Present and Future Computing Requirements for IMPACTS and CLIMES

(Investigation of Magnitudes and Probabilities of Abrupt Climate Transitions)
(Center at LBNL for Integrative Modeling of the Earth System)

William D. Collins

Berkeley Lab and UC Berkeley

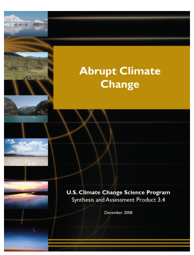
NERSC BER Requirements for 2017
September 11-12, 2012
Rockville, MD

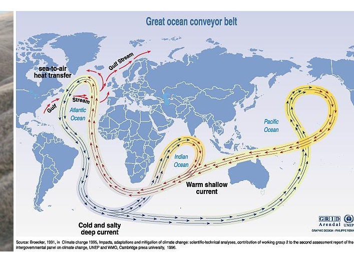
1. Project Description

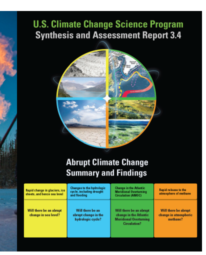
IMPACTS PI: William Collins

IMPACTS leads: William Riley (boreal), Philip Cameron-Smith (clathrates), William Lipscomb/Steve Price (land ice), and Ruby Leung (droughts)

- Scientific objectives of IMPACTS
  - **Project the risk of abrupt climate change over the 21st Century**
    - *Disintegration of marine ice sheets*
    - *Melting permafrost leading to releases of CO2 and CH4*
    - *Destabilization of methane deposits in Arctic-circle oceans*
    - *Large-scale megadroughts in North America*
  - **Enhance global models of these rapid climate transitions**



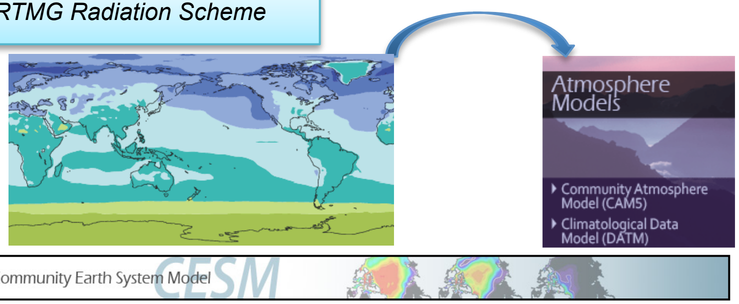
1. Project Description (cont.)

CLIMES PI: William Collins
CLIMES leads: David Romps (fast physics), William Collins (extreme/local change), William Riley and Jeffrey Chambers (terrestrial Earth systems)

- Scientific objectives of CLIMES
  - **Advance simulations of climate forcing, response, and feedback:**
    - *Ultra high-resolution global climate simulation*
    - *Frameworks for robust regional climate modeling*
    - *Quantification of critical uncertainties in the carbon cycle*
    - *Representation of clouds, aerosols, and the cryosphere in climate models*
  - **Advance projections of climate mitigation measures:**
    - *Improved representations of human-Earth system interactions*
    - *Integrated assessment model development, intercomparison, and diagnostics*

Contributions to CESM Science and Capabilities

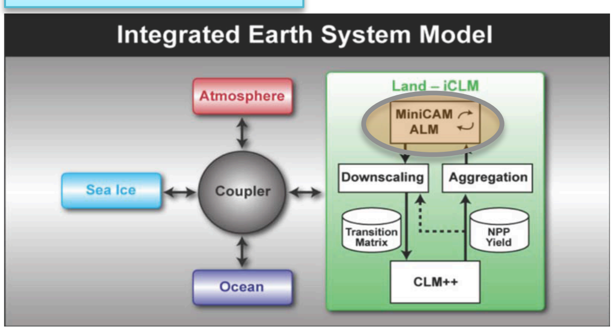
RRTMG Radiation Scheme



Radiation:

- LBNL led the integration of the ASR-funded RRTMG radiation parameterizations into the DOE-NSF CESM.
- RRTMG is included in the public release of CESM.

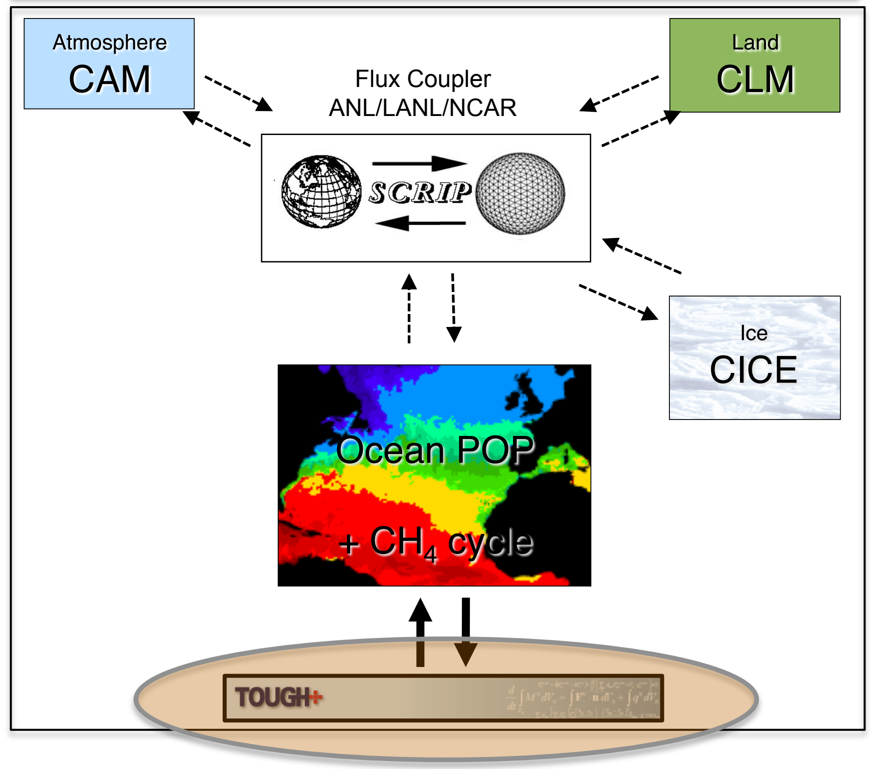
Integrated Assessment



iESM:

- LBNL led the integration of the BER-funded GCAM IA model into the DOE-NSF CESM.
- LBNL has shared this capability with iESM ST.

