignored. Also, back at command line (still within /app/um), run rose macro --validate to check for other issues. Another way of checking is to look at actual number of items within rose-app.conf for each directory (i.e. original and new) by using grep umstash_streq rose-app.conf | wc -l. This number should be the same for both
Note: When exporting/importing, the number it reads through will not be the same as this number.
Note: All of this is explained in Exercise 6.2 from the Rose Conversion course, at http://cms.ncas.ac.uk/documents/training/RoseMay2017/practicals/further-exercises.html#stash-manipulation-macros. However, this is wrong! Instead, follow above.

- As always, best to fcm commit before and after doing this
Note: All of above needs to be done before submitting suite, because making such a change mid-run would require a reconfiguration step as it changes the contents of the restart dump.
Note: Based on experience, changing to this minimal STASH table reduces runtime of coupled task by ~1 hour, and reduces output to ~70G.

Run length, cycling and wallclock time

- Current setup is for 1 year run length, 6 months cycling and 5 hours wallclock time (in Eocene suite GUI at suite conf > Run initialisation and cycling). Will need to change this to appropriate values. Currently using long24 queue (see above), but don't want to request a wallclock time near this limit because time in the queue will be longer. Experience shows that using above number of processes (e.g. 32 x 20

atmosphere processes and 20 minute ocean timestep), model can do 1 year in ~7.5 hours (which includes coupled task, and postprocessing/pptransfer (see below), which overlaps next coupled task). So, to be safe, request 15 hours wallclock time with 1 year cycling and run length of whatever e.g. 100 years
Note: Worth a little trial and error: one suite with yearly cycling and another with 3-monthly. If there isn't a clear difference after a couple of years then it probably doesn't matter. Number of nodes will make a difference (see above), so many reduction in queuing time due to reducing the node count will more than offset the slower runtime of the model.

Housekeeping

- Housekeeping (in housekeeping > prune): currently set-up to remove remote log directories (in /home/d05/<USERNAME>/cylc-run<SUITEID>/log/job/<CYCLE>) from previous 3 months (specified by -P3M) an archive log directories from previous 9 months (specified by -P9M).

Postprocessing & transfer

- Need to follow instructions at: http://cms.ncas.ac.uk/wiki/Docs/PostProcessingApp. In order to find out whether upgrading stage is needed, look in /home/d05/<USERNAME>/roses/<SUITEID>/app/fcm_make_pp/rose-app.conf and /home/d05/<USERNAME>/roses/<SUITEID>/app/postproc/rose-app.conf to find out which version the me is pointing to. Likely to need upgrading, because NEXCS postprocessing went in at postproc_2.2 which is not tied to any UM version/configuation. If unsure, just run upgrade command as instructed (if already upgraded it will do nothing). Then follow steps at: http://cms.ncas.ac.uk/wiki/Docs/PostProcessingAppNexcsSetup to explain how to set up suite.rc file to allow non-interactive authentication from NEXCS to JASMIN
- Different streams are reinitialised at different times and thus are archived at different times, e.g. *pa is reinitialised every 30 days so should get a *pa file archived every month from month 2 onwards, *pb is reinitialised every 90 days so that won't get archived until month 4, *pm is reinitialised every 10 days hence why 2 are archived for the first month

Part 3: Running Eocene suite

Once everything is ready, need to commit all changes to repository - if suite is just saved, changes are only saved to local working directory, not trunk. Use fcm commit, which can be done even if suite is running, to commit this new branch to repository. If FCM error appears e.g. [FAIL] svn: E215004: No more credentials or we tried too many times then try re-caching MOSRS password using mosrs-cache-password (or just logout and login again). When asked, add relevant notes to file (e.g. what has been done). Don't actually need to do this in order to run suite, because all settings are taken from local working copy. But do need to do this if wanting to copy suite (e.g. to make another version) - if not, the copied version with revert back to previous committed version. To reverse a previous committed change, need to do

```
fcm merge --reverse --revision <rev_number>
fcm commit   # commit again
```

If a copy is made and needing to be compared to original, use diff -r <old_suiteID> <new_suiteID>.

Run! Assuming everything submits correctly, suite should begin building model with various fcm tasks, followed by install_ancil, followed by recon (if this is switched on) followed by coupled and finally postprocessing and pptransfer. Quite likely that validate_suite_info task[11] will fail - just need to right click on it, and reset the state to "Succeeded".

Basic commands

In following, if not already in <SUITEID> then need to add --name=<SUITEID> at end.

- rose suite-run : Runs suite from start
- rose suite-run -new : Runs suite from start, but completely afresh (i.e. does the same as doing rose suite-clean first)
- rose suite-shutdown: Shutdown suite

[11] This is just there for CMIP6 production runs and checks the metadata in rose-suite.info which is helpful information when data are disseminated but doesn't affect the run itself.

