The Energy and Resources Institute

TERI-NFA Working Paper No. 7

Earth System Model Installation at TERI HPC: Norway – India Collaboration

Authors: Ingo Bethke[1], Mats Bentsen[1] and Vidyunmala Veldore[2]

[1]Bjerknes Center for Climate Research

[2]The Energy and Resources Institute

Project Members: Eystein Jansen, Arabinda Mishra, Suruchi Bhadwal, Helge Drange, Michel dos Santos Mesquita, Ingo Bethke, Mats Bentsen, Vidyunmala Veldore, Saurabh Bhardwaj, Nehru Machineni, Martin King, M. S. Madhusoodhanan, Alexander Oltu and Torleif M Lunde

January 2013

Acknowledgements

This paper was written as part of a project "Climate change impacts assessments, including climate change, process and earth system modeling" under the Program of Activities, Framework Agreement between the Norwegian Ministry of Foreign Affairs (MFA) and The Energy and Resources Institute (TERI), briefly referred to as the Norwegian Framework Agreement (NFA). We are grateful for BCCR, BCCS (Bjerknes Center for Computational Sciences) for providing constant support and guidance throughout the present work. We also extend our thanks to WIPRO for providing the technical support in maintaining the TERI High Performance Computing system. We would like to thank the Norwegian Ministry of Foreign Affairs for funding this project. We would also like to express our thanks to Dr. Jerry Tjiputra, BCCR for his constant help and providing inputs in taking forward these activities.

Corresponding Author, [**Dr. Vidyunmala Veldore**], is an **Associate Fellow** at TERI, New Delhi.

Email: {vidyunmala.veldore@teri.res.in}

Contacts
The Energy and Resources Institute
Darbari Seth Block
India Habitat Centre
Lodhi Road
New Delhi 110 003
Tel: + 91 - 11- 24682100 / 41504900

Table of Contents

Summary:

The new generation NorESM-1M and NorESM1-ME model (without and with carbon cycle) have been successfully ported in TERI high performance computing (HPC) system. The objective of this report is to document the performance of these model configurations and the changes made to achieve the optimized results in TERI HPC. It was found that the performances in TERI-HPC are 13-simulation years/day for the low-resolution NorESM, 7-simulation years/day for the intermediate-resolution NorESM (with and without carbon cycle), and 11-simulation years/day for the standard CESM1 at intermediate resolution. The benchmarking results suggest that the TERI-HPC outperforms the Norwegian CRAY for small CPU counts. As we increase the number of CPUs the performance does not vary a lot. This technical report describes the NorESM installation process, verification and benchmarking results. In the next phase, the optimized and tested NorESM will be used to perform a 150-year historical simulation using the restart conditions from BCCR researchers. Preliminary results of the first 30 year simulation with NorESM-1ME restart conditions show that the simulation outcomes are similar to Norwegian CRAY based simulations and the climates are indistinguishable.

Scope of the report:

The purpose of this report is to document the installation of the Norwegian Earth System Model on the TERI cluster. The first two sections describe the model and computational system. Section three provides detailed information on the porting and installation of the model system while section four elaborates on performance optimization, benchmarking and verification of the installation. Proposed future activities with the model system installed at TERI are offered in section 5.

1. Introduction to the Norwegian Earth System Model

1.1 General

A climate model is a mathematical presentation of our Earth's climate system based on physical, thermodynamic, chemical, and biological principles. In general, climate models do not give an exact presentation of the real climate and the realism and fidelity of a climate model depends on its complexity, the accuracy of resolved processes, and its parameterizations of unresolved processes. For complex climate models, the model equations are typically solved numerically with the help of computers.

Earth System Model (ESM) refers to the newest generation of climate models, which has a special focus on the interaction of the natural carbon cycle with human induced emissions and with physical climate change. The ESM extends the Atmospheric Ocean General Circulation Model (AOGCMs) class by the inclusion of a fully interactive carbon cycle, which requires additional ocean and land biogeochemistry as well as atmospheric chemistry components.

The Norwegian Earth System Model (NorESM; Bentsen et al. 2012, Iversen et al. 2012) is one out of ~20 climate models that has produced output for the CMIP5 (http://cmip-pcmdi.llnl.gov/cmip5). The CMIP5 project provides a framework for coordinated climate change experiments for the next five years and includes simulations that will be used in the upcoming Fifth Assessment Report of the IPCC (http://www.ipcc.ch), due 2013/14. The development of the NorESM started under the now finalized research project NorClim (http://www.norclim.no) and is continued as part of the ongoing EarthClim project (http://gfi.uib.no/EarthClim). EarthClim has ten partners: Uni Research AS (coordinator); Bjerknes Centre at the University of Bergen; CICERO Center for Intern Climate and Environmental Research; Norwegian Meteorological Institute; Department of Geosciences, University of Oslo; Norwegian Computing Center; Norwegian Institute for Air Research; and Norwegian Polar Institute, of which the Norwegian Meteorological Institute and the Bjerknes Centre have the main responsibility for the model development. The organizational framework for NorESM is the virtual Norwegian Climate Centre (NCC) consisting of nine of the ten institutions participating in EarthClim.

1.2. Model components and configurations

NorESM consists of the four main components: atmosphere, land, ocean and sea ice. These components can be understood as individual models that run in parallel and communicate with the help of a model coupler tool. The coupler exchanges fields between the atmosphere, land and sea ice components every 30 minutes of integration (model time), and with the ocean component every day of integration. The main components contain a series of subcomponents that are tightly integrated in the code of the main components. The subcomponents share data variables with the main components and use the same grid spacing. The schematic shown in Fig. 1 illustrates the communication of NorESM's main components and the integration of selected subcomponents. In the following, the individual components are presented in more detail. NorESM is built upon the Community Earth System Model version 1 (CESM1; see http://www.cesm.ucar.edu; Gent et al., 2011) developed at the

National Center for Atmospheric Research (NCAR) and share substantial parts of model components and supporting framework.

