CLM-NorwayDoc Documentation

Release 1.0.0

May 25, 2021

Welcome to CLM-Norway's documentation!

Note This documentation is not meant to be static and is a collaborative effort of the total team. If you spot anything wrong or would like to add information to the documentation, please send us a pull request or file an issue in our repository.

Chapter 1

Get CTSM

That's where you should start.

NordicESMhub maintains a CTSM repository with all the configuration files for running on machines in the Nordics. For now we support:

- saga (sigma2, Norway)
- fram (sigma2, Norway)

If your machine is not in the list and you would like us to support it, please contact us.

For compatibility, load git version 2.23.0 or higher on your super computer

```
git/2.23.0-GCCcore-8.3.0
```

1.1 How to get CTSM (for users)

To get the FATES EMERALD platform version, CLONE from NordicESM hub

```
git clone -b release-clm5.0 https://github.com/NordicESMhub/ctsm.git
${HOME}/ctsm_fates_emerald
```

In this example we are checking out the release-clm5.0 tag and create a new local branch (recomended). The destination of the checkout is a directoy (e.g. ctsm_fates_emerald) in our home directory.

1.1.1 How to get a specific branch

Change into the created ctsm directory

```
cd ${HOME}/ctsm_fates_emerald
```

Check all existing branches

```
git branch --all
```

To checkout the FATES EMERALD platform (in this example release 2.0.1) into a new local branch (e.g. new_branch_name)

```
git checkout release-emerald-platform2.0.1 -b new_branch_name
```

For later reference, it is usefull to choose new_branch_name according to function and include the version and your username, e.g. username_release-emerald-platform2.0.1.

To fetch the proper externals (CIME, FATES, etc.) run

```
./manage_externals/checkout_externals
```

from the main ctsm directory (we are going to call this $CTSM_ROOT from now on). All should be set by this and you should be able to create your first case.

1.1.2 Which branch do I run?

1.2 How to get CTSM (for developers)

Dependent on which project you wish to contribute to you might want to start your development from different versions of CTSM. For the CLM-Norway team we have to mainly two startingpoints:

- The NordicESM-hub (note that this is a project for developers in the Nordics)
- The latest version of the original CTSM (this is the original version of CTSM developed by NCAR)

1.2.1 From the NordicESM-hub

Follow the steps above, but checkout the fates_emerald_api instead

```
git checkout fates_emerald_api -b new_branch_name
```

For later reference, it is usefull to choose new_branch_name according to function and include the version and your username, e.g. username_fates_emerald_api. After checking out the externals, change to cime directory and create your own branch to record all your changes

```
cd externals/cime
git checkout -b username_cime
```

Change to fates directory and create your own branch to record all your changes

```
cd externals/fates
git checkout -b username_fates
```

If you do not create your own branch for “cime” and “fates”, running “./manage_externals/checkout_externals”, will overwrite your previous “cime” and “fates”. You should be ready to create your first case.

1.2.2 From the latest version of CTSM

This tutorial assumes that you are logged into one of the clusters (fram or saga) at sigma2. For access to those see (future referance to prerequisites section).

Start from your home folder

```
cd
```

Clone CTSM from ESCOMP

```
git clone --origin escomp https://github.com/ESCOMP/CTSM.git CTSM
```

Change into the new directory

```
cd CTSM
```

Create a local branch

```
git checkout master -b my_branch_name
```

For later reference, it is usefull to choose my_branch_name according to function and include the version and your username.

To fetch the proper externals (CIME, FATES, etc.) run

```
./manage_externals/checkout_externals
```

1.2.3 Porting of cime

Now you need to add machine specifics for the norwegian clusters. This can be done in two ways(check the original documentation for a more detailed explanation):

- 1. You can replace some default configuration files with configuration files that contain details for these clusters.
- 2. You can create a *.cime* folder with the machine configurations under your home diretory.

For method (1) above execute the following steps:

```
cd cime/config/cesm/machines
```

Delete the default files

```
rm config_machines.xml config_batch.xml config_compilers.xml
```

Fetch replacementfiles from this repository

```
git init
git remote add origin https://github.com/gunnartl/config_files_sigma2.git
git pull origin main
```

For method (2) above:

Clone this repository and consult the *README.md* file for details for making a new case.