IMPACTS Model of the Global CH4 Cycle for Abrupt Climate Change

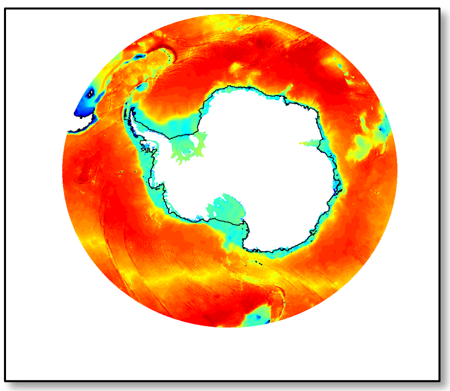


Abrupt Climate Change:

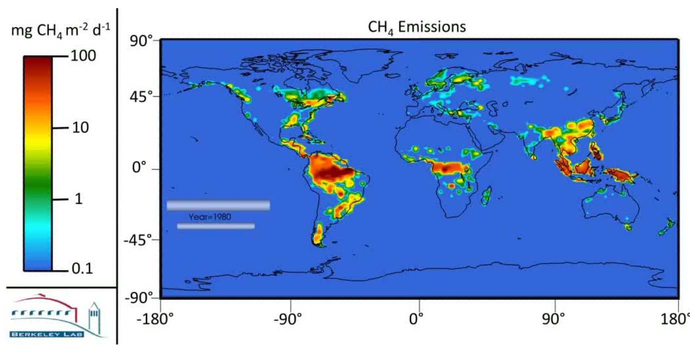
- LBNL led the integration of the 1D TOUGH+ model for ocean hydrates into the DOE-NSF CESM.
- LBNL heads the development of terrestrial CH4 cycle with treatments of permafrost, thermokarst, peat, etc.

1. Project Description (cont.)

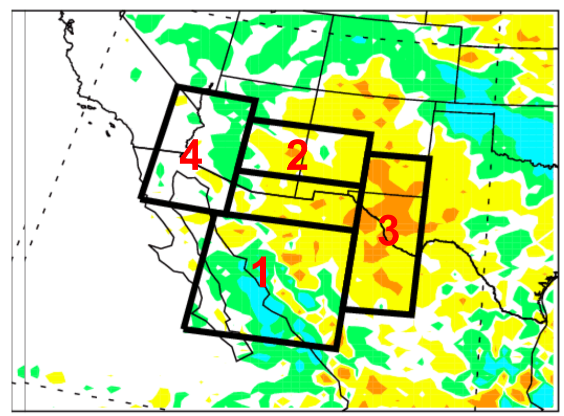
- **Our present focus in IMPACTs is to perform:**
  - 1st coupled projections of Earth's methane cycle
  - 1st sea-level rise projections including Antarctica
  - Simulations of the future of western forests



New POP bathymetry
White region: new land mask
Black line: border of original land mask



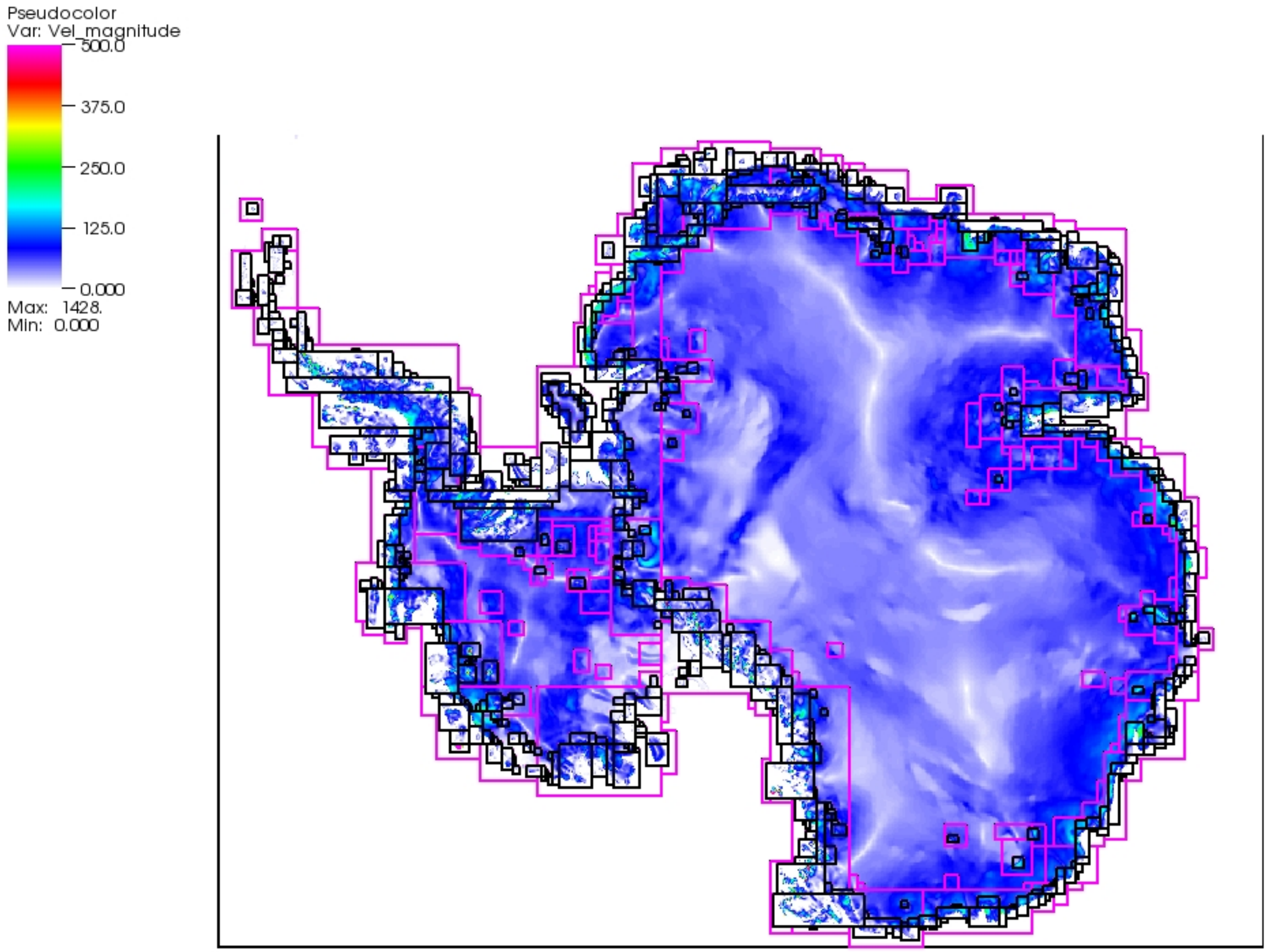
Summer precipitation change in the NAM region