- rose suite-run --reload: Get suite to pick up a change, unless change requires a rebuild (this can be done whilst any task is running, and next time task runs it will use pick up new modification)
- rose suite-clean : Completely clean suite i.e. i.e. delete everything in /home/d05/<USERNAME>/cylc-run/
- rose sgc - if cylc it is not open, reopen it
- rose suite-restart : restart suite, either after fixing a failure or a shutdown

Differencing 2 suites

3 ways:

- Simplest way: diff -r <SUITEID1> <SUITEID2> however this produces just a list and is therefore difficult to read if there are lots of differences
- Website way: if e.g. <SUITEID1> = ab001 and <SUITEID2> = ab002, then go to https://code.metoffice.gov.uk/trac/roses-u/changeset/HEAD/a/b/0/0/1/trunk?old=HEAD&old_path=a/b/0/0/2/trunk
- Much easier way (within <SUITEID> directory): xxdiff --exclude=.svn --recursive <SUITEID1> <SUITEID2>

Output monitoring

- All *log* files and other output etc go to /home/d05/<USERNAME>/cylc-run/<SUITEID>/log/job/<CYCLE>/<APP>/NN on NEXCS (the NN directory always refers to version). To see how long each coupled cycle is taking, and also how long is being spent in the queue, look at /home/d05/<USERNAME>/cylc-run/<SUITEID>/log/job/<CYCLE>/<APP>/NN/job.out (this file is large, so much so that emacs cannot read it and nedit can, but fails to show all of it, without any warnings - instead, do tail –n 80 job.out to show last 80 lines (or however many is needed). One job.out file is produced for every cycle, not the entire run, e.g. if running for 50 years with 1 year cycling (i.e. resubmission), there will ultimately be 50 cycles corresponding to each. However, only last cycle will be shown, as previous cycles are automatically tarred and zipped - they will be in /home/d05/<USERNAME>/cylc-run/<SUITEID>/log and ending *.tar.gz
- When suite is queueing within Cylc, it only shows the previous, current and next cycle

- All *actual output* is found under 2 locations on NEXCS: firstly latest output (e.g. last cycle of run) is found in /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data/History_Data (with subdirectories CICEhist for CICE output and NEMOhist for NEMO output) and secondly all other output goes to the temporary disc at /projects/nexcs-n02/<USERID>. The directory is called share because files under here are shared between all cycles of the suite run. These are the standard locations for the output data and it is NOT recommended to change these. But for reference, the temporary disc space is specified at postproc > Post Processing - common settings > Archer Archiving)
- If automatic transfer to JASMIN has been turned on (see Part 2c), and NEXCS account on JASMIN has been granted, then pp files will automatically be transferred from /projects/nexcs-n02/<USERID> to /group_workspaces/jasmin2/nexcs/<USERID> (pathname is specified in postproc > Post Processing - common settings > JASMIN transfer)

Restarting

Note: There is no way to get DJF or annual mean out for first year of any run. Reason for no DJF mean is obvious, as December would be before start of the run. Reason for no annual mean is because automatically-calculated annual mean (i.e. that produced by model in output) goes from December-November, not January-December, therefore again there is no December. Moreover, automatically-calculated annual mean is built from DJF + MAM + JJA + SON, so cannot be produced because there is no DJF. When run is restarted mid-way through with rose suite-run*, model is effectively starting new run and knows nothing about previous portions. 2 options to resolve this:*

a) *Save annual mean from archived previous year (e.g. rename or move it), then restart this earlier year e.g. if blowup occurs in August 1880, save archived output from 1879, then rewind and restart from January*
b) *Calculate DJF / annual mean manually from monthly means (if this is done, manually-calculated annual mean will not be identical to automatically-calculated annual mean, because of above difference in meaning periods e.g. November-December vs January-December)*

Note: When it comes to changing ancillary files (e.g. adding in a modified ancil to an already running suite), then technically speaking rose suite-run --reload *is fine, and the model will start using it on the next cycle. But whether that's scientifically a good idea will depend on the experiment, because an instantaneous jump in the boundary conditions is being imposed. It depends how far it has run. For example, in the case of ozone, if it has already warmed much more than 1K then there may already be an excess of moisture in the lower stratosphere, and not sure how long the model would take to reduce this once the excess moisture supply at the cold point is cut off.*

If model falls over

- Shut down suite (using rose suite-shutdown), fix problem, then restart using rose suite-run. Using this, it knows where it got to and will start from the beginning of failed cycle, not beginning of suite. Cylc window will then reopen, showing failed stage - need to right click on it and select Trigger (now)

If suite has successfully completed but want to run for longer

- If wanting to restart from same endpoint (e.g. to do another 10 years), just need to change the run length (still in suite conf > Run initialisation and Cycling) then just do rose suite-run --restart - it will now know where to start from, so no need to input a restart dump. Note: The "Total Run Length" is from the Model Basis time not from where it left off, so need to add on new number of years to that already done e.g. if 100 years has been done and want to do another 50, set total run length to 150 years
- Occasionally, model might fail straightaway - see Fixing errors below

If suite needs to be restarted from a previous restart dump, or suite has successfully completed but want to run again using new suite

- If suite needs to be restarted from a previous restart dump - full instructions on how to do this are at: https://code.metoffice.gov.uk/trac/moci/wiki/tips_CRgeneral#Restartingfromarchivedrestarts. Basically need to find 5 restart dumps: atmosphere, ocean physics, ocean tracers, CICE and icebergs. If restarting from previous cycle/year, which has already been archived to JASMIN using pptransfer, then these will be on JASMIN at

/group_workspaces/jasmin2/nexcs/<USERID>/<SUITEID>. If restarting from same cycle/year, then restart dumps will be in /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data on NEXCS.