Chapter 2

Setup CTSM

2.1 Accounting

For running CTSM, you usually need to have access to a High-Performance Computer.

If you are working in Norway, you can use either saga, fram, or betzy (depending on demands on parallel computing) and need to be part of an active account:

- Notur

To check which project you can use:

```
projects
```

This will return something like:

```
nn2806k
nn1000k
```

In the example above, two projects can be used (nn2806k and nn1000k). Then make sure you choose the right project when running CTSM.

Recommended spin-up routine for CTSM

2.1.1 Why do we do a spin-up?

- The idea behind doing a so-called ***spin-up*** is to reach an equilibrium state in your model.
- Typically any perturbation (e.g. 2x CO2) experiment starts from such a state.
- Running our perturbed system into equilibrium again will give you the maximum impact of the perturbation.

2.1.2 Which variables are important to check?

- Since we are interested in the equilibrium state of our system, the "slowest" variables matter the most (e.g. terrestrial soil carbon).
- If you are running in BioGeoChemistry (BGC) mode, it is important to check the "health" of your vegetation.
- The variable directly associated with the health is Total Leaf Area Index (TLAI).
- If this is 0 your vegetation is dead and something went really wrong.

2.1.3 How long should a spin-up be?

- The overall duration (years before an equilibrium state is reached) varies.
- As mentioned above, the variables with the slowest turnover rate will define the time you'd need to spend for your spin-up.
- The rule of thumb for single-cell simulation is the higher the latitude the longer the spin-up.
- Spinning up a single cell case in the Tropics will approximately take 2x100 years after which the terrestrial carbon pools are in equilibrium (Fig.1).
- In the polar regions, it will take a couple of 1000 years!

2.1.4 What does equilibrium mean?

- Typically you'd cycle through a range of given meteorological years - referred to as time-slice (e.g. 1990-2000).
- After an initial rapid increase/decrease in your variables, you will start to see the interannual variability of the chosen time slice (wiggling).
- Equilibrium is reached as soon as consecutive time slices differ by less than a threshold that you have to choose (< 0.1-1%).

2.1.5 How to perform a spin-up?

- Usually, we do a two-staged spin-up.
- See the necessary XML change commands below.

1. Accelerated decomposition (reduced size of carbon pools)

```
./xmlchange CLM_FORCE_COLDSTART="on"
./xmlchange CLM_NML_USE_CASE="2000_control"
./xmlchange DATM_CLMNCEP_YR_START="1991"
./xmlchange DATM_CLMNCEP_YR_END="2010"
./xmlchange DATM_PRESAERO="clim_2000"
./xmlchange CCSM_CO2_PPMV="369."
./xmlchange STOP_OPTION="nyears"
./xmlchange RUN_REFDATE="0001-01-01"
./xmlchange RUN_STARTDATE="0001-01-01"
#Additional xml changes for AD mode
./xmlchange CLM_ACCELERATED_SPINUP="on"
#Number of years for spin-up
#For Tropics only STOP_N=100 is needed!
./xmlchange STOP_N=100
#Frequency of restart files (1/4 of STOP_N)
./xmlchange REST_N=25
#Setting wall clock time (on saga 5 years / 30 min)
#./xmlchange --subgroup case.st_archive JOB_WALLCLOCK_TIME=36:00:00
./xmlchange JOB_WALLCLOCK_TIME=36:00:00
./case.setup
./preview_namelists
```

2. Normal spin-up (original size of carbon pools)

```
# Switch off cold start
./xmlchange CLM_FORCE_COLDSTART="off"
./xmlchange CONTINUE_RUN="FALSE"
```

```
#Re-use directory but re-define reference time
./xmlchange RUN_REFDATE="0401-01-01"
./xmlchange RUN_STARTDATE="0401-01-01"
#Additional xml changes for AD mode
./xmlchange CLM_ACCELERATED_SPINUP="off"
./xmlchange STOP_N=100
#Frequency of restart files (1/4 of STOP_N)
./xmlchange REST_N=25
#Setting wall clock time
#./xmlchange --subgroup case.st_archive JOB_WALLCLOCK_TIME=36:00:00
./xmlchange JOB_WALLCLOCK_TIME=36:00:00
./case.setup
#Point to restart file (IMPORTANT: Do not forget the single-quotation marks '<path
 to restart file>')
echo "finidat = '${restart_file}'" >> user_nl_clm
echo "Restart file: ${restart_file}"
```

