

Summary of the 1st SMILE meeting at EGU 2019

The 1st SMILE (Single Model Initial-Condition Large Ensemble) meeting was held at EGU 2019 on April 10th as a splinter meeting and was open to anyone working with large ensembles. Around 30 participants attended the meeting in Vienna, three participants from Ouranos and NCAR joined the discussion remotely via video connection.

Part I - current state

The first part started with three short presentations introducing different large ensembles, including the lessons learned while creating these ensembles.

Andrea Dittus (University of Reading): The SMURPHS ensemble

The ensemble consists of 20 historical simulations with the HadGEM3-GC3.1 model (135 km atmosphere, 1° ocean). Aerosol forcing is scaled by 5 different factors (0.2, 0.4, 0.7, 1.0, 1.5), 4 realisations with initial conditions from PiControl are produced for each scaling ("Macro" initialization). The ensemble is using CMIP6 historical forcing.

Motivation for the ensemble:

- Uncertainties in the magnitude of historical aerosol forcing are very large
- Role of aerosol forcing in global climate variability is still an open question

Output:

- Available on request (a.j.dittus@reading.ac.uk)
- Monthly CMIP6 requested variables, some daily and sub-daily variables

Challenges:

- Experiments with state-of-the-art climate models are expensive, how to balance sampling internal variability and wide range of aerosol forcing?
- Need to balance length of simulations, model complexity (sophisticated aerosol scheme), number of ensemble members
- Selecting variables to save, data post-processing and archival

Nicola Maher (Max Planck Institute for Meteorology): The Max Planck Institute Grand Ensemble (MPI-GE)

- The ensemble consists of 100 realisations with the MPI-ESM1.1-LR (~1.8° atmosphere, 1.5° ocean) for each of 5 forcings (historical, 1% CO₂, RCP2.6, RCP4.5, RCP8.5). The initial conditions are sampled from a PiControl run ("Macro" initialization). The ensemble is using CMIP5 forcing.

Motivation:

- Multiple scenarios to compare multiple futures and investigate if they differ when taking internal variability into account
- 100 realisations to assess 1 in 100 year events at each time-step

Output:

- Available in CMOR format (<https://www.mpimet.mpg.de/en/grand-ensemble/>)
- Problem reports and questions can be addressed to grandensemble@mpimet.mpg.de

Lessons learned:

- Ask the communities before deciding which output to save
- Ensure that control simulation is equilibrated before initialising

Challenges:

- Data storage
- Data sharing

Raul R. Wood (LMU Munich, Department of Geography): CRCM5-LE regional climate model LE

The ensemble consists of 50 realisations of the regional CRCM5 model (12km atmosphere resolution) with two domains: Europe and Québec. Boundary conditions are provided by global CanESM2 simulations ("Macro/Micro" initialization). The 50 members cover the period 1955-2099 and use CMIP5 historical and RCP8.5 forcing.

Motivation:

- Confirm knowledge on whether and how climate change contributes to higher magnitudes and frequencies of extreme events
- Distinguish between the effects of internal variability and a 'clear' climate change signal
- Improve methods to analyse hydro-meteorological extreme events and provide robust estimates of flood frequency

Output:

- Available in netCDF format <http://www.climex-project.org> (1 variable, 1 member, 1 month per file)
- Variables with 1 hour to 1 day resolution

Lessons learned:

- Post-processing is very time consuming
- Only possible with access to big data storage and cluster computing
- Long-term plan for data storage and access is needed
- Demand for intensive collaboration among the single working groups

Open discussion for part I:

This summary is listing some topics that were discussed, but may not reflect the full discussion for each topic.

- **LE-MIP** (or variability-MIP): should we, as a community, aim for a LE-MIP?
 - Topics that could be covered: variability, extreme events, ...
 - Should this build on the existing simulations or include new, coordinated experiments?

- A **large ensemble review paper** summarising the advances in the field over the last years would be useful
 - US Clivar working group is planning to submit a perspectives paper
- A **special issue on large ensembles** based on the EGU session has been proposed for *Earth System Dynamics*. This issue is open to submissions which exploit the new opportunities offered by large ensembles and explore how a combined analysis of the existing large ensembles can advance our knowledge in different fields. The special issue particularly invites submissions that use new methods to investigate these topics.
 - Tentative contributions, including a title and author list, can be sent to Nicola Maher (nicola.maher@mpimet.mpg.de), Ralf Ludwig (r.ludwig@lmu.de), or Sebastian Milinski (sebastian.milinski@mpimet.mpg.de)
- So far, **only few studies use multiple large ensembles**; with the output of several large ensembles now publicly available, studies with multiple large ensembles are now possible and could improve our understanding of model differences and robustness of results across models. The collection of output created by the US CLIVAR working group on large ensembles provides access to the output from several models:

https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm4.CLIVAR_LE.html
- **Storing and sharing of output** is one of the main challenges encountered by groups running large ensembles. In the discussion we identified some user preferences and possible approaches to data handling in the future.
 - general preference for CMOR-like output format (single files per variable to avoid downloading big files with many variables if only one variable is needed)
 - share processed climate indices (who does the processing? How do we agree on definitions for the indices?)
 - should we prioritize which fields are made available first?
 - pre-industrial controls must be made available: this also allows groups to run their own extra ensemble members and subsets (given restart files are saved occasionally)
 - regridded output for the ocean is useful for some users; however, closed budgets can only be computed on the native grid
- We should have a **community survey** to get a representative picture of the needs and plans of the large ensemble community. This survey should cover potential joint efforts, including joint applications for funding, distributing computing load across supercomputing centers, e.g., CESM LE and MPI GE had some ensemble members contributed by outside groups (needs model experimental setups that are easily transferable), explore different options to facilitate the storing and sharing of output, and using alternative analysis approaches like the CMIP6 ESMVal tool or CVDP LE (in development).

Part II - applications and future plans

The second part started with two short presentations introducing the plans for a new large ensemble and exploring perspectives for large ensemble modelling.

Keith Rodgers (ICCP, Busan, South Korea): Future Plans for a GCM Large Ensemble

- IBS & NCAR collaboration to run a 100 member large ensemble with CESM2.1 for 1850-2100
- Scenario and initialization procedure are still under discussion
- Intention to save high-frequency output from both ocean and atmosphere to facilitate research on extreme events
- Open access to output for international community

Erich Fischer (ETH Zurich): Towards a better quantification of changes in extremes?

- Large ensembles help to quantify return periods, understand model differences, quantify and narrow uncertainties, and detecting and attributing signals.
- In the context of extreme events, large ensembles help to identify small signals in dynamical drivers and allow a better sampling of the tails of the distribution.
- Future directions could include
 - targeted downscaling to bridge the gap between low-resolution large ensemble and small ensembles with high-resolution models
 - large ensemble boosting (targeted re-initialisation to investigate rare events with limited resources)
 - developing narratives and storylines of extreme pathways
 - combining with large ensembles with observations (e.g. observational large ensembles, *McKinnon et al. 2017*)
 - testing robustness of machine learning algorithms for dynamical adjustment (*Sippel et al.*) or to produce super-ensembles (*Beusch et al.*)

Open discussion for part II:

This summary is listing some topics that were discussed, but may not reflect the full discussion for each topic.

- To be usable as **boundary conditions for RCMs**, GCM large ensembles need to provide high frequency output for specific variables that are necessary as input for RCMs. Otherwise downscaling is not possible. (e.g. for ongoing downscaling experiments at MPI-M, some of the large ensemble GCM experiments were re-run with high-frequency output)
- Do we want to have **coordinated experiments** or not?
 - SSP scenarios?
 - 1% CO₂
 - step forcing (to train emulator)
 - pacemaker experiments
 - experiment design needs to be primarily driven by a scientific question. This should also determine whether coordinated experiments should be performed or not.

- Should we establish a common reference experiment that can be used to compare different LEs of different generations (CMIP5 vs. CMIP6 forcing)?
 - 1% CO₂, 30 members? Forcing is clearly defined and will not change in CMIP6,7,...
 - Is the piControl a candidate or do we need transient forcing?
- length of piControl simulation must be long enough to provide a robust reference
- LE (low-res) vs. high-res: potential to gain new insights from combined analysis
 - CESM1.3 10 members high-res (0.1°) available in ~1.5 years
 - MPI-ESM has 10 historical members with high-res
- ESM large ensemble with prescribed GHG emissions rather than concentrations to investigate internal variability with free carbon cycle (experiments with CMIP6 historical forcing have been started with the MPI-ESM)
- include air pollution and health community - currently a gap in the LE projects
- Organising community efforts:
 - A joint large ensemble session has been proposed for the AGU Fall Meeting 2019 (lead: Keith Rodgers)
 - A 2nd SMILE meeting has been proposed as a town hall for the AGU Fall Meeting 2019 (lead: Raul R. Wood)
 - potential Large Ensemble workshop in Europe in autumn 2019 to complement the LE workshop at NCAR in July and potential AGU events in December with an event in Europe. Details will be determined together with the community. The workshop might take the form of a webinar to allow everyone to attend without the need to travel.

The SMILE mailing list can be found at <https://listserv.gwdg.de/mailman/listinfo/smile>

The 1st SMILE meeting was organised by:

Sebastian Milinski (MPI for Meteorology, sebastian.milinski@mpimet.mpg.de)

Andrea Dittus (University of Reading, Department of Meteorology, a.j.dittus@reading.ac.uk)

Flavio Lehner (NCAR, US CLIVAR, flehner@ucar.edu)

Nicola Maher (MPI for Meteorology, nicola.maher@mpimet.mpg.de)

Raul R. Wood (LMU Munich, Department of Geography, raul.wood@lmu.de)