Note: Atmosphere and ocean tracers are defined in one place, as described at above website and below. However, ocean physics, CICE & icebergs can be defined in 2 places, either in top-level config file (in GUI at suite conf > Run initialisation and cycling, or /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf) AND in NEMO app (in GUI at nemo_cice > Restart files, or /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf). Not clear which overrides which, and need to be very careful here! If, for example, icebergs restart is forgotten, model will fail with following error in /home/d05/<USERID>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/ocean.output:

```
==⇒>> : E R R O R
===========
        iom_open ~
        File ./restart_icebergs.nc* not found
```

If this happens, advice from CMS (ticket) is wrong! They say to leave the NEMO start dump at nemo_cice > Restart files as blank, and instead insert the NEMO dump the top-level config. Although technically this works (i.e. it does run), leaving the nemo_cice > Restart files as blank is very very bad scientifically, because this means the ocean (and therefore icebergs, which is why it works without finding the appropriate restart file) STARTS FROM REST/CLIMATOLOGY! If this is done, therefore, every time model is restarted is essentially the same as a BRAND-NEW RUN, in terms of ocean spin-up. Therefore, DO NOT DO THIS! Instead, follow instructions at above website carefully, but with one small change: instead of defining these 3 restarts within the nemo_cice > Restart files, instead define them in top-level config and then point to them within nemo_cice > Restart files, i.e.

```
$NEMO_ICEBERGS_START
$NEMO_START
$CICE_INIT
```

(or, within /home/d05/<USERNAME>/roses/<SUITEID>/app/nemo_cice/rose-app.conf):

```
NEMO_ICEBERGS_START=$NEMO_ICEBERGS_START
NEMO_START=$NEMO_START
```

CICE_START=$CICE_INIT

Then define actual 3 files within top-level config.

- Atmosphere: There is one restart dump, output at the beginning and end of the year e.g. January and December. Apart from first year of run, when it will be in year 2's cycle, these are all within current cycle. Filename is <SUITEID>.da<YEAR>1201_00 e.g. /group_workspaces/jasmin2/nexcs/cwilliams2011/sweet/gc31/u-bh301/18500101T0000Z/bh301a.da18501201_00. This needs to be transferred back to NEXCS and moved into appropriate directory e.g. /home/d05/<USERID>/restarts and then pointed to in GUI at um > namelist > Model Input and Output > Dumping and Meaning > astart
- NEMO: There are 2 restart dumps, one for physics and one for tracers[12], and (if on JASMIN) each year contains beginning of that year (i.e January) and end of the previous year (i.e December). Filenames are <SUITEID>o_<YEAR>_restart_<PE>.nc (physics) and <SUITEID>o_<YEAR>_restart_trc_<PE>.nc (tracers) e.g. /group_workspaces/jasmin2/nexcs/cwilliams2011/sweet/gc31/u-bh301/18500101T0000Z/bh301o_18500101_restart.nc and bh301o_18500101_restart_trc.nc.

  Note: If wanting to look at those from NEXCS (i.e. in between yearly archiving) directory will contain multiple restart files, one for each PE, but they are all at very coarse resolution (40,42) rather than expected 362,332. This is because, apart from the first cycle (when it reads a global restart), each PE reads and writes its own restart, so roughly 40x40 points in each file. Each PE has its own MPI domain (see above). Here, there are 9 PEs in the x direction and 8 PEs in the y direction, so 360 / 9 = 40 points (similarly for the y direction). The postproc app combines them into a single file once per year for archiving. If files are wanting to be looked at in between yearly archiving, need to rebuild it in the same way as the meshmask files e.g. find most recent NEMO restart physics files (or tracers) at /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data/History_Data/NEMOhist on NEXCS and then run

[12] Contains passive diagnostic tracers for e.g. age of water - contains very few fields, but still needed for restarting

~hadtq/bin/local_rebuild/rebuild_nemo_login <FILENAME> <NUMBER> e.g. if there are 72 ocean restart files, then ~hadtq/bin/local_rebuild/rebuild_nemo_login bk568o_18610501_restart 72.