Chapter 3

Run CTSM

NB! this example is connected to project nn2806k (for your own project, change the project code. To see available projects and resources, use cost -p):

```
export CESM_ACCOUNT=nn2806k
```

3.1 Run your very first CTSM case

LOAD externals of CTSM (FATES and so on; only necessary first time), in folder ~/ctsm. If you are updating FATES go here first: (https://github.com/NordicESMhub/ctsm-dev/blob/master/Updating_FATES.md)

```
./manage_externals/checkout_externals
```

navigate to ~/ctsm/cime/scripts/

3.2 Inputdata

(only first time or whenever it disappears in your workdir i.e. 45 days)

```
cd ~/ctsm/cime/scripts
./link_dirtree $CESM_DATA /work/users/$USER/inputdata
```

3.3 Make a case

```
./create_newcase --case ~/cases/I2000Clm50BgcCruGs  --compset I2000Clm50BgcCruGs
 --res f19_g17 --machine abel --run-unsupported --project $CESM_ACCOUNT
```

navigate to ~/cases/I2000Clm50BgcCruGs

3.3.1 1) check the configuration

```
./xmlquery --l #(--l list --f file)
```

eg

```
./xmlquery STOP_OPTION
```

3.3.2 2) Change configuration

- For instance, to change the duration of a simulation to 5 days:

```
./xmlchange STOP_OPTION=ndays #(nyears, nmonths)
./xmlchange STOP_N=5 #(then 5 days)
```

or edit the xml files is another way to change these parameters (**not recommended**).

3.3.3 3) setup case

```
./case.setup  #(--reset)
```

3.3.4 4) edit user_nl_clm

add this below

```
hist_mfilt=5 #(number of output files)
hist_nhtfrq=-24 #(means daily outputs)
```

hist_mfilt allows you to specify the number of output files and *hist_nhtfrq* the frequency; here *-24* means daily outputs.

3.3.5 5) case build

```
./case.build
```

Remark: if your build fails or if you make changes and need to rebuild, make sure you clean the previous build:

```
./case.build --clean
```

3.3.6 6) run case

```
./case.submit
```

3.4 Run fates

NB! Fates is not automatically checked out with the latest version (as it is still under development), and this has to be done manually.

Follow https://github.com/NordicESMhub/ctsm-dev/blob/master/Updating_FATES (based on https://github.uio.no/huit/clm5.0_notes/issues/26 and https://github.com/ESCOMP/ctsm/wiki/Protocols-on-updating-FATES-within-CTSM)

```
./create_newcase --case ../../../ctsm_cases/fates_f19_g17 --compset
2000_DATM%GSWP3v1_CLM50%FATES_SICE_SOCN_MOSART_SGLC_SWAV --res f19_g17 --machine
abel --run-unsupported --project $CESM_ACCOUNT
```

3.5 Run a single cell case

CLM supports running using single-point or regional datasets that are customized to a particular region.

In the section below we show you how to run ready to use single-point configurations (out of the box) and then show you how to create your own dataset for any location of your choice.

3.5.1 Out of the box

To run for the Brazil test site do the following:

```
export CESM_ACCOUNT=nn2806k

./create_newcase -case ~/cases/testSPDATASET -res 1x1_brazil -compset
I2000Clm50SpGs  --machine abel --run-unsupported --project $CESM_ACCOUNT
```

Remark: make sure you set **CESM_ACCOUNT** to your project.

Change configuration

```
./xmlchange NTASKS=1 #(number of CPU's, can be increased if excitation error)
```

3.6 Run a regional case

Here, the regional case procedure is explained for running the model over Fennoscandia and Svalbard region at 0.25 degree spatial resolution on Abel.

