MANUAL: A Gaussian process emulator for simulating ice sheet-climate interactions on a multi-million year timescale: CLISEMv1.0

Contact: Van Breedam Jonas (jonas.van.breedam@vub.be)

This manual describes how to use the Gaussian Process emulator CLISEMv1.0 as used in the study presented in GMDD (Van Breedam et al., 2021). The code is written in R and tested with software versions 3.4.4 and 3.5.1. NetCDF libraries (ncdf4) might have to be replaced for other R model versions. The source code is stored in file *pca_emul.R* which was originally written by Michel Crucifix (https://github.com/mcrucifix/gp) and is extended for this study. All routines and input data needed to reproduce the emulated temperature and precipitation forcing fields as input to the ice sheet model are explained below.

CLISEMv1.0 is calibrated based on a number of precursor climate model runs with the HadSM3 climate model for varying orbital parameters, atmospheric CO_2 concentrations and ice sheet geometries. Basically, monthly temperature and monthly precipitation fields are emulated from the global climate model on the ice sheet model grid, since they are key in the determination of the mass balance over the ice sheet. These variables then serve as input to a 3-D thermomechanical ice sheet model, where the mass balance is calculated based on a Positive Degree Day Model. The interaction between the ice sheet model and the emulator consists of exchanging the ice volume and/or ice area (ice sheet parameter) to the emulator and the precipitation and temperature fields to the ice sheet model.

The emulation process consists of a calibration phase during which the optimal parameters for the nugget and the length scales are chosen, and a prediction phase which is used during the coupled ice sheet-emulator runs. The calibration phase makes use of the climatic information of the precursor climate model runs and the file containing the forcing (orbital parameters, CO_2 values and ice volume/ice area) for these precursor climate model runs.

The output data from the precursor climate model runs are stored in three main directories for the three different emulators (EMULATOR_70, EMULATOR_100a and EMULATOR_100b), where 70 and 100 indicates the number of precursor climate model runs performed with the respective emulator:

- output_emulator_70
- output_emulator_100a
- output_emulator_100b

The scripts to compute the forcing for the ice sheet model are stored in 10 different directories (three different emulators calibrated on three different ice sheet parameter definitions + the simulations including the emulator uncertainty):

- emulator_70_icevolume
- emulator_70_icearea
- emulator_70_icevolar
- emulator_100a_icevolume
- emulator_100a_icearea
- emulator_100a_icevolar
- emulator_100b_icevolume

- emulator_100b_icearea
- emulator_100b_icevolar
- emulator_100b_icevolume_uncertainty

Each directory (except for emulator_100b_icevolume_uncertainty) contains the following scripts for the calibration and prediction of temperature and precipitation for each month:

- **calibration.R**
  - uses *xaem_aism_<month>.R* to read the temperature and precipitation data form the precursor climate model runs
  - uses the file *emul_input_5variables.txt* to read the forcing parameters (e sinϖ, e cosϖ, ε, CO_2, ice volume/ice area) of the precursor climate model runs
  - uses the file *emul_input_6variables.txt* to read the forcing parameters (e sinϖ, e cosϖ, ε, CO_2, ice volume, ice area) of the precursor climate model runs, when the ice sheet parameter is defined by both the ice area and the ice volume
  - writes the scripts *E_list_temp.Rda* and *E_list_precip.Rda* containing the mean temperature and precipitation data and the PCA scores for each of the principal components that are taken into account and for all months.
  - writes *temp_<month>_all_aism.nc* and *precip_<month>_all_aism.nc* containing all temperature and precipitation data from the precursor climate model runs as a diagnostic.
- **prediction.R**
  - uses the files *E_list_temp.Rda* and *E_list_precip.Rda* to get the PCA scores for the calibrated temperature and precipitation respectively for all months
  - uses the file *emul_input.txt* to read the forcing parameter time series (e sinϖ, e cosϖ, ε, CO_2). The file *emul_input.txt* contains values for the forcing parameters between 40 Myr and 33 Myr ago at an interval of 250 years. In the standard experiments, the coupling time is 1000 years. Hence, the data are read each fourth value from the file *emul_input.txt*.
  - uses the file *EMULICE* to read the ice sheet parameter (ice volume, ice area, or both ice volume and ice area). The current time index and ice volume and/or ice area are written to the file *EMULICE* by the ice sheet model. The file *emulice_full_dt<timestep>.txt* (where the timestep equals 500, 1000 or 2000) contains the full time series matrix of the ice sheet size parameter. This file can be used to calculate the temperature and precipitation forcing during the whole length of the experiments.
  - Writes the data *MONTHLY_TEMP.nc* and *MONTHLY_PRECIP.nc* that serves as input to the ice sheet model.