Once rebuilt, need to be transferred back to NEXCS and moved into appropriate directory e.g. /home/d05/<USERID>/restarts and then pointed to in GUI, at suite conf > Run initialisation > NEMO start dump (for the physics file) and ocean_passive_tracers > env > Initialisation Settings (for the tracers file)

- CICE: There is one restart dump, and again (if on JASMIN) each year contains beginning of that year (i.e January) and end of the previous year (i.e December). Filename is <SUITEID>i.restart.<YEAR>-00000.nc e.g. /group_workspaces/jasmin2/nexcs/cwilliams2011/sweet/gc31/u-bh301/18500101T0000Z/bh301i.restart.1850-01-01-00000.nc. This needs to be transferred back to NEXCS and moved into appropriate directory e.g. /home/d05/<USERID>/restarts and then pointed to in GUI at suite conf > Run initialisation > CICE start dump
- Icebergs: There is one restart dump, and again (if on JASMIN) each year contains beginning of that year (i.e January) and end of the previous year (i.e December). If on NEXCS, individual PEs will again need to be rebuilt, as above. Resulting filename is <SUITEID>_icebergs_<YEAR>_restart.nc e.g. /group_workspaces/jasmin2/nexcs/cwilliams2011/sweet/gc31/u-bk944/18700101T0000Z/bk944o_icebergs_18700101_restart.nc. This needs to be transferred back to NEXCS and moved into appropriate directory e.g. /home/d05/<USERID>/restarts and then pointed to in GUI at suite conf > Run initialisation > NEMO icebergs start dump

  Note: This will only be needed, and indeed will only exist in output (on either JASMIN or NEXCS) if iceberg modelling is turned on e.g. for PI, Pliocene or others. For Eocene, iceberg modelling is turned off, so icebergs restart dump will not exist in output and is not needed for restarting

- Follow rest of instructions at above website e.g. changing start date to match restarts, switch off compilation/building (although will need to switch these back on if a rebuild is required e.g. if anything other new ancillaries have been added), switching

off reconfiguration, switching off testing (turning BITCOMP_NRUN = true, in either GUI at suite conf > Run initialisation and cycling, or /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf*)*

- When ready, resubmit as a fresh run: rose suite-run
  Note: No need to reconfigure if just updating ancillary files, but if adding new fields need to run reconfiguration. If running reconfiguration, need to change the path of ainitial, if not then it's variable astart. If adding new ancillary files, can just specify name and location within Rose, don't need to use FCM to create new branch.
- If suite has successfully completed but want to run again using new suite - need to firstly copy suite (see above) then point new suite to above restarts, and submit

If suite needs to be restarted but with a perturbation to the atmosphere (e.g. to resolve a grid point storm)

If model fails during coupled stage with no obvious error message in job.err (e.g.

```
_pmiu_daemon(SIGCHLD): [NID 07624] [c11-2c2s2n0] [Mon Aug 24 12:11:58
2020] PE RANK 1009 exit signal Aborted
atpAppSigHandler: Back-end never delivered its pid. Re-raising signal.
[NID 07624] 2020-08-24 12:11:58 Apid 113542160: initiated application
termination
[FAIL] run_model # return-code=137
2020-08-24T12:12:08Z CRITICAL - failed/EXIT
```

then worth looking in ocean.output file in case of oceanic blowup, at /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<DATE>/coupled/ocean.output. Possible error might be:

```
===>>> : E R R O R
      ===========
 stpctl: the zonal velocity is larger than 20 m/s
 ======
kt=****** max abs(U):  3.5340E+06, i j k:   118  289   10
```

Can check this by rebuilding the NEMO dump, as above. But if wanting to restart:

- check latest version has been committed, then copy suite to a new ID (this is not actually necessary, but wise)

- Copy usual restart dumps (atmosphere, ocean physics, ocean tracers, ice, icebergs) to another directory e.g. /home/d05/<USERID>/restarts, then then point to these within new suite (as well as checking all usual restart points, see above. Make sure that these are from year before blowup e.g. if blowup occurs in 1900, restart from 1899, as otherwise annual/seasonal output from first year will be missing, causing a gap
- Make sure, if it isn't already, that BITCOMP_NRUN=false (search for this within GUI)
- Set stphseed=1 (either search for this within GUI once UM app is open, or look in */home/d05/<USERNAME>/roses/<SUITEID>/app/um/rose-app.conf*)
- Run as usual

Note: Might be that most recent restart dump contains no sign of problem, as normally would only expect to see obviously bad gridpoints a few timesteps before blowup. Fortunately NEMO is much more helpful than the UM in this regard, and spits out entire dump of the state when it fails cleanly like this. Look at set of output.abort_*.nc files in /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<DATE>/coupled and rebuild as above.