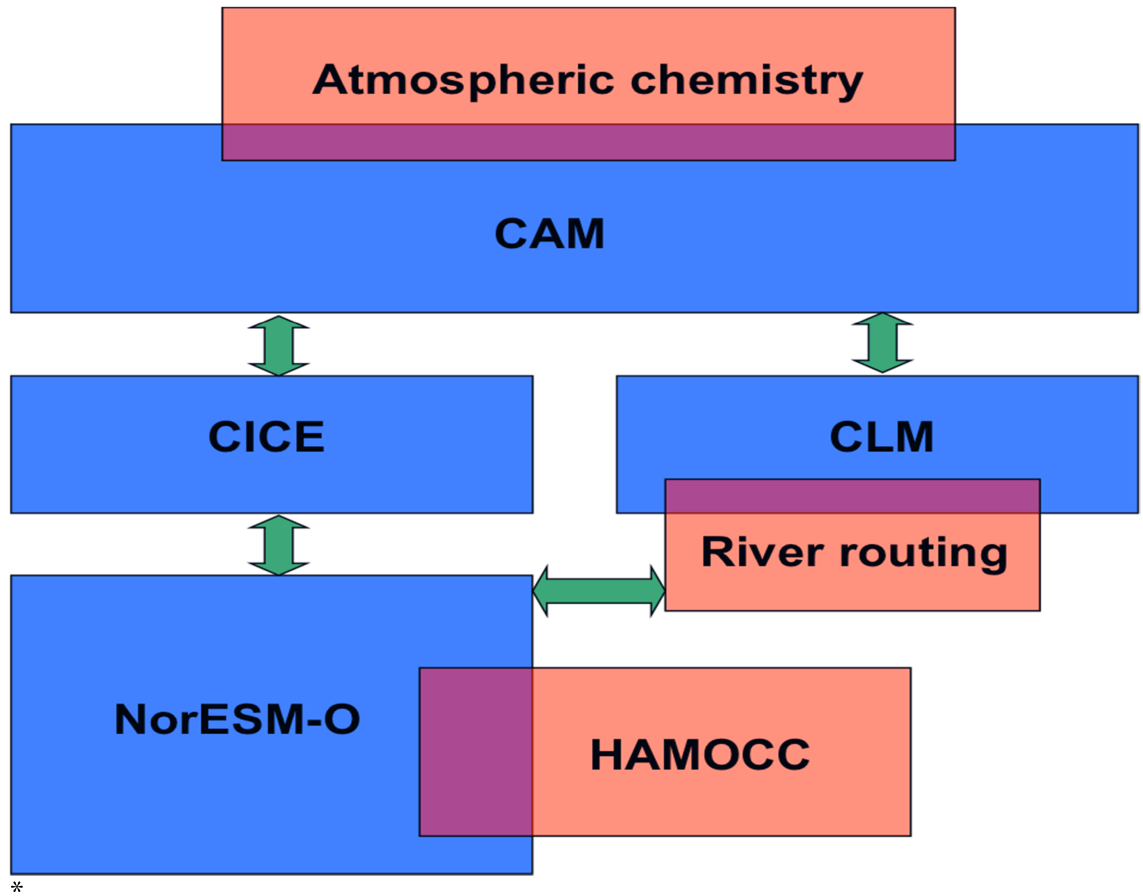

Figure 1: Schematic of NorESM. Blue boxes denote the main model components. Green arrows indicate exchange of information between the components through the model coupler. Red boxes denote selected subcomponents that are embedded in the main components.

The atmospheric component is the version 4 of the Community Atmosphere Model (CAM4; Neale et al., 2010, 2012) with modified chemistry-aerosol-cloud-radiation interaction schemes developed at the University of Oslo and the Norwegian Meteorological Institute (CAM4-Oslo; Seland et al., 2008; Kirkevåg et al., 2012). With CAM4-Oslo the indirect effect of aerosols can be simulated. The land component is the version 4 of the Community Land Model (CLM4; Oleson et al., 2010; Lawrence et al., 2011). CLM4 is configured on the same horizontal grid as the atmospheric component, except for the embedded river transport model of CLM4 that is configured on its own horizontal grid.

As in CESM1, a modified version 4 of the Los Alamos National Laboratory sea-ice model (CICE4; Hunke and Lipscomb, 2008) is used as the sea-ice component of NorESM. NCAR extensions of this model include the delta-Eddington shortwave radiation transfer (Briegleb and Light, 2007), melt pond and aerosol parameterizations (Holland et al., 2012).

The sea-ice component is configured on the same horizontal grid as the ocean component. In NorESM, an isopycnic coordinate ocean model, NorESM-O, replaces the CESM1 ocean component. NorESM-O is based on MICOM (Bleck and Smith, 1990; Bleck et al., 1992) but with significantly modified dynamical core and physical parameterizations (Bentsen et al., 2012). Embedded in NorESM-O is the ocean carbon cycle model HAMburg Ocean Carbon Cycle (HAMOCC; Maier-Reimer, 1993; Maier-Reimer et al., 2005), adapted to an isopycnic ocean model framework (Assmann et al., 2010; Tjiputra et al., 2012).

Three configurations of NorESM exists at the moment labelled NorESM1-L, NorESM1-M, and NorESM1-ME. NorESM1-L is a low resolution configuration tailored for millennium scale experiments using a 3.75º □ 3.75º atmospheric horizontal grid with spectral dynamical core (T31) and a nominal 3º oceanic horizontal grid. NorESM1-M and NorESM1-ME are medium resolution configurations using a 1.9º □ 2.5º atmospheric horizontal grid with a finite volume dynamical core and a nominal 1º oceanic horizontal grid. NorESM1-M uses prescribed concentrations of greenhouse gases and is configured without interactive carbon cycle, while NorESM1-ME is purely emission driven and includes an interactive carbon cycle.

1.3. Climate applications

The need to understand climate variability and change are gaining high priorities than anticipated. Many significant impacts have been highlighted which substantiates the influence of climate variability and change on different spatial and temporal scales. More precisely it was found that the understanding of decadal variability and shorter time scale variabilities need to be better understood, such as droughts (linkages to teleconnections), extreme short term events (frequency, intensity etc.). In the advent of increasing computational efficiency the complexities and processes represented in the climate system models or the new generation Earth System Models are also increasing. However, when we consider the coupling of atmosphere, land, ocean, cryosphere and biosphere the scales of climate variability at different temporal scales plays an important role, i.e., the low frequency variability of decadal scales should also be well represented. Though there is a definite progress in determining the interannual variabilities by many Earth System Models, even the interdecadal variabilities need to be represented well. It is also important that not only the global teleconnections need to be better represented in such models, but also the regional scale processes that are important for regional scale climate variability and change is desirable.