Dust Impact on Rain (mm/day)

Dust Impact on Rain (mm/ day)

Dynamics of Antarctica and Sea-Level Rise



IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

Implementation of Land-Ice/Ocean Interface in POP

Objectives:

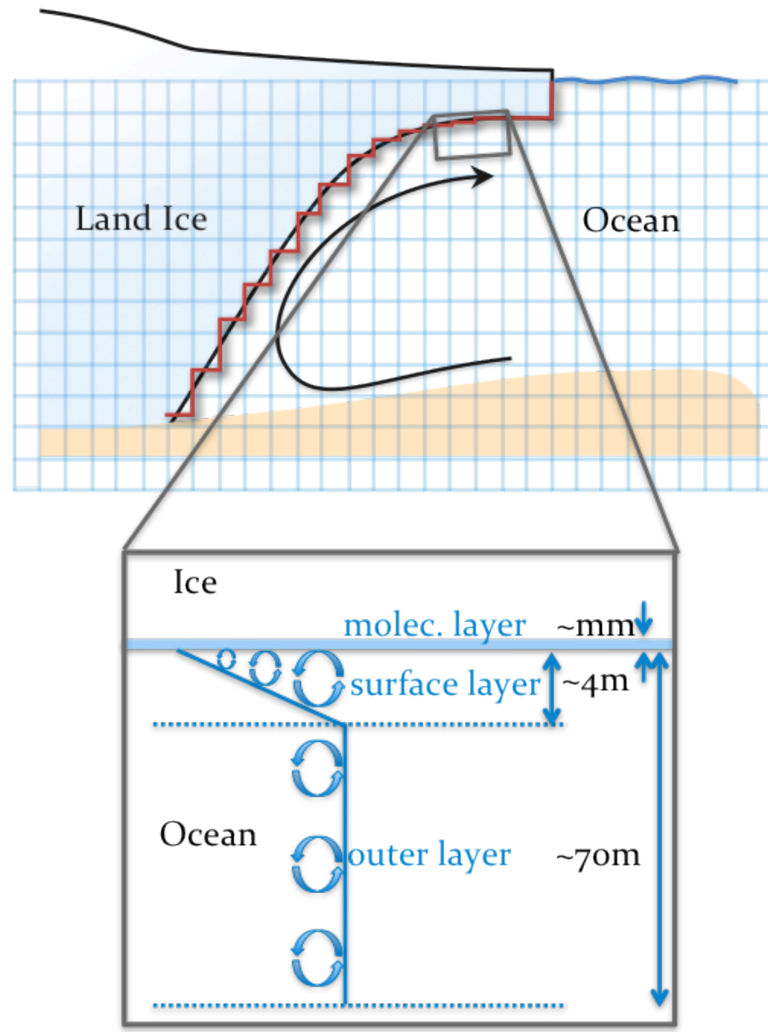
- Projections of sea-level rise due to Antarctic and Greenland ice sheets
- Examine abrupt climate feedbacks related to land-ice/ocean interactions

Implementation:

- Adding *dynamic* land-ice/ocean interface to POP
    - partial cells
    - immersed boundaries
- Developed efficient algorithms for representing turbulent ocean boundary layers under ice shelves

Experiments:

- Underway: expts. with fixed, idealized geometries for model comparison
- Next: expts. with fixed, realistic geometries for data comparison



top: land-ice/ocean interface (using partial cells)
bottom: turbulent boundary layer under the ice

IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

Atmospheric Impact of Methane Clathrate Emissions

Objective:

- Study the impact of methane clathrate emissions on the atmosphere.

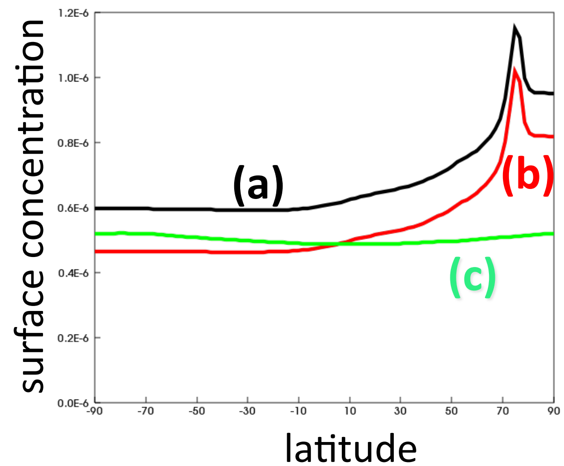
Implementation:

- Implemented Fast Methane Chemistry in CESM1_0_beta14, using CAM4 physics and RRTMG radiation.

Simulation Experiments :

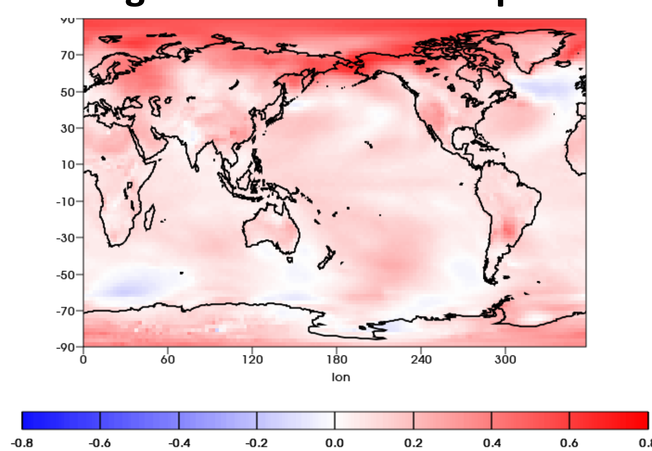
- **Control Case:** CESM, full ocean, fast methane chemistry, 2 degree resolution.
- **Arctic Methane Emission Case:** Simulating Arctic Methane Emission (22% increase),
- Impacts CH_4, T, rainfall, air-quality.
- **Uniform Methane Emission Case:** 22% increase in emissions spread uniformly,
- Impacts depend on emission location.