Stopping/killing a suite

- If needing to stop/kill the suite: right click on task then 'Kill' (if Cycl GUI has frozen, just close window then run cylc gscan or rose sgc to reopen it). Then rose suite-shutdown --name=<SUITEID> to completely shutdown the suite (or just rose suite-shutdown within /home/d05/<USERNAME>/roses/<SUITEID>). This may not completely kill the suite (when trying to resubmit, error will appear saying it is still running). If so:
    - Firstly, within /home/d05/<USERNAME>/roses/<SUITEID>, do cylc stop '<SUITEID>'
    - Might also need to kill actual processes - check with ps -flu <USERID> | grep <SUITEID> and then kill any running processes (2nd number)
    - Might also need to delete contacts file at /home/d05/<USERNAME>/cylc-run/<SUITEID>/.service/contact

Charles Williams
Last updated: 13/8/20

FIXING PROBLEMS

Building

- [FAIL] mirror.target = : incorrect value in declaration
  [FAIL] config-file=/working/d05/cwilliams/cylc-run/u-bm327/work/18800101T0000Z/fcm_make_pp/fcm-make.cfg:4
  [FAIL] config-file= - file:///home/d04/fcm/srv/svn/moci.xm/main/trunk/Postprocessing/fcm_make/postproc.cfg@2381:12
  [FAIL] config-file= - - file:///home/d04/fcm/srv/svn/moci.xm/main/trunk/Postprocessing/fcm_make/inc/remote.cfg@2381:6
  [FAIL] fcm make -f /working/d05/cwilliams/cylc-run/u-bm327/work/18800101T0000Z/fcm_make_pp/fcm-make.cfg -C /home/d05/cwilliams/cylc-run/u-bm327/share/fcm_make_pp -j 4 # return-code=9
  2019-08-30T11:56:19Z CRITICAL - failed/EXIT
- Implies failure during compilation/building phase. Due to /home/d05/<USERNAME>/roses/<SUITEID>/site/meto_cray.rc where host = line in the HPC section is set to localhost but needs to be xcs-c for the compilation (FCM) to go through. To fix, add these 2 lines to [[HPC_SERIAL]] block, and [[OCEANBUILD_RESOURCE]] block:

  ```
  [[[remote]]]
    host = xcs-c
  ```

- However, once all the compilation has gone through, need to change [[HPC_SERIAL]] block back to host = localhost as otherwise the postprocessing/pptransfer stages won't submit automatically (because it will be trying to connect to xcs-c from xcs-c (i.e. from itself), which it won't like. So as soon as first coupled stage is running, change this back then reload the suite (rose suite-run --reload). Don't need to shut down first, because change will be picked up next task. Won't need to do this again, as long as suite is committed - possibly not true, even if it is committed this might need to be changed again, twice. If this 2nd step isn't done, tasks won't automatically submit (see error below)
- (xcs-c) 2018-10-29T14:51:41Z [STDERR] qsub: error: [PBSInvalidProject] 'climate' is not valid for collaboration trustzone on XCS

Means submission of tasks (during FCM stages) to XCS has failed. Check project code and default account (see Part 1a)

- [FAIL] Unable to obtain a Cray Compiling Environment License
  Means pool of licences has dried up, probably because too many people are running compilations. Try retriggering task again. Small chance it will have got stuck in an inconsistent state, so stop suite and start it again with rose suite-run –new
- [FAIL] svn: warning: W170000: URL 'file:///home/d04/fcm/srv/svn/um.xm/main/branches/dev/marcstringer/vn10.7_oasis3_geto2a_warn_bug' non-existent in revision 61207
  Means one of branches (in Eocene suite GUI at fcm_make_um > Sources), and in particular revision number, is either missing or out of date. To find it, go to UM trac system and search for missing branch: browse source tab, put 61207 in the "View revision" box (else you get only the latest revision), then click down through directories, then click on "Revision log", then click on ticket number (to describe purpose of the branch), then add latest revision number into GUI
- Ancillary File does not exist.
  File : '/home/d05/cwilliams/gc31/ancils/vegetation/fractions_igbp/qrparm.veg.frac'
  This might either be in job.err or job.out. If it says it can't find this ancillary, but certain that it is there, doublecheck ancilfilename= (in /home/d05/<USERNAME>/roses/<SUITEID>/app/um/rose-app.conf) is surrounded by 2 quotes only
- To check differences between 2 suites: xxdiff --exclude .svn -r <SUITEID> <SUITEID>
- To check suite is running/monitor it: grep -i Atm_Step /home/d05/<USERNAME>/cylc-run/SUITEID>/<CYCLE>/coupled/pe_output/<SUITE>.fort6.pe0000 | tail
- (oasis_advance_map) ERROR: in nwgts and coupling terms 1, 2
  If this error is shown in /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/debug.root.01, suggests a crash in the OASIS coupler. Doublecheck that namcouple replacement is correct (see Part 1c)
- Error from routine: IOS_INIT:IOS_RUN
  Error message: IO Server is not active
  If number of processes for OpenMP threads = 1 (see above), but IO server processes has been left as >1, then this error will be given as IO server won't work unless there

are multiple threads. So make sure IO server processes = 0 in Eocene suite GUI (or IOS_NPROC=0 in rose-suite.conf)

Restarting

- Error message: [FAIL] Unable to find top restart files for this cycle. Must either have one, or as many as there are nemo processors (73). [FAIL] Found 72 iceberg restart files