In order to understand the sensitivity of the climate system to different GHG concentration and to the newly developed Representative Concentration Pathways (RCPs) scenarios, a series of experiments have been carried out at the BCCR institute which is also submitted to the CMIP5 set of experiments (Fig 2). A spin-up of the NorESM model has been carried out for every horizontal resolution to ascertain that the coupled processes are stable and there is no climate drift, the historical simulations of NorESM1-M model is then carried out for three ensemble members and extensions of these simulations from historical to the CMIP5 based experiments are carried forward. The global climate projections from these models are used for providing regional climate projections over different parts of the world with the use of Weather Research Forecasting (WRF) Regional climate model.

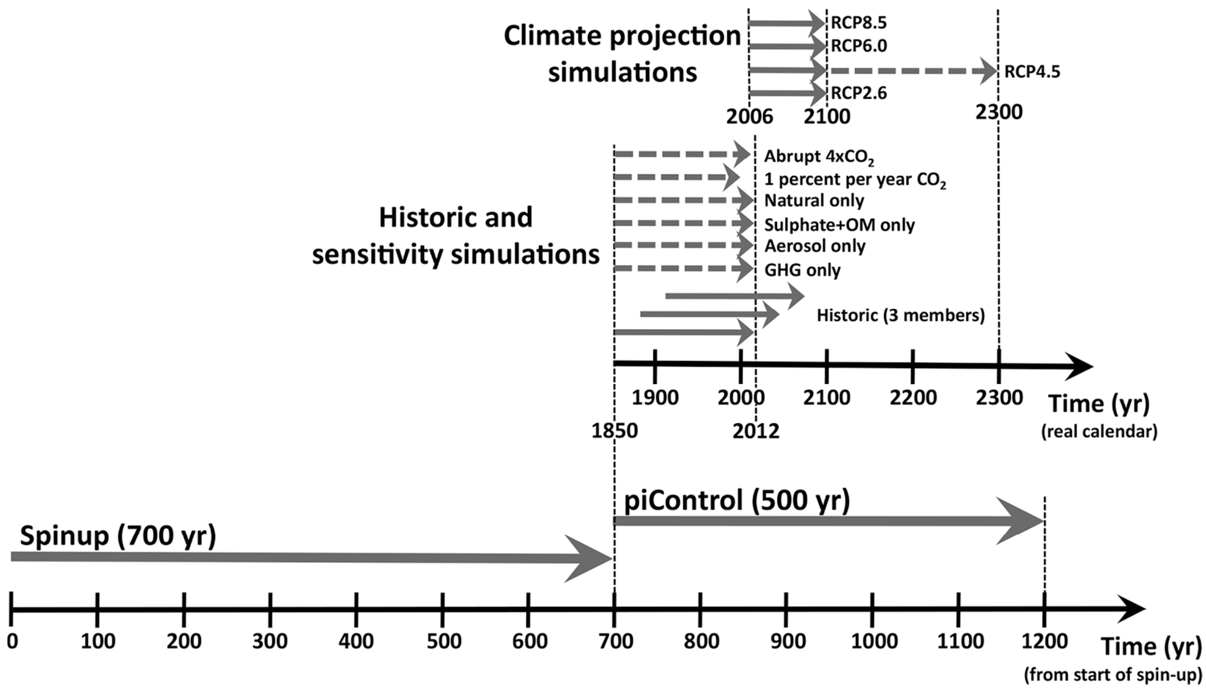

Figure 2 Schematic of the pre-industrial spinup and CMIP5 experiments carried out with NorESM1-M.

2. Computational resources for climate modelling at TERI

Climate change impacts through global in nature may affect regional scales in different ways. Thus, understanding of climate change impacts requires a cascaded assessment of climate information at different spatial and temporal scales. The Energy and Resources Institute (TERI) is a non-profit research organization endeavoring to understand climate change issues and adapt to sustainable world for the better future. TERI has been a front runner in innovation when it comes to capacity building. It has teamed up with the Bjerknes Centre for Climate Research, in Bergen, Norway. Through their Indo-Norwegian project, the TERI climate change division has been able to acquire a new high-performance computing infrastructure for the large demands of their climate modeling activities. The vision of climate modeling in TERI is "To build capacity in climate modeling skills to understand the links between global and regional climate changes and their links to policy at different scales".

The project aims "To understand the process in Earth system modeling and provide high resolution regional climate change scenarios for tropical region in order to assess climate change impacts on different sectors ". The project also aims to provide capacity building through collaborative climate research schools for early career scientists and researchers in and around South Asia. Two research schools have been conducted in the past two years (2011-2012) in (1) Global models, (2) Regional climate models. We are also preparing for our next phase of research school for 2013 with emphasis on Earth System models. In order to carry forward the project objectives TERI has increased its computing infrastructure to hex core clusters and dedicated for performing the climate runs.

2.1 High Performance Computing Lab at TERI:

TERI' computational lab for climate modeling and analysis activities started early 2007 with two Linux based servers. In 2008, within the Norwegian Framework Agreement in collaboration with Royal Norwegian Embassy India, the High Performance-Computing (HPC) lab was extended with state-of-art computing facility that can perform both global and regional climate simulations at higher speed.

The HPC is a cluster based architecture setup by WIPRO (stands for Western Indian Products Limited- a leading multinational institute in India providing information technology services). The cluster comprises of multiple compute servers connected together over a high speed interconnect. A large mesh is submitted to the Master Server, wherein it is broken into smaller sub-domains and each sub-domain is submitted to a different processor for computing. The JOB Decomposition is done within the application (Cluster Version), it is migrated to various processors using the system middleware and low latency interconnects.

A grid engine based architecture is followed to provide policy based workload management and dynamic provisioning of application workloads. Computing tasks or jobs are distributed across the grid in accordance with resource requirements for the job, user requests for the job, and administrative/managerial policies. Usage accounting data is stored and made available so that it is possible to determine what resources were used in the execution of a job, and for whom the job was run. Through the pooling of departmental resources into larger enterprise grids, multiple users, teams, and departments can share common resources while working on different projects with different goals and schedules, providing maximum resource availability to all users in a flexible, policy-based environment. Productivity can be dramatically increased compared to pre-grid approaches.

The HPC consists of 512 cores parallel cluster machine with 128 cores of Intel Xeon quad-core E5440 processors @2.83GHz and 384 cores of Intel Xeon hex-core E5530 @2.93GHz processors. Each core has a minimum of 2 GB RAM that provides the entire machine a ram capacity of over 1000 GBs, the total compute nodes are 46 (Fig 3), with two master nodes (Master 1 and Master 2). The processors and nodes are connected through high-performance infiniband (6 voltair 24 port infiniband) switches and a Panasas parallel file system (PanFS) having a total storage of 32 TB. Apart from a backup storage of 24TB, the whole machine provides a storage capacity of over 50 TB. The total system peak performance is around 5 teraFLOP/s.