Methane Increase from Arctic Clathrates



(a) 100 year zonal mean
(b) 20 year zonal mean
(c) 20 year Uniform Emission Case

Change in Surface Air Temperature



100-year difference between Arctic methane emission & control cases

IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

Inclusion of a Terrestrial CH_4 Model into CESM1 (CLM4Me)

Objectives:

•Identify uncertainties

•Predict 21st century CH_4 emissions

•Quantify potential for abrupt feedbacks.

Implementation:

- Vertically resolved biochemical model
- 2 reactions and 3 transport processes
- Implementation designed to integrate with future land model improvements.

Experiments:

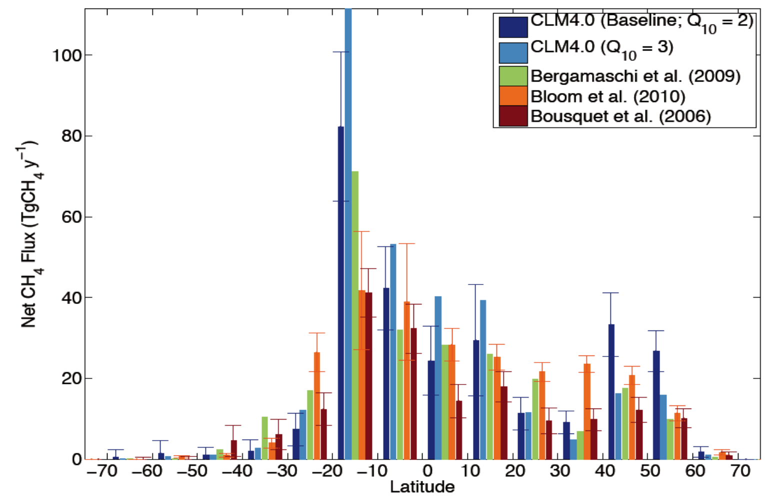
•Compared present CH_4 emissions to 15 sites and 3 atmospheric inversions.

•Identified critical uncertain parameters

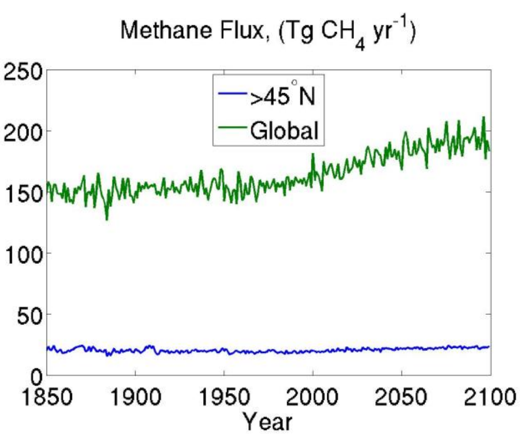
•Showed declines in high-latitude inundation may limit 21st century increases in emissions

•Developing subgrid peatland ecosystem model

Comparison of CH_4 emissions from CLM4Me and several atmospheric inversions.



RCP 4.5 emissions from vegetated ecosystems without old soil carbon source.



IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

Belowground Carbon Processes

Objectives:

- Represent processes responsible for growth and loss of permafrost C, which is a large (>1000 Pg) and vulnerable fraction of the terrestrial C pool.

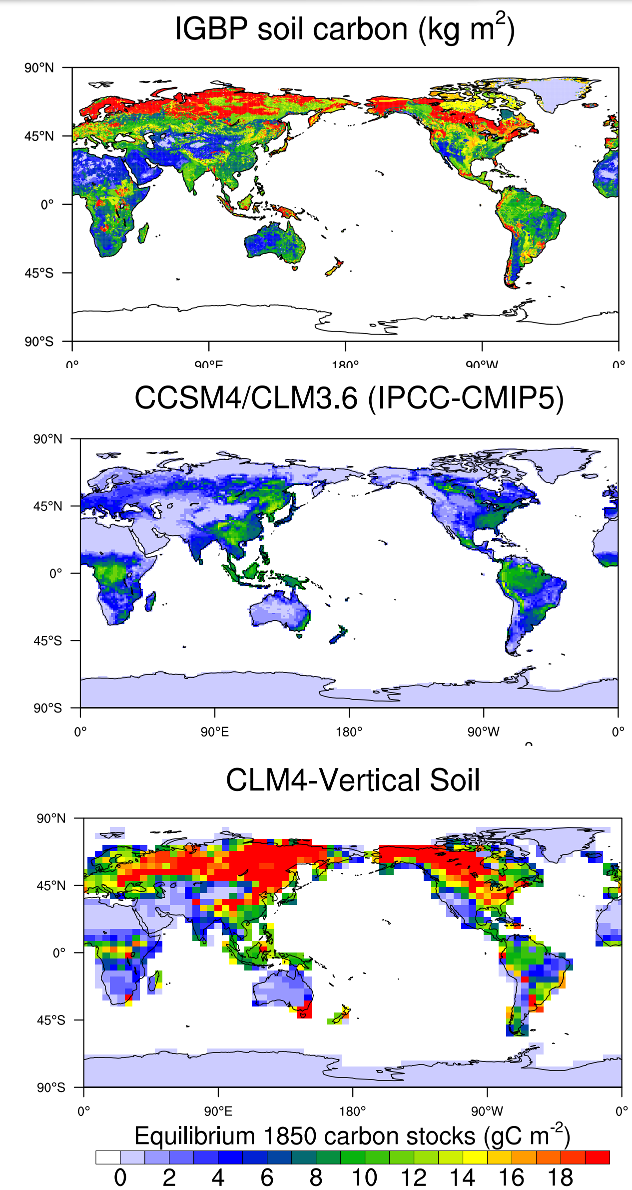
Implementation:

- Developed vertically-resolved belowground biogeochemistry, mixing.
- Improved SOM dynamics, growth of Permafrost C pools
- Improved N cycle at high latitudes leads to better productivity

Experiments and Next Steps:

•Equilibrium experiments, sensitivity to parameters and model structure

•Next Steps: Future scenarios; Coupling between soil and wetland biogeochemistry; coupled soil BGC and soil physics



IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

The Role of Surface Water – Groundwater Interactions on Long Term Droughts

Objectives:

- Study the role of surface water – groundwater interactions on long term droughts

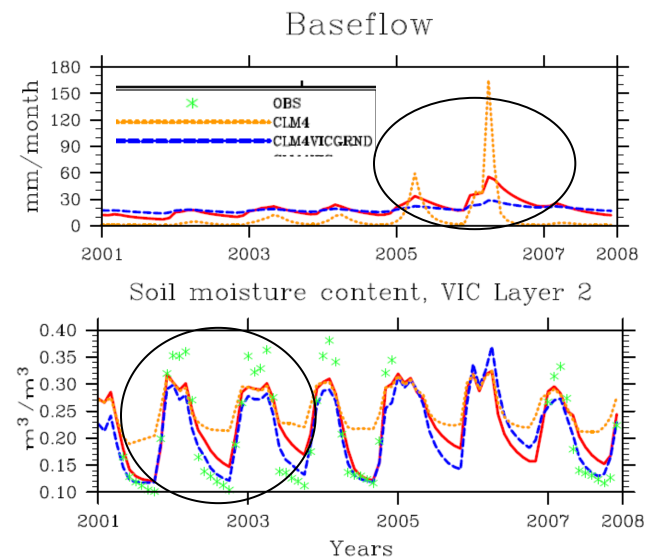
Implementation:

- Implemented the VIC runoff and groundwater parameterizations to CLM4
- CLM4 has been coupled to WRF using the CCSM flux coupler

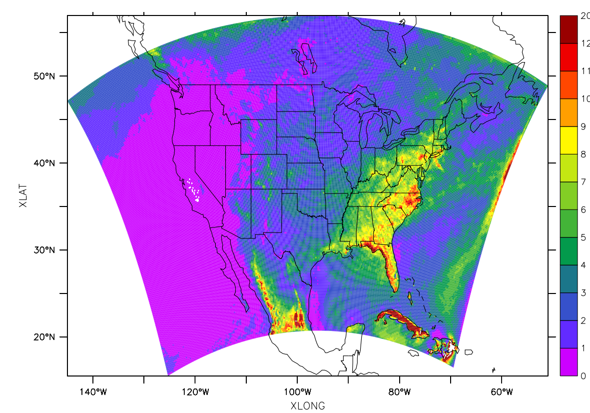
Experiments and Results:

- CLM4VIC-ground has been applied to flux tower sites for comparison with CLM4 and CLM4VIC and showed improvements in simulating seasonal soil moisture
- WRF-CLM has been configured for the US using a new global 0.05° CLM input data
- WRF-CLM simulates realistic precipitation and surface temperature in North America
- WRF-CLM will be used to perform numerical experiments to study the role of surface water – groundwater interactions on long term droughts

Comparison of CLM4, CLM4VIC, and CLM4VIC-ground at Tonzi, CA



2003 JJA precipitation simulated by WRF-CLM



IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

Impacts of Great Basin Dust on North American Summer Monsoon Precipitation

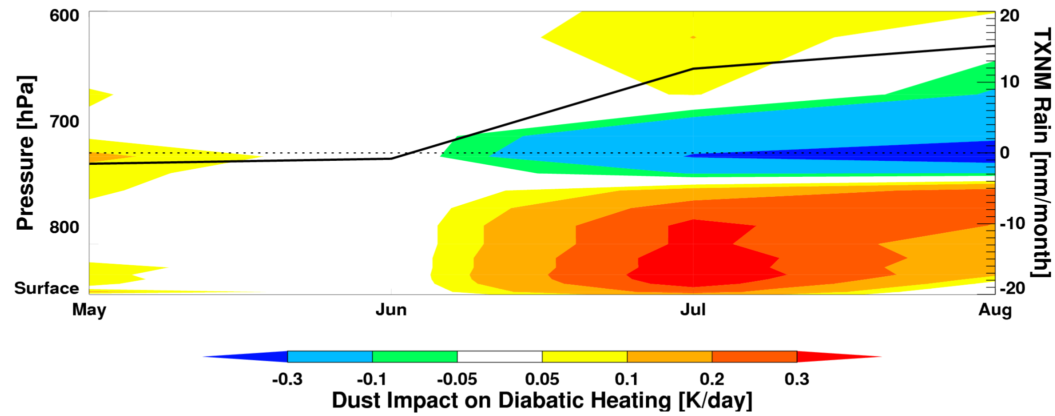
Objectives:

- Study the impacts of dust on North American summer monsoon precipitation

Implementation:

- WRF-Chem is used to conduct numerical experiments using the modal MADE/SORGAM scheme coupled with the GOCART dust emission scheme

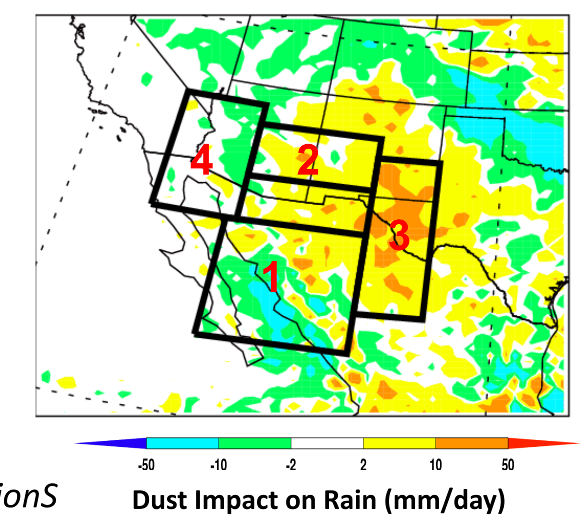
Dust induced diabatic heating over the desert (contours) and precipitation changes over Texas-New Mexico (Region 3 shown below)



Experiments and Results:

- Simulations were performed for April – September of 1995 – 2009 at 36km resolution with and w/o dust
- The simulated dust concentration and AOD compare well with observations
- Dust from the Great Basin induces surface cooling of 1 W/m^2 and atmospheric heating of 0.4 W/m^2
- Dust heating of 0.3 K/day in the lower atmosphere strengthens the low-level meridional winds, leading to a 10-40% increase in NAM precipitation