Required input files (domain and surface data) are already produced and stored under the main inputdata directory. So, start with linking the main inputdata files to your working directory.

```
cd ~/ctsm/cime/scripts
./link_dirtree $CESM_DATA /work/users/$USER/inputdata
export MYCTSMDATA=/work/users/$USER/inputdata
```

3.6.1 Definitions

As mentioned earlier in this documentation, you need to define your project account (e.g. nn2806k) before running the model. To be sure that you are using the correct version of netcdf and practical purposes, we will export the paths and some environment variables :

```
export CESM_ACCOUNT=nn2806k
export PROJECT=nn2806k
export CTSMROOT=/cluster/home/$USER/ctsm

export INC_NETCDF=/cluster/software/VERSIONS/netcdf.intel-4.3.3.1/
export LIB_NETCDF=/cluster/software/VERSIONS/netcdf.intel-4.3.3.1/
export NETCDF_ROOT=/cluster/software/VERSIONS/netcdf.intel-4.3.3.1
```

3.6.2 Make a case

To run the model with Satellite Phenology mode, do the following

```
cd $CTSMROOT/cime/scripts
./create_newcase -case ~/cases/SCA_SpRun -res CLM_USRDAT -compset I2000Clm50SpGs
--machine abel --run-unsupported --project $CESM_ACCOUNT
```

Here, the compset **I2000Clm50SpGs** initialize the model from 2000 conditions with GSWP3 atmospheric forcing data. If you want to run CTSM with BGC mode and force with CRU NCEP v7 data set, use **I2000Clm50BgcCruGs** compset. To see available compset aliases and their long names, type the following :

```
$CTSMROOT/cime/scripts/query_config --compsets clm
```

Note that you can use only the compsets with **SGLC** (Stub glacier (land ice) component) for single point and regional cases. Otherwise, simulation fails.

3.6.3 Set up the case

Now, we are setting up our test simulation for three years between 2000-2002.

```
cd ~/cases/SCA_SpRun
export GRIDNAME=0.25_SCA
export SCA_DATA=$MYCTSMDATA/share/domains/
export ATMDOM=domain.0.25x0.25_SCA.nc

./xmlchange DIN_LOC_ROOT=$MYCTSMDATA
./xmlchange ATM_DOMAIN_PATH=$SCA_DATA,LND_DOMAIN_PATH=$SCA_DATA
./xmlchange ATM_DOMAIN_FILE=$ATMDOM,LND_DOMAIN_FILE=$ATMDOM
./xmlchange CLM_USRDAT_NAME=$GRIDNAME
./xmlchange STOP_OPTION=nyears
./xmlchange STOP_N=3
./xmlchange NTASKS=16
./xmlchange JOB_WALLCLOCK_TIME="03:59:00"
./xmlchange PROJECT=$CESM_ACCOUNT
./xmlchange RUN_STARTDATE="2000-01-01"
./xmlchange RESUBMIT="0"
./xmlchange DATM_CLMNCEP_YR_ALIGN="2000"
./xmlchange DATM_CLMNCEP_YR_START="2000"
./xmlchange DATM_CLMNCEP_YR_END="2002"
./xmlchange DOUT_S="FALSE"
./xmlchange GMAKE_J="8"
./case.setup
./preview_run
```

3.6.4 Customizing the CLM namelist

In order to define the location of surface data file for our Scandinavia region, we add the **fsurdat** setting to the land model's namelist.

```
cat << EOF > user_nl_clm
fsurdat="$MYCTSMDATA/lnd/clm2/surfdata_map/surfdata_0.25_SCA_hist_16pfts_Irrig_CMIP6_simyr2000_c
EOF
```

If you run your simulation with **BGC** mode, then you need to set another surface data map as follows:

```
cat << EOF > user_nl_clm
fsurdat="$MYCTSMDATA/lnd/clm2/surfdata_map/surfdata_0.25_SCA_hist_78pfts_CMIP6_simyr2000_c190819
EOF
```

If you want to print out only selected variables in hourly resolution in the model output files for each year:

```
cat << EOF > user_nl_clm
fsurdat="$MYCTSMDATA/lnd/clm2/surfdata_map/surfdata_0.25_SCA_hist_16pfts_Irrig_CMIP6_simyr2000_c
hist_empty_htapes = .true.
hist_fincl1 = 'TSA', 'TSKIN', 'EFLX_LH_TOT', 'FSH', 'WIND', 'TWS', 'SNOWLIQ',
'SNOWICE', 'SNOW_DEPTH', 'TSOI', 'H2OSOI'
hist_nhtfrq = -1
hist_mfilt = 365
EOF
```

Note that the variable names (e.g. 'TSA', 'TSKIN', 'EFLX_LH_TOT', etc.) are following the CESM name convention.