2.2 Software:

2.2.1 Operating system:

CentOS (Community enterprise Operating System) is an Enterprise-class Linux Distribution derived from sources freely provided to the public by a prominent North American Enterprise Linux vendor. CentOS is developed by a small team of developers and is supported by an active user community and supported by linux administrators, network administrators, enterprisers, scientific community etc.

CentOS5.5 version is a stable version of the CentOS release and has been installed in 46 nodes and two master nodes of the HPC. The HPC is 64-bit architecture hence the operating system is 64-bit OS with libraries and developing libraries for 64-bit has been installed.

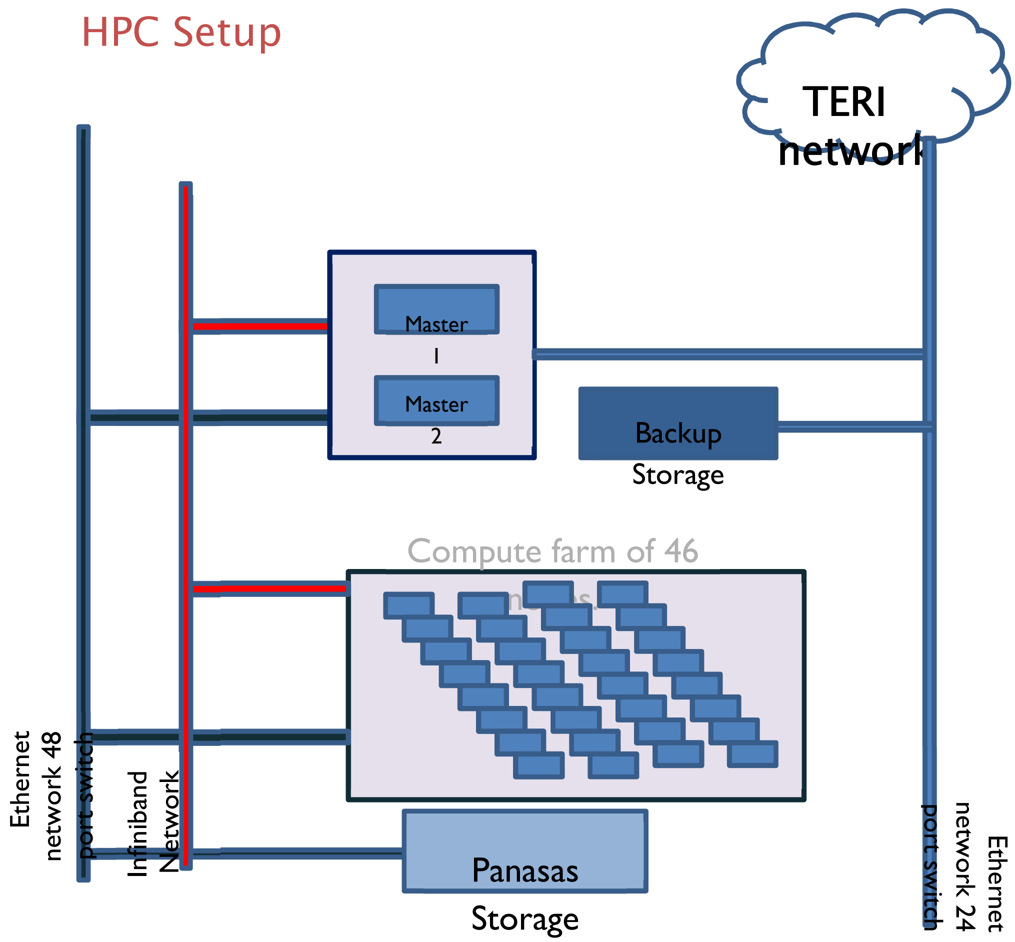

Figure 3: HPC Setup and architecture layout at TERI (Vinoth, WIPRO 2011).

2.2.2 Software packages installed

In order to carry forward the model installations, parallel processing software (open Message Passing Interface) system has been installed. The software we have used for implementing MPI is open-mpi and MPICH. The intel compilers for C, C++ and Fortran has also been installed in the HPC. The paths for MPI with intel compilers and intel compiler path folders are:

#source /opt/intel/Compiler/11.0/081/bin/ifortvars.sh intel64

#source /opt/intel/Compiler/11.0/081/bin/iccvars.sh intel64

MPI installed

Openmpi 1.4.3	-/opt/openmpi-1.4.3-intel and /opt/openmpi-1.4.3-gcc
Openmpi 1.4.4	-/opt/openmpi-1.4.4-intel and /opt/openmpi-1.4.4-gcc
Intel compiler 11	-/opt/intel/impi/3.2.1.009/bin64

A Module based environment is installed in the system to provide the dynamic modification of users environment via module files. This has been implement with the help of Alexander Oltu, BCCS, and Norway who was involved in the TERI-BCCR project from BCCR. Modules can be loaded and unloaded dynamically and automatically in a clean manner. Scientific softwares like CDO (Climate Data Operators), netCDF (network Common Data Format), HDF (Hierarchical Data Format), FERRET scientific visualization software, ncview (netcdf viewer software) with intel/gcc compilers are installed and can be loaded on the HPC with the use of simple module command.

module load "software_of_interest"

2.2.3 Monitoring Software

Ganglia is a scalable distributed monitoring system for high-performance computing systems such as clusters and Grids. It is based on a hierarchical design targeted at federations of clusters. It relies on a multicast-based listen/announce protocol to monitor state within clusters and uses a tree of point-to-point connections amongst representative cluster nodes to federate clusters and aggregate their state. It leverages widely used technologies such as XML for data representation, XDR for compact, portable data transport, and RRDtool for data storage and visualization. It uses carefully engineered data structures and algorithms to achieve very low per-node overheads and high concurrency. The implementation is robust, has been ported to an extensive set of operating systems and processor architectures, and is currently in use on over 500 clusters around the world. It has been used to link clusters across university campuses and around the world and can scale to handle clusters with 2000 nodes. The ganglia system is comprised of two unique daemons, a PHP-based web frontend and a few other small utility programs. This ganglia system is installed in TERI-HPC for monitoring the nodes and is maintained both by TERI and WIPRO team.