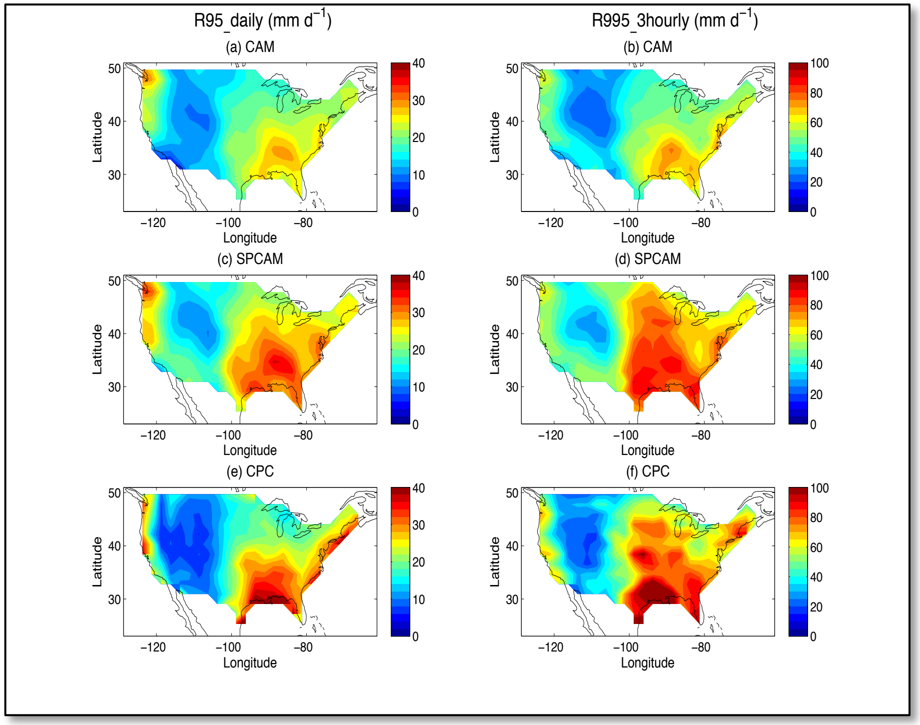
Summer precipitation change in the NAM region



IMPACTS: Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS

1. Project Description (cont.)

- Our present focus in CLIMES is to create:
  - New and more robust simulations of climate extremes
  - Enhanced models of regional moisture and precipitation
  - Better representations of carbon-cycle processes
  - First coupled energy-climate models with 2-way interactions



Reactive transport modeling-CLM4BeTR

Objectives:

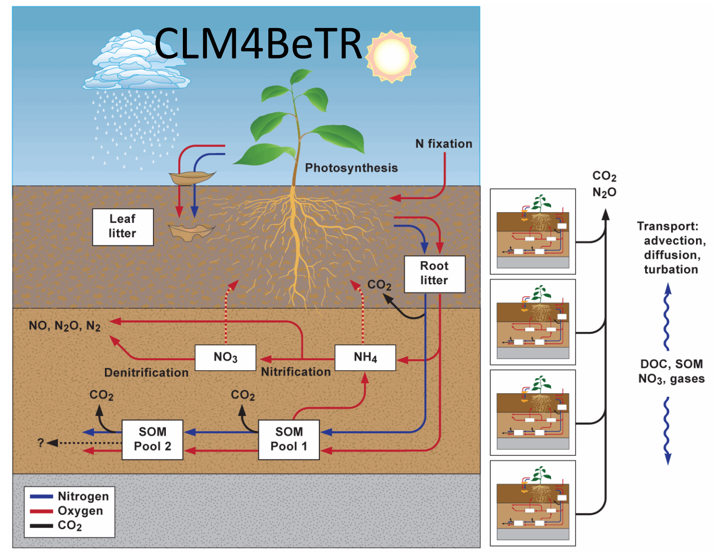
- Uniform implementation of vertically-resolved underground biogeochemistry
- Multiphase description of C-Nutrient dynamics, gaseous, aqueous, sorbet

Implementation:

•Operator splitting approach
•Two-layer bi-directional modeling of atmosphere-surface exchange for different tracers, e.g. NH_3, N_2O, CH_4
•Evaluation against analytical results (successful) and measurement data

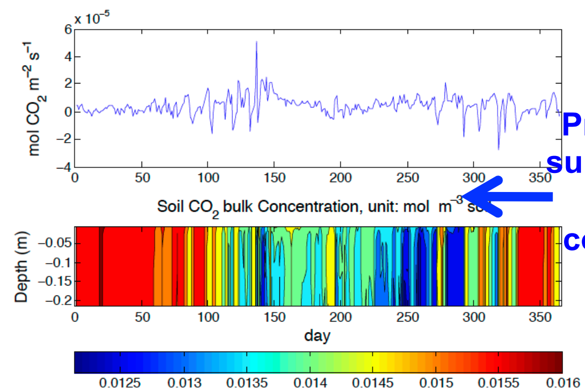
Experiments and :

•Tested single-point simulation of N_2O, CO_2 transport with vertically resolved C and N biogeochemistry
•Test the functionality of isotope fractionation and merge with CLM4Me