  For unknown reasons, when starting an existing suite from its own endpoint, occasionally (not always) above error occurs. Error message is actually misleading, because there is an error in the error message (!) where the number of processes (e.g. 73) is actually 1 greater that the actual number of processes, 72. Model will only run if there is 1 restart dump, or 72 (or however many processes there actually are). However, because (once the model successfully finishes) the PE files are rebuilt into one global file, at /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data/History_Data/NEMOhist/, there are suddenly 73 present and model falls over. To resolve, simply move the rebuilt iceberg file AND the rebuilt tracers file, leaving the remaining PE files where they are (don't need to remove the physics file). Then do rose suite-run --restart and trigger failed task.

Running

- Error message: Convergence failure in BiCGstab, omg is NaN
  This is a common point for the model to fail if it has ingested or developed NaNs or infinities. See the following URL for more information:
  https://code.metoffice.gov.uk/trac/um/wiki/KnownUMFailurePoints
  Very common, catch all, error. Website doesn't help! Most common reason a mismatch land-sea mask with one of the ancils, meaning the model is expecting missing data but gets data, or vice versa. But doesn't say which ancil. Several options for tracking this down:
  - Look at fields exchanged via the coupler, since they get exchanged at timestep 0 - i.e. there will be some things arriving from the ocean before the start of the first timestep (effectively initial conditions), to check if anything looks wrong either as they come out of ocean (before regridding) or go into atmosphere (after regridding). To see these, in namcouple replace any occurrences of

"EXPORTED" with "EXPOUT", which will tell coupler to produce netcdf files containing coupling fields. Files will appear at /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled and will be derived from field names in namcouple: anything with "toyatm" = field on atmosphere side of coupling exchange, anything with "toyoce" = field on ocean side of coupling exchange e.g. model01_O_SSTSST_toyoce_01.nc = SST coming out of ocean and ocn_sst_toyatm_01.nc = SST as it goes into dear after being regridded. If fields aren't obvious from file names, can pair up outgoing/incoming fields using number at end of name (e.g. _64 indicates topmelt category 4) or by cross referencing contents of namcouple file

- If everything looks okay here, error is most likely due to one newly-created/modified ancillaries. Need to go through each one very carefully, checking things like: land-sea mask (if there is one) matches actual land-sea mask, data are not upside down, latitudes are not upside down, longitudes are not reversed/displaced, data are consistent with other ancillaries (e.g. soil parameters versus vegetation fraction, see Part 2b), something has gone wrong xancil stage. Best to use a bisection technique when running model to find which ancillary is problematic i.e. revert to modern/unmodified versions for half of them, see if it works, then try other half, see if it works, then half these, see if it works, and so on

- lib-4611 : UNRECOVERABLE library error
  Missing opening (left) parenthesis in format.
  PE RANK 21 exit signal Killed
  [NID 00675] 2018-10-30 12:57:19 Apid 44028491: initiated application termination
  [FAIL] um-recon # return-code=137
  2018-10-30T12:57:20Z CRITICAL - failed/EXIT
  This is a generic, and unclear error. Often job.out is more useful than job.err, though for coupled suites job.out sometimes doesn't contain atmosphere or ocean output - instead, have to look in /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/pe_output/* (for UM / JULES) and /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/ocean.output (for NEMO / CICE). If job.err suggests failure might come from one of the UM PEs (e.g. pe6), switch on output from all PEs by setting prnt_writers=1 (in /home/d05/<USERNAME>/roses/

<SUITEID>/app/um/rose-app.conf) and/or change diagnostics to "Extra" for both model and reconfiguration (in Eocene suite GUI at um > env > Runtime controls > Atmosphere only and Reconfiguration only). Alternatively, search for PRINT_STATUS and RCF_PRINTSTATUS, respectively in GUI[13]. Don't forget to turn these back to "Normal" once suite is running, to save space

- [FAIL] run_model # return-code=137
  2018-12-07T15:03:43Z CRITICAL - failed/EXIT
  Again, generic error. If nothing obvious in job.err or job.out, might be a problem with the ocean, so look in /home/d05/<USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/ocean.output and see if there are any ERRORS. If so, it should give indices of where blow-up has occurred. Best way of looking at this is to rebuild ocean files - if model has failed due to this, there will be 72 output.abort_*.nc files (one for each PE). These can be rebuilt into a global file using the same script as above i.e. ~hadtq/bin/local_rebuild/rebuild_nemo_login ocean.abort 72 and then viewed in xconv or whatever (use varying colour scales to identify problem area). If error message in ocean.output contains something like max velocity is at level 33, first thing to check is currents, temperature and salinity at that level. One possible reason: lot of noise and a nasty temperature gradient in the eastern Mediterranean, due to the connection (thanks to the new bathymetry) between the Mediterranean and the Indian Ocean - because they have very different densities, nasty gradients might appear. Solution is to apply a zonal mean to the initial temperature and salinity fields (see Part 2b)
- Error message: North/South halos too small for advection. See the following URL for more information:
  https://code.metoffice.gov.uk/trac/um/wiki/KnownUMFailurePoints
  This is an atmosphere error, coming from when the wind is too fast such that the halo rows are not wide enough for the advection. It might be that the Eocene has slightly faster winds than the modern world, or it may be a symptom of a dynamical instability. Original halo size: East-West = 4, North-South = 5. Might be resolved by simply rewinding and restarting (see above), but with a perturbation[14] i.e. set