2.2.4 Torque scheduler

TORQUE Resource Manager provides control over batch jobs and distributed computing resources. It is an advanced open-source product based on the original PBS project* and incorporates the best of both community and professional development. It incorporates significant advances in the areas of scalability, reliability, and functionality and is currently in use at tens of thousands of leading government, academic, and commercial sites throughout the world. Torque software scheduler is installed in the TERI HPC cluster to take forward the activities related to jobs distribution in the HPC.

Torque benefits

- Initiate and manage batch jobs, create route execute modify or delete
- Define and implement resource policies that determine how much of each resource can be used by a job.
- Apply job to resources across multiple servers to accelerate job completion.
- Collects information about the nodes within the cluster to determine which are in use and which are available.
- Major tasks include
    - Job submission
    - Job monitoring
    - Job deletion

2.3 Technical support and service from WIPRO

TERI has signed a contract with WIPRO for continuous support and maintenance of HPC. Everyday maintenance will be taken forward by climate modeling researchers and IT department at TERI. However, the technical support for hardware issues, and software issues is also provided by BCCR, Alexander Oltu from BCCS and WIPRO on time-to-time basis.

3. Model installation and configuration

3.1 NorESM installation – how and where

The installation of NorESM at TERI HPC comprises the model environment, the model code and the model input data. The model environment consist of a collection of scripts and configuration files that are necessary for setting up new experiments and for running the model. The model environment and model code are installed under ~vidya/NorESM and use approximately 1 Gb of disk space. Two model versions are available in this location: noresm1-m, which is the CMIP5 version without carbon cycle (as described in Bentsen et al. 2012); and the noresm1-me, which is the CMIP5 version with carbon cycle (as described in Tjiputra et al. 2012). Since the NorESM is to a large extent based on NCAR's model system, the NorESM installation also includes the complete CESM (formerly CCSM) system and it uses the same model environment.

All model components that are specific to NorESM (i.e., MICOM, CAM-OSLO and HAMOCC) can be activated optionally. This is done during the creation of a new experiment (termed "case") with the help of the create_newcase script. This script creates a case catalogue outside the model system (under ~vidya/NorESM/cases) that contains all experiment specific configurations and model modifications as well as the run script for the experiment.

The model is executed in a case specific run-catalogue which is located under */work/vidya/cesm*. The model output is initially stored in this catalogue. After the model execution is finished, the output is moved to an archive area which is found in */work/vidya/archive*.

The model input data is stored in */opt/noresm/inputdata* and uses approximately 570 Gb of disk space. The data comprises initial and boundary conditions, grid information, historical forcing products and other data that is required for running the model in various configurations. For a single experiment, however, only a small subset of this data is necessary. Observational data that can be used for model validation is stored in */opt/noresm/obs*.

Documentation on the model environment and how to set up and run new experiments in the CESM framework is available at *http://www.cesm.ucar.edu/models/cesm1.0*. Any modifications of the model environment are documented in the files *ccsm2NorESM.log* and *cesm2NorESM.log* which are found in *~vidya/NorESM/noresm-1me*. Further NorESM specific information is collected in the folder *~vidya/NorESM/noresm-1me/scripts/README-NorESM*.

3.2 Customization of configuration files

Several configuration files stored in the folder *noresm1-me/scripts/ccsm_utils/Machines* had to be changed and customized for the installation on the TERI cluster:

kbatch.tesim
purpose: instructions for generation of run-script
changes: customized mpi-run command for TERI
comment: mkbatch.tesimNOOPT version for debugging with optimization turned off

env_machopts.tesim
purpose: sets path information to libraries, loads modules, etc
changes: customized path information for TERI

Macros.tesim
purpose: sets compiler options
changes: customized compiler options for TERI

config_machines.xml
purpose: contains machine specific information
changes: added a section "tesim" for TERI cluster

config_pes.xml
purpose: contains cpu options for different grids and different machines
changes: added options for TERI (machine identifier="tesim")

3.3 Changes in model code

Some minor changes had to be applied to the model code because of the use of a different compiler (Intel instead of Portland-Group) and because some source files contained hard-coded path information to input data. These modifications have no effect on the actual computations.

Following source files of the component CAM-OSLO had to be changed (stored in *noresm1-me/models/atm/cam/src/physics/cam_oslo)*:

aerosoldef.F90	*genaero_intr.F90*	*initlogn.F90*	*opttab.F90*
aqrat.F90	*initaeropt.F90*	*intccn1to3.F90*	*oxidants.F90*
emissions.F90	*initccn.F90*	*intccn4.F90*	*physpkg.F90*
gaschem.F90	*initdryp.F90*	*intccn5to10.F90*	*wetdepaer.F90*

In the component MICOM, the files c*ommon_bud.h a*nd *mod_xc.F* had to be modified. These files are found in *noresm1-me/models/ocn/micom/phy*.

In the subcomponent HAMOCC, the file *carchm.F90* had to be modified. This file is found in *noresm1-me/models/ocn/micom/hamocc*.

The original files are still present, but with the string ".ORIG" added at the end of their names.

3.4 System maintenance

The changes in model code and configuration files will be transferred to the official NorESM repository. Thus, in the future, it will be possible to upgrade the NorESM installation at TERI by downloading the latest NorESM version from the subversion server. This can be done by issuing the command "svn checkout https://svn.met.no/NorESM/noresm/trunk .". Access to the repository can be requested by emailing to helpdesk@met.no. Since the TERI cluster currently does not support direct internet access, the model has to be first downloaded to a different server and then copied to the cluster.

Input data, that is required for specific experiments or specific grid resolutions, is normally downloaded automatically from NCAR's subversion server. However, since the TERI cluster does not support direct internet access, any additional input data needs to be downloaded manually from NCAR's subversion server (https://svn-ccsm-inputdata.cgd.ucar.edu/trunk/inputdata/) and then moved to the input data location on the TERI cluster, which is */work/shared/cesm/inputdata.* The access to NCAR's repository requires registration at http://www.cesm.ucar.edu/models/cesm1.0. Any input data that is particular to the NorESM system can be requested from noresm-ncc@met.no.

4. Testing and performance

4.1 Verification of model installation

We use the following checklist to evaluate if the porting of the model system to TERI was successful:

Does the model compile and run without errors?