Before making predictions, the emulator needs to be calibrated by executing the calibration.R script. To execute the above described scripts to calculate the temperature and precipitation time series, we make use of the shell script meta_prediction.sh. This script reads the file *emulice_full_dt<timestep>.txt* and executes the script *prediction.R* for each timestep (500, 1000 or 2000 years) during the run. Only EMULATOR_100a and EMULATOR_100b contain time series for $\Delta t = 500$ and $\Delta t = 2000$ years. The outcome are netCDF files of temperature (*monthly_temp_dt<timestep>_<index>.nc*) and precipitation (*monthly_precip_dt<timestep>_<index>.nc*) at each timestep (for a run of 3 Myr length, this

script will produce 3000 netCDF files containing the monthly temperature and 3000 netCDF files containing the precipitation)

The (January temperature) uncertainty simulations are performed with the scripts stored in emulator_100b_icevolume_uncertainty. The uncertainty with respect to January temperatures are emulated using the scripts *prediction_jan_montecarlo_dt500.R* and *prediction_jan_montecarlo_dt2000.R*. First the script *calibration_jan.R* has to be executed to perform the calibration and produce the necessary input files for the temperature prediction. We make again use of the shell scripts *prediction_jan_mc_dt500.R* and *prediction_jan_mc_dt2000.R* to calculate the forcing time series for the entire length of the simulations. This creates the files *temp_jan_dt500_<index>.nc* and *temp_jan_dt2000_<index>.nc*.

Overview of initial scripts in each folder:

calibration.R	Calibrate the emulator for temperature and precipitation for all months
prediction.R	Make temperature and precipitation predictions for all months
xaem_aism_<month>.R	Read all the climatic input data from the precursor climate model runs
meta_prediction.sh	Run the script to calculate temperature and precipitation forcing files for the entire length of the experiments
pca_emul.R	Source code for the PCA emulator

Overview of the initial files in each folder:

emul_input.txt	Input forcing for the simulations
emul_input_5variables.txt	Input forcing of the precursor climate model runs
emulice_full_dt1000.txt	ice sheet parameter given at an interval of 1000 years

Overview of initial scripts in the folder for the Monte Carlo simulations (January temperature uncertainty):

prediction_jan_montecarlo_dt500.R	Make January temperature predictions with the inclusion of the uncertainty for a coupling time step of 500 years
prediction_jan_montecarlo_dt2000.R	Make January temperature predictions with the inclusion of the uncertainty for a coupling time step of 2000 years
meta_prediction_mc_dt500.sh	Run the script to calculate January temperature forcing files for the entire length of the experiments (dt=500 years)
meta_prediction_mc_dt2000.sh	Run the script to calculate January temperature forcing files for the entire length of the experiments (dt=2000 years)
xaem_aism_janR	Read all January temperatures from the precursor climate model runs
pca_emul.R	Source code for the PCA emulator

Overview of initial files in the folder for the Monte Carlo simulations (January temperature uncertainty):

emul_input.txt	Input forcing for the simulations
E_stochastic.Rda	Contains the calibrated data for January temperature
emulice_full_dt500.txt	ice sheet parameter given at an interval of 500 years
emulice_full_dt2000.txt	ice sheet parameter given at an interval of 2000 years