3.6.5 Build and submit the case

Final step is to build the case and submit the job to the queue.

```
./case.build
./case.submit
```

Chapter 4

Recommended spin-up routine for CTSM

Single cell example: Brazil

Error opening image file: [Errno 2] No such file or directory: '/__w/rtd-github-pages/rtd-github-pages/docs/doc/figures/spinup_brazil_example.png'

Figure 4.1. Fig.1 Example spin-up for single cell case Brazil. The red boxes indicate 100 years of accelerated spin-up.

4.1 Why do we do a spin-up?

- The idea behind doing a so-called ***spin-up*** is to reach an equilibrium state in your model.
- Typically any perturbation (e.g. 2x CO2) experiment starts from such a state.
- Running our perturbed system into equilibrium again will give you the maximum impact of the perturbation.

4.2 Which variables are important to check?

- Since we are interested in the equilibrium state of our system, the "slowest" variables matter the most (e.g. terrestrial soil carbon).
- If you are running in BioGeoChemistry (BGC) mode, it is important to check the "health" of your vegetation.
- The variable directly associated with the health is Total Leaf Area Index (TLAI).
- If this is 0 your vegetation is dead and something went really wrong.

4.3 How long should a spin-up be?

- The overall duration (years before an equilibrium state is reached) varies.
- As mentioned above, the variables with the slowest turnover rate will define the time you'd need to spend for your spin-up.
- The rule of thumb for single-cell simulation is the higher the latitude the longer the spin-up.
- Spinning up a single cell case in the Tropics will approximately take 2x100 years after which the terrestrial carbon pools are in equilibrium (Fig.1).

- In the polar regions, it will take a couple of 1000 years!

4.4 What does equilibrium mean?

- Typically you'd cycle through a range of given meteorological years - referred to as time-slice (e.g. 1990-2000).
- After an initial rapid increase/decrease in your variables, you will start to see the interannual variability of the chosen time slice (wiggling).
- Equilibrium is reached as soon as consecutive time slices differ by less than a threshold that you have to choose (< 0.1-1%).

4.5 How to perform a spin-up?

- Usually, we do a two-staged spin-up.
- See the necessary XML change commands below.

1. Accelerated decomposition (reduced size of carbon pools)

```
./xmlchange CLM_FORCE_COLDSTART="on"
./xmlchange CLM_NML_USE_CASE="2000_control"
./xmlchange DATM_CLMNCEP_YR_START="1991"
./xmlchange DATM_CLMNCEP_YR_END="2010"
./xmlchange DATM_PRESAERO="clim_2000"
./xmlchange CCSM_CO2_PPMV="369."
./xmlchange STOP_OPTION="nyears"
./xmlchange RUN_REFDATE="0001-01-01"
./xmlchange RUN_STARTDATE="0001-01-01"
#Additional xml changes for AD mode
./xmlchange CLM_ACCELERATED_SPINUP="on"
#Number of years for spin-up
#For Tropics only STOP_N=100 is needed!
./xmlchange STOP_N=100
#Frequency of restart files (1/4 of STOP_N)
./xmlchange REST_N=25
#Setting wall clock time (on saga 5 years / 30 min)
#./xmlchange --subgroup case.st_archive JOB_WALLCLOCK_TIME=36:00:00
./xmlchange JOB_WALLCLOCK_TIME=36:00:00
./case.setup
./preview_namelists
```

2. Normal spin-up (original size of carbon pools)

```
# Switch off cold start
./xmlchange CLM_FORCE_COLDSTART="off"
./xmlchange CONTINUE_RUN="FALSE"
#Re-use directory but re-define reference time
./xmlchange RUN_REFDATE="0401-01-01"
./xmlchange RUN_STARTDATE="0401-01-01"
#Additional xml changes for AD mode
./xmlchange CLM_ACCELERATED_SPINUP="off"
./xmlchange STOP_N=100
#Frequency of restart files (1/4 of STOP_N)
./xmlchange REST_N=25
```

```
#Setting wall clock time
#./xmlchange --subgroup case.st_archive JOB_WALLCLOCK_TIME=36:00:00
./xmlchange JOB_WALLCLOCK_TIME=36:00:00
./case.setup
#Point to restart file (IMPORTANT: Do not forget the single-quotation marks '<path
 to restart file>')
echo "finidat = '${restart_file}'" >> user_nl_clm
echo "Restart file: ${restart_file}"
```

- genindex
- modindex
- search