The model compiles on the TERI cluster and no crashes have been registered during more than ten test simulations that have been performed to benchmark the system. Furthermore, the system has been compiled and tested with debug options that enforce checking of array bounds and floating point operations. No error has been triggered so far.

Can the model reproduce results that are generated on the same machine?

In case that output gets lost, it is generally desirable to regenerate this output, especially if it builds the basis of scientific publications. By making use of restart (or checkpoint) information one tries to avoid having to rerun entire experiments, in particular if only a small part of the output is affected.

We have performed following test, which the TERI setup passed successfully:

Step 1: a simulation A is run for 10 model days, starting from certain initial conditions.
Step 2: a simulation B is run for 5 model days, starting from the same initial conditions as A and also otherwise identically with A. At the end of the 5th day, restart data is written.

Step 3: a simulation C is initialized with restart information that was written by A and then run for another 5 model days.

Step 4: the final result of A and C are compared, e.g., by computing checksums of the output files.

Can the model reproduce results that are generated on other machines (i.e., outside of TERI)?

In the majority of cases, it is not possible to produce bit-identical simulation results on different computational systems. In this respect, the porting of NorESM to the TERI cluster is no exception. Due to the chaotic nature of weather, microscopic deviations in the model state quickly lead to large differences in the model weather. These differences typically stem from rounding and truncation errors, which depend on processor architecture, compiler optimizations and model parallelization.

More important than bit-reproducibility, however, is the question whether scientific publication results can be reproduced. For the reproducibility of scientific results it is in most cases not necessary to reproduce the exact state of the “model weather”. Hence, one can relax the reproducibility requirement, stating that is sufficient to be able to reproduce the same climate evolution and climate response (i.e., the statistics of the model weather, but not the model weather itself).

To verify the reproducibility of the model climate, we have analyzed the top-of-atmosphere radiative balance, which is a good measure for the stability of the climate (i.e., values close to zero means stable while values larger than ~1 W/m2 produce significant climate drift) and very sensitive to minimal model changes. Fig. 4a shows the comparison of radiative balance from a NorESM1-ME historical test simulation performed on the TERI cluster and from a corresponding simulation performed on the Norwegian system HEXAGON, for the years 1850-1879. Both radiation curves are close to zero, indicating a stable climate. Apparent oscillations around zero are primarily a consequence of internal ENSO variability and hence not expected to be in phase on the two machines. The mean radiative balance is 0.06 +- 0.07 W/m2 on the TERI cluster and indistinguishable from the one on HEXAGON,. Consistent with the top of atmosphere radiative balance, the evolutions of the global surface air temperatures are broadly similar (Fig. 4b) and residual differences can be attributed to the presence of chaotic, internal climate variability. Preliminary simulations with NorESM-1ME restart run from Dr. Jerry with start year of simulation being 931 year has been performed and the results are similar to the one obtained from Norwegian HEXAGON system. Based on these three criteria, we can conclude that the porting of the NorESM system to the TERI cluster has been successful.

4.2 Performance optimization

The performance optimization currently applied to the NorESM configuration on the TERI cluster is rather conservative. The compiler optimization level is set to the moderate O2 value, and the option "-fp-model precise" is activated, which disables any optimizations that affect rounding and truncation results. More aggressive optimization could potentially lead to a performance gain of 10-20 %. However, this would require thorough testing to guarantee that stability, reproducibility and general reliability of the system are retained.

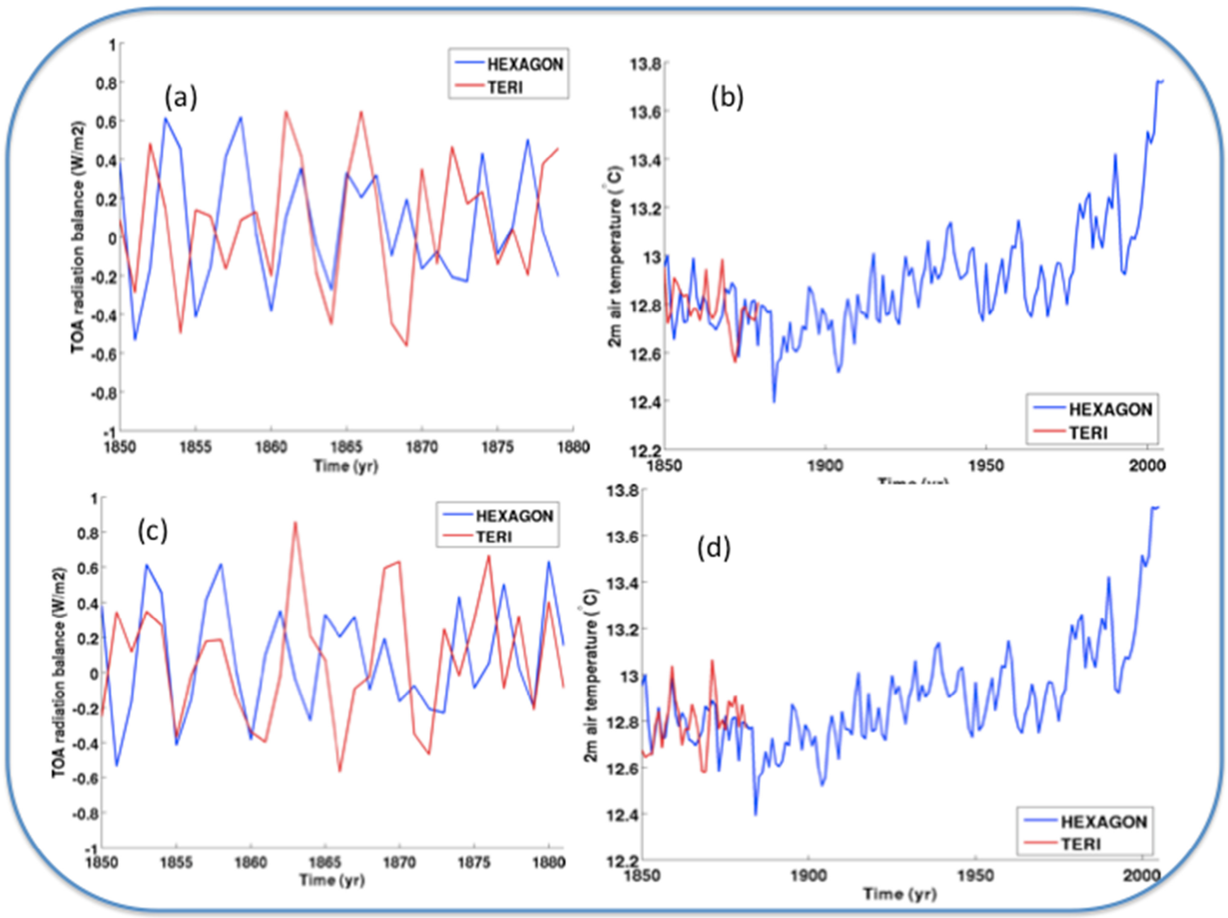

Figure 4: (a) Annual mean top-of-atmosphere radiation balance for the first 30 years of the NorESM1-ME historical experiment, comparison between TERI-HPC and Hexagon in Norway, (b) 2m Surface Temperature comparison between TERI and Hexagon in Norway, (c) Annual mean top-of-atmosphere radiation balance for the NorESM-1ME restart run starting from year 931, and (d) Similar to (b) for restart run starting from 931 year.