[13] Note: When searching in GUI, search box is case-sensitive (so try both upper and lower, or a mixture) and will only work if app is open (e.g. using above example, must click on UM app first)
[14] This tells the stochastic physics scheme to use a new seed instead of the one stored in the dump - might be enough of a kick to change the simulation to avoid the specific conditions which trigger the failure. Remember to reverse these changes next time you want to restart without changing results.

BITCOMP_NRUN=false (in GUI at suite conf > Run initialisation and cycling, or /home/d05/<USERNAME>/roses/<SUITEID>/rose-suite.conf and stphseed=1 (in /home/d05/<USERNAME>/roses/<SUITEID>/app/um/rose-app.conf) If this doesn't work (i.e. it blows up again in same location/time), try increasing this by an integer of 1 for each, in Eocene suite GUI at um > namelist > Top-level model configuration > Parallel communication options. Main problem with this option is that it will require processes to be decreased (see below), possibly in both directions, which will dramatically slow everything down.

Note: first thing to do (which CMS website says to do) is to restart suite but with extra diagnostics turned on (in GUI at um > env > Runtime controls > Atmosphere only, or just search for "PRINT_STATUS" in GUI). Don't need to shut down or restart, can just do rose suite-run --reload.

Quite likely that, when resubmitted, model will fail at the beginning of coupled task, giving:

- Error from routine: DECOMPOSE_FULL
  Error message: Too many processors in the North-South direction (28) to support the extended halo size (6). Try running with 23 processors.

  This is because there is presumably a limit to how big the halos can be relative to the PE domain size - if the halo is so big that it straddles more than one neighbour then the code won't handle that. Error is suggesting a smaller number of PEs so that each domain is larger. Follow instructions in error i.e. reduce the number of atmosphere processes in North-South direction. It might fail again, giving same error but from East-West direction, so again follow instructions.

  Note: May well need to go through the entire process more than once, i.e. if model again blows-up with halo error, may need to increase halo sizes again e.g. to East-West = 6, North-South = 7, and then subsequently reduce atmosphere nodes according to error instructions.

  Note: However, if this error occurs more than twice, possibly suggestive that the Eocene configuration has significantly stronger maximum winds than the modern. If there isn't an obvious scientific explanation, increasing the halo size might just paper over cracks, and it might be better to investigate first in case there is a malign cause. First question to answer: did errors occur in the same region of the world? To check this, need to look at problematic processor ID (which will either be in actual error message e.g. Error from processor:, or will be in /home/d05/<USERNAME>/

cylc-run/<SUITEID>/work/<CYCLE>/coupled/pe_output files) and use this to work out longitude/latitude of blow-up (see below for instructions). Each time suite is resubmitted, this will be overwritten, so need to make a date of processor number, resubmit, then make a note of next processor number if error occurs again.

- Error from routine: set_thermodynamic

  Error message: A total of 1 points had negative mass in set_thermodynamic. This indicates the pressure fields are inconsistent between different levels and the model is about to fail. See the following URL for more information: https://code.metoffice.gov.uk/trac/um/wiki/KnownUMFailurePoints

  This is another general error, and again the website doesn't help. Website implies that this only occurs due to hardware problems resulting in data corruption, which might be true for a modern run but not necessarily for an Eocene run. Website recommends to try re-triggering the failed coupled task - this may or may not work (i.e. if model is at very edge of an instability, restarting from a previous dump might work)

  Note: With 12- or 6-monthly cycling, need to start from January restarts, and best to go back to the January before the start of the cycle it failed on.

  But if restarting doesn't work, or if this error continues to appear, suggestive that there is an instability building up over time (especially if, when restarted, coupled task runs for a few hours i.e. gets a few months in). In which case, need to work out exactly where error is occurring, by looking at PE ID and calculating longitude/latitude (see below). If error has occurred more than once, need to record location at each failure to see if it's occurring in the same place. Then, in most recent dump (e.g. output from day or month before blow-up occurred, which won't been archived to JASMIN yet so will still be in /home/d05/<USERNAME>/cylc-run/<SUITEID>/share/data/History_Data), look at possible culprit fields e.g. surface temperature, streamfunction or subgrid-scale noise in horizontal/vertical winds. Possible culprit: SST/deep ocean too warm due to too much warmth in ocean initialisation fields, and/or atmosphere to warm due to CO_2 level being too high?

- Error: Disc quota exceeded.