Various run-time options are set to optimize the mpi-communication between the processors. This is done with the help of Modular Component Architecture (MCA) parameters that are passed to the *mpirun* command which launches the model on the compute nodes.

The MCA parameters specified for NorESM are:

Name of MCA parameter	*Value*
btl	openib,self,sm
btl_openib_if_include	mthca0
btl_sm_use_knem	1
mpi_paffinity_alone	1
btl_tcp_if_include	etho0
oob_tcp_if_include	eth0

Further performance optimization can be achieved by optimizing the load balancing between the different model components. If some components are slower than other components then the faster components spend time waiting for the slower components before they can exchange data. Therefore, one always tries to balance the number of cpus between the components such that the waiting time is minimized. The optimal cpu layout depends on the computational system. In particular, the number of cpus per node is critically because the communication between cpus of the same node is much faster than the inter-node communication. On the TERI cluster, some preliminary load balancing has been performed for the NorESM1-ME. The best result (6.7 model year per integration day, with a cost of 600 cpu hours per model year) has been obtained for a processor count of 168 cpus, whereof 72 cpus are used for the ocean, 96 cpus for the atmosphere, and 80+12 cpus for the sea ice+land (which run in serial with the atmosphere and therefore use the same cpus as the atmosphere).

4.3 Performance benchmarking

A series of test runs with different model and cpu configurations have been carried out to evaluate and benchmark the model performance on the TERI cluster. The timing results have been compared with those obtained on the norwegian supercomputer HEXAGON, which has a CRAY XE6 architecture with 32 cpus on each computational node (for more information see http://www.bccs.uni.no/hpcdoc/Available_resources).

It should be emphasized that performance scaling is a complex function due to the different speeds of intra versus inter-node communication and due to changes in the load balancing between components when the number of cpus are changed. The benchmarking results are summarized in the table below.

The performance of the model system on the TERI cluster ranges between 6 and 13 model years per integration day (y/d), depending on the model configuration (resolution and active components) and the processor count and layout. For the intermediate resolution setups (NorESM1-M and NorESM1-ME), the TERI cluster outperforms the CRAY for low processor counts while the CRAY outperforms the TERI for higher processor counts. For the low resolution setup (NorESM1-L), the results indicate that the parallelization gain on the TERI cluster has reached saturation at 64 cpus, which is the lowest processor count that was tested.

Table 1: Performance benchmarking results between Norway-hexagon and TERI-HPC.

	cpus	speed-TERI (yr/d)	speed-NORWAY (yr/d)	cost-TERI (cpu-hrs/sim-d)	cost-NORWAY (cpu-hrs/sim-d)
NorESM1-L	64	**13.2**	19.5	116	79
	128	8.2	33.3	373	92
NorESM1-M	160	**6.9**	5.0	553	770
	310	7.4	8.5	1002	877
NorESM1-ME	160	6.2	3.9	618	976
	168	**6.7**	4.4	602	922
	310	7.3	7.0	1013	1057
	396	-	8.3	-	1136
CESM1 - 2deg	176	-	11.7	-	362
	214	**10.6**	-	483	-
	416	-	18.7	-	534

Consistent with the hardware specifications, these results indicate that the serial cpu performance of the TERI cluster is better than that of the CRAY, while the communication between the processors is slower on the TERI cluster. The latter limits the scalability of the model on the cluster. Hence, from a cost-benefit point of view, one should use fewer processors on the TERI cluster than on the CRAY. We recommend using the lowest processor count (second lowest for NorESM1-ME) that is specified for each respective model configuration in Table 1. This has the additional advantage that more than one experiment can be run at the same time on the cluster. The performance is then approximately 13 yr/d for the low-resolution NorESM, 7 yr/d for the intermediate-resolution NorESM (with and without carbon cycle), and 11 yr/d for the standard CESM1 at intermediate resolution. The reason for a higher performance of CESM over NorESM is that NorESM contains additional aerosol-cloud chemistry components that are not present in CESM.

5. Ongoing work and future plans

In the Norwegian Framework of Agreement supported by the Royal Norwegian Embassy, TERI-BCCR collaborative project aims to a) strengthen the climate modeling skills, b) enhance the understanding of the earth system processes (teleconnection linkages, systematic biases etc), c) provide high resolution regional climate change scenarios which in turned be tailored to impact assessments and their links to policy. These multiple objectives are quite difficult to achieve within the given frame considering the varied nature and the human capacity required to achieve the goals.

We have successfully completed the porting of CESM, NorESM and WRF models in TERI HPC and also have completed a few test runs to check the system performance. The benchmarking simulation for NorESM has been an important task and we were able to complete this with good performance (6-7 model years per integration day) for intermediate resolution. We are planning to continue the 150-year historical simulation of NorESM1-ME with restart conditions from 931 year of spin-up run and also then perform an RCP simulation with RCP 8.5 scenario. The approximate time to complete these simulations would be five weeks. The restart conditions considered as 931 year from Dr. Jerry's simulation, since the historical simulation performed at BCCR has been from 900 year and by performing the 931-year based simulation, we would create an ensemble member for the carbon cycle based runs, which could be also shared with the BCCR researchers. It is important to note that individual 150 year simulation of an ESM run would require both high computational time and space to store the model outcomes.

In the current project time period, we have strengthened the relationship between scientists from BCCR and TERI to take forward such large range of activities. We are able to achieve the mentioned targets for the first objective i.e., installation and validation of NorESM and strengthening the climate modeling skills (both human and infrastructure), a on-going working report on the second objective will also be submitted along with this report for the work on understanding teleconnections in NorESM-1M model. The third objective is related to the WRF-Regional Climate run which is a successive component of the NorESM1-M simulations and the progress in this work will be discussed on another working paper. Overall, the targets mentioned to achieve in the current project is varied and a start up of all the initiatives have been carried in this exercise, and in a 3-year project it is challenging to meet the detailed objectives related to improved understanding of earth system processes. Thus, continuous collaborative efforts from both ends are required to be able to achieve this target. As part of this project, we have also developed a new stream of capacity building the researchers in South Asia on short-term five- day courses to understand the climate modeling processes better. The approach followed in these five day courses are to ensure the participants are able to get the leading edge knowledge on climate change processes and modeling tools. The Secretary, Ministry of Environment, India and also Norwegian Embassy have also recognized this initiative and expressed the importance of such exercises in developing country perspective.

References

Assmann, K. M., Bentsen, M., Segschneider, J., and Heinze, C.: An isopycnic ocean carbon cycle model, Geosci. Model Dev., 3, 143–167, doi:10.5194/gmd-3-143-2010, 2010.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjansson, J. E.: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation, Geosci. Model Dev. Discuss., 5, 2843–2931, doi:10.5194/gmdd-5-2843-2012, 2012.

Bleck, R. and Smith, L. T.: A wind-driven isopycnic coordinate model of the North and Equatorial Atlantic Ocean, 1. Model development and supporting experiments, J. Geophys. Res., 95, 3273–3285, 1990.

Bleck, R., Rooth, C., Hu, D., and Smith, L. T.: Salinity-driven thermocline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic, J. Phys. Oceanogr., 22, 1486–1505, 1992.

Briegleb, B. P. and Light, B.: A Delta-Eddington Mutiple Scattering Parameterization for Solar Radiation in the Sea Ice Component of the Community Climate System Model, Tech. Rep. NCAR/TN-472+STR, National Center for Atmospheric Research, Boulder, Colorado, USA, 2007.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z. L., and Zhang, M.: The community climate system model version 1, J. Climate, 4, 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

Holland, M. M., Bailey, D. A., Briegleb, B. P., Light, B., and Hunke, E.: Improved sea ice shortwave radiation physics in CCSM4: the impact of melt ponds and aerosols on arctic sea ice, J. Climate, 25, 1413-1430, doi:10.1175/JCLI-D-11-00078.1, 2012.

Hunke, E. C. and Lipscomb, W. H.: CICE: the Los Alamos Sea Ice Model, documentation and software, version 4.0, Tech. Rep. LA-CC-06-012, Los Alamos National Laboratory, Los Alamos, New Mexico, USA, 2008.

Iversen, T., Bentsen, M., Bethke, I., Debernard, J. B., Kirkevåg, A., Seland, Ø., Drange, H., Kristjánsson, J. E., Medhaug, I., Sand, M. and Seierstad, I. A.: The Norwegian Earth System Model, NorESM1-M – Part 2: Climate response and scenario projections, Geosci. Model Dev. Discuss., 5, 2933–2998, doi:10.5194/gmdd-5-2933-2012, 2012.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E., Struthers, H., Ekman, A. M. L., Ghan, S., Griesfeller, J., Nilsson, E. D., and Schulz, M.: Aerosol-climate interactions in the Norwegian Earth System Model - NorESM, Geosci. Model Dev. Discuss., 5, 2599–2685, doi:10.5194/gmdd-5-2599-2012, 2012.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in version 4 of the community land model, J. Adv. Model. Earth Syst., 3, M03001, doi:10.1029/2011MS000045, 2011.

Maier-Reimer, E.: Geochemical cycles in an ocean general circulation model. Preindustrial tracer distributions, Glob. Biogeochem. Cycles, 7, 645–677, doi:10.1029/93GB01355, 1993.

Maier-Reimer, E., Kriest, I., Segschneider, J., and Wetzel, P.: The HAMburg Ocean Carbon Cycle Model HAMOCC5.1 – Technical Description Release 1.1, Berichte zur Erdsystemforschung 14, ISSN 1614– 1199, Max Planck Institute for Meteorology, Hamburg, Germany, 50 pp., 2005.

Neale, R. B., Richter, J. H., Conley, A. J., Park, S., Lauritzen, P. H., Gettelman, A., Williamson, D. L., Rasch, P. J., Vavrus, S. J., Taylor, M. A., Collins, W. D., Zhang, M., and Lin, S.-J.: Description of the NCAR Community Atmosphere Model (CAM 4.0), Tech. Rep. NCAR/TN-485+STR, National Center for Atmospheric Research, Boulder, Colorado, USA, 2010.

Neale, R. B., Richter, J., Park, S., Lauritzen, P. H., Vavrus, S. J., Rasch, P. J., and Zhang, M.: The mean climate of the Community Atmosphere Model (CAM4) in forced SST and fully coupled experiments, J. Climate, submitted, 2012.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stöckli, R., Wang, A., Yang, Z.-L., Zeng, X., and Zeng, X.: Technical Description of version 4.0 of the Community Land Model (CLM), Tech. Rep. NCAR/TN-478+STR, National Center for Atmospheric Research, Boulder, Colorado, USA, 2010.

Seland, Ø., Iversen, T., Kirkevåg, A., and Storelvmo, T.: Aerosol-climate interactions in the CAM-Oslo atmospheric GCM and investigation of associated basic shortcomings, Tellus A, 60, 459–491, doi:10.1111/j.1600-0870.2008.00318.x, 2008.

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D., Lorentzen, T., Schwinger, J., Seland, Ø., and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), Geosci. Model Dev. Discuss., 3035–3087, doi:10.5194/gmdd-5-3035-2012, 2012.

Documentation resources

A collection of scientific documentation on the NorESM is hosted by Geoscientific Model Development (GMD) and made available at:
http://www.geosci-model-dev-discuss.net/special_issue21.html
http://www.geosci-model-dev.net/special_issue20.html

Scientific documentation on the CESM is hosted by the American Meteorological Society (AMS) and made available at:
http://journals.ametsoc.org/page/CCSM4/CESM1

CESM user information is made available by NCAR at:
http://www.cesm.ucar.edu/models/cesm1.0

Information concerning the Coupled Model Intercomparison Project version 5 (CMIP5) is made available by PCMDI at:
http://cmip-pcmdi.llnl.gov/cmip5