  This means that disc quota has been exceeded in either /projects/nexcs-n02/<USERID> or home directory on NEXCS. In home directory: can delete tarred up files in /home/d05/<USERNAME>/cylc-run/<SUITEID> e.g. log.*.tar.gz. Can also delete log directories in /home/d05/<USERNAME>/cylc-run/<SUITEID>/log/job but

just keeping the last few. Shouldn't fill up /projects/nexcs-n02/<USERID>, because if pptransfer has been set up (see Part 2c) then all data should be automatically archived to JASMIN

- Error: "submit-failed" on certain tasks (e.g. postprocessing) which appears to happen randomly.

  This is not a random problem with host machine, but rather due to "mkstemp" error which occurs from time to time if host is set to xcs-c - causes suite to login from xcs to xcs-c and causes problems with the creation of jtmp file. To fix, see above
- Error: "archive_integrity" failure after pptransfer.

  The archive_integrity app only works with data that is being put into MASS, which is not needed here, so should be switched off. It is currently set to run every 10 years, hence this error will only appear every 10 years, not straightaway. Need to change "Archive_Integrity - Post Proc App Gap checker" option (in suite conf > Build and run > Postproc app) to False. Then do rose suite-run --reload - no need to stop and restart the suite
- No obvious error in job.err, job.out or similar

  If no error is obvious, might well be a dynamical blowup in ocean. Look in /home/d05/USERNAME>/cylc-run/<SUITEID>/work/<CYCLE>/coupled/ocean.output for any possible errors

To work out longitude/latitude based on PE ID

- PEs are ordered left-to-right from the bottom-left (the SW corner of the grid to the NE corner) according to domain decomposition, so check number of PEs in each row and column respectively (see above for where Eocene GUI or top-level conf file) e.g. if it is East-West = 24, North-South = 12 (which = ATM_PROCX=24, ATM_PROCY=12), then there will be 24 PE columns and 12 rows:

  24 25 26 27 28 ….. 47

  0 1 2 3 4 ….. 23

  Alternatively:

  2xPROCX 2xPROCX+1 2xPROCX+2 ..…

  PROCX PROCX+1 PROCX+2 ..…

  0 1 2 ….. PROCX-1

so need to find exact x,y (or i,j if that convention is preferred!) location of PE, then convert this indices to a longitude/latitude based on grid spacing - this will be determined by 360 divided by number of East-West processes and 180 divided by number of North-South processes, i.e. 360 / 24 = 15 and 180 / 12 = 15, so grid spacing is 15° x 15°. Know that grid starts at 0,-90, so use this to get longitude/latitude of the SW corner of PE by doing:

lon = (x-1) * dx where dx=360 / PROCX

lat = -90 + (y-1) * dy where dy=180 / PROCY

Example 1 (EW = 24, NS = 12): Problem is occurring at processor 28 (*Error from processor: 28*). Above shows that this would be x=5 (out of 24 PEs in x direction), and y=2 (out of 12 PEs in y direction) - i.e. starting in bottom-left corner and working along each row, then after 28 PEs it is in the 5th column and 2nd row. So now have our x,y: 5,2. So, using above:

lon = (x-1) * dx = 4 * (360/24) = 4 * 15 = 45

lat = -90 + (y-1) * dy = -90 + 1 * (180/12) = -90 + 15 = -75

So SW corner of PE begins at 45°E, 15°S and, as grid spacing is 15°, PE extends to 60°E, 0°S. Above equation can be used regardless of domain decomposition i.e. number of PEs

Example 2 (EW = 36, NS = 23): Problem is occurring at processor 342 (*Error from processor: 342*). Above shows that this would be x=19 (out of 36 PEs in x direction), and y =10 (out of 23 PEs in y direction) - i.e. starting in bottom-left corner and working along each row, then after 342 PEs it is in the 19th column and 10th row. So now have our x,y: 19,10. So, using above:

lon = (x-1) * dx = 18 * (360/36) = 18 * 10 = 180

lat = -90 + (y-1) * dy = -90 + 9 * (180/23) = -90 + 70.43 = -19.56

So SW corner of PE begins at 180°E, 19.56 °S and, as grid spacing is 10° x 7.83°, PE extends to 190°E, 11.73°S.

Example 3 (EW = 32, NS = 20): Problem is occurring at processor 305 (*Error from processor: 305*). Above shows that this would be x=18 (out of 32 PEs in x direction), and y = 10 (out of 20 PEs in y direction) - i.e. starting in bottom-left corner and working along each row, then after 305 PEs it is in the 18th column and 10th row. So now have our x,y: 18,10. So, using above:

lon = (x-1) * dx = 17 * (360/32) = 17 * 11.25 = 191.25

lat = -90 + (y-1) * dy = -90 + 9 * (180/20) = -90 + 81 = -9

So SW corner of PE begins at 191.25°E, 9°S and, as grid spacing is 11.25 ° x 9°, PE extends to 202.5°E, 0°S.