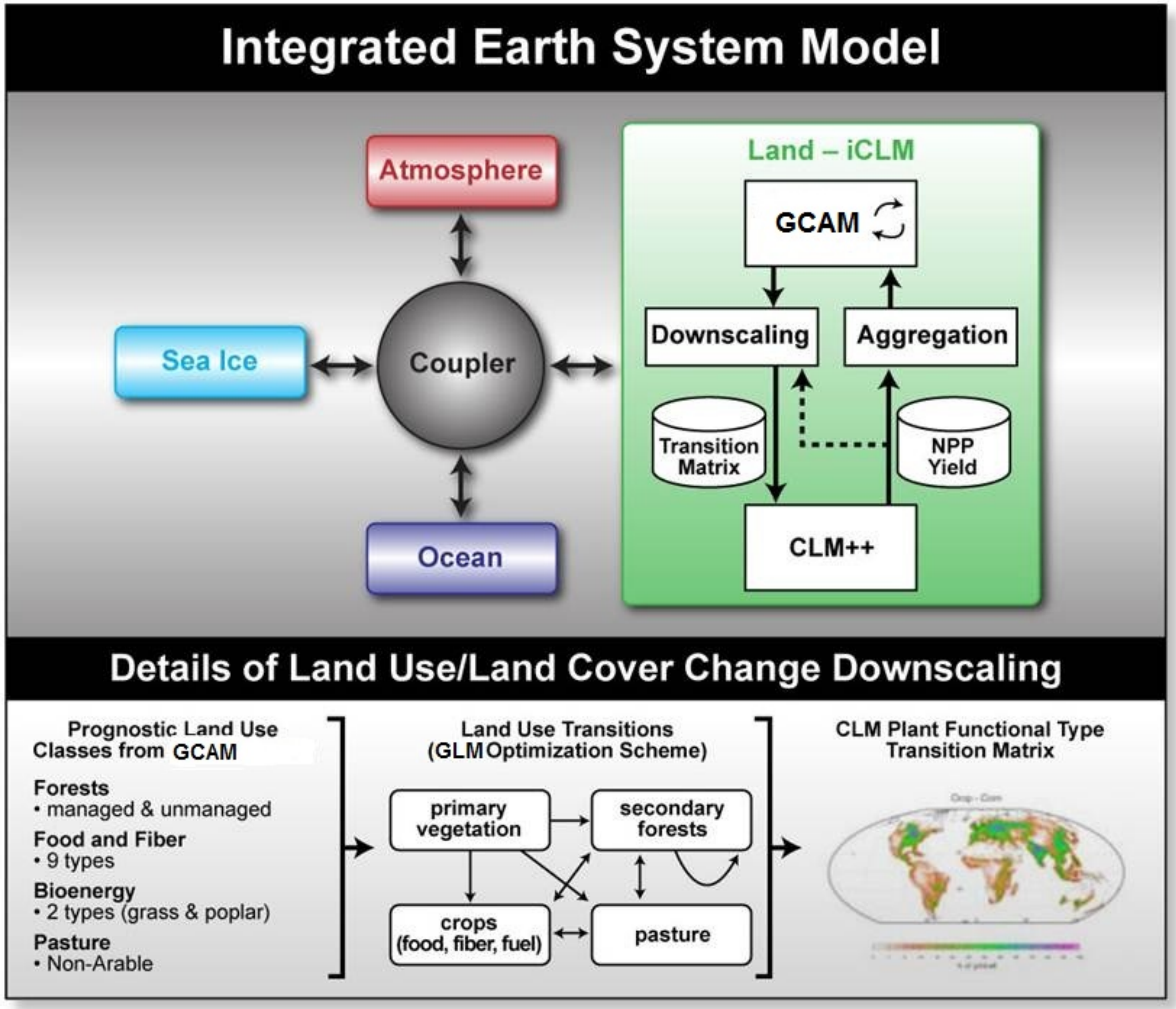
Vertical mixing & transport

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial t}\left(D\frac{\partial C}{\partial z}\right) - \frac{\partial}{\partial t}(uC) + R$$



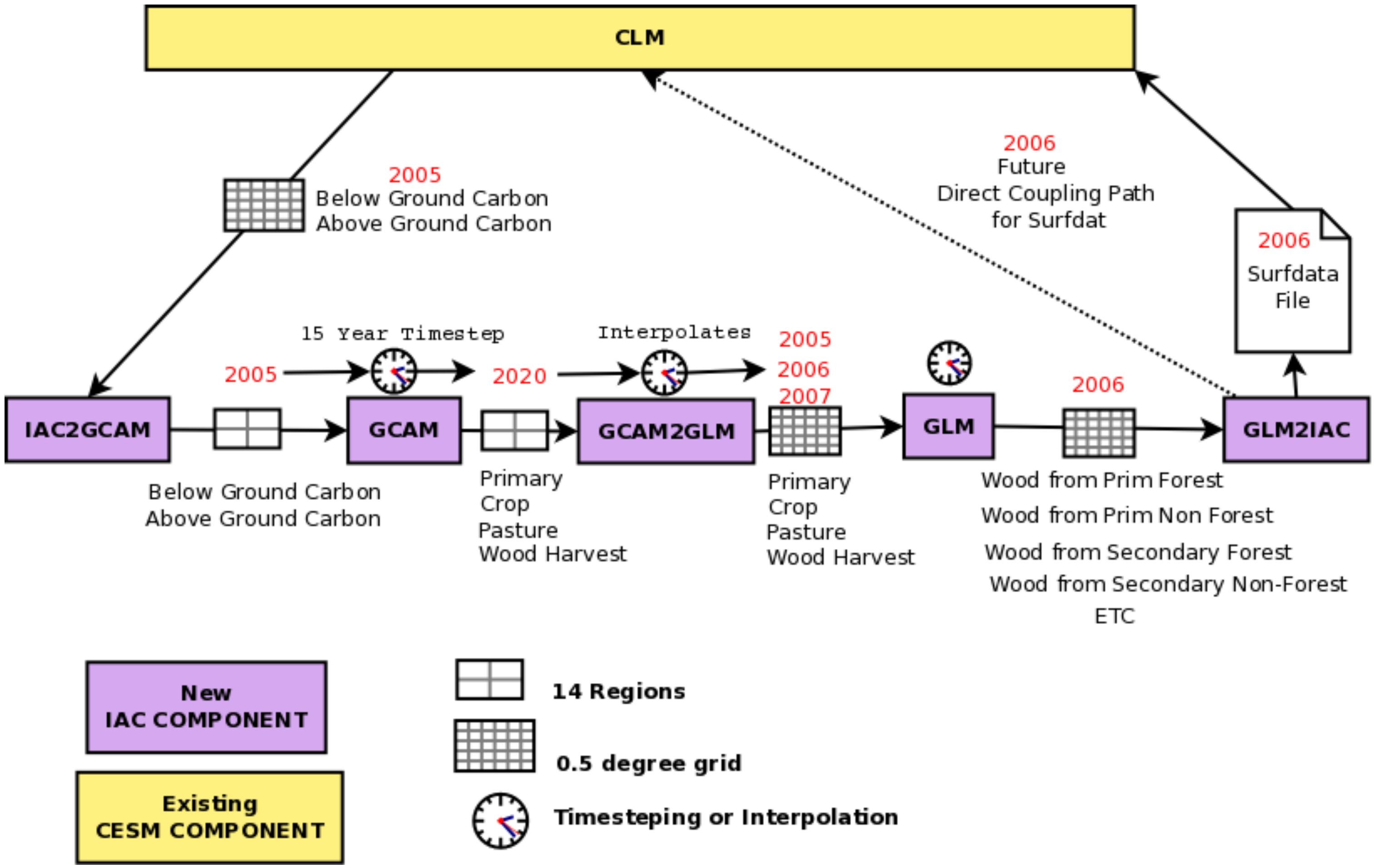
Predicted CO_2 surface flux and soil-gas concentration

iESM Schematic



iESM Coupling:

Emulation of Sneaker-net Coupling

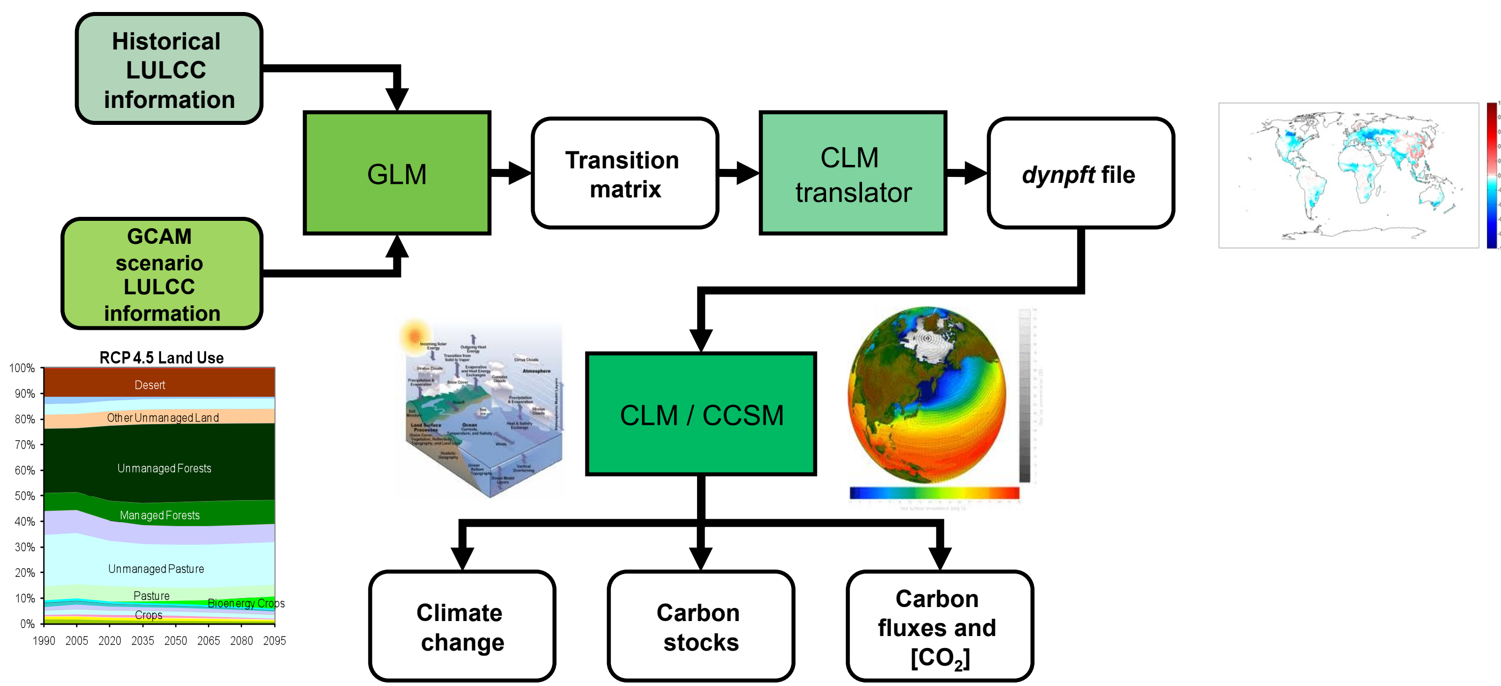


Status:

- We emulate sneaker-net using 15-year timesteps.

1. Project Description (cont.)

- By 2017 we expect to:
  - Develop probabilistic risks of abrupt climate change
  - Conduct local and regional projections of extreme rainfall
  - Simulate the $CO_2/CH_4/N_2O$ feedbacks in a warmer climate
  - Develop more integrated scenarios for climate mitigation



1. Project Description (cont.)

- Performance enhancement to support developments
  - SLR: $\geq$10X for high-resolution land-ice / ocean models
  - Extremes: 10 to 30X for superparameterized models
  - Chemistry: 10X for reactive chemistry and transport
  - Scenarios: 10 to 100X for scenario development

2. Computational Strategies

- We climate simulation computationally using models that solve the Euler equations, constituent equations, and 1st of thermodynamics for ocean, atmosphere, and ice.

- The primary code we use is the DOE-NSF joint Community Earth System Model (CESM): http://www.cesm.ucar.edu/

- Distinctive features of simulations:
  - *Duration:* Centuries to millennia
  - *Time steps:* Minutes (atmosphere) to hours (ocean)
  - *Experiments:* Response to time-evolving boundary conditions
  - *Metrics:* Non-deterministic statistics of the solutions

2. Computational Strategies (cont.)

- **CESM is comprised of 4 (now 5) components and a coupler:**
  - Atmosphere: (200 x 288 x 30 = 1.7e6)
  - Ocean: (180 x 360 x 40 x 0.7 = 1.8e6)
  - Sea & land ice (= ocean/land grids)
  - Land (200 x 288 x 10 x 0.3 = 1.7e5)

CCSM Graphic

- **The dynamical frameworks are / are evolving to:**
  - Atmosphere: Spectral element dycore on cubed sphere (SNL)
  - Ocean: Unstructured mesh/Voronoi tessellation (LANL)

- **Implementation of parallelism:**
  - Choice of MPI, OpenMP, MPI/OpenMP hybrid throughout.
  - Components run in arbitrary mix of serial and parallel processor layouts.
  - Parallel NetCDF for I/O.

2. Computational Strategies (cont.)

- **Our biggest computational challenges are:**
  - Ensemble sizes required for uncertainty quantification (1000s)
  - 100x increase in throughput required for cloud/eddy-resolving models
  - Barrier to long-time integrations from flat trends in clock rates

- **Current implementations exhibit scaling to $O(10^5)$ processors:**
  - CESM scales to 30K Cray / 60K Blue Gene cores *(Dennis et al, 2012)*
  - Spectral element dycore scales to 256K processors *(Taylor et al, 2011)*

- **Major changes anticipated / contemplated by 2017:**
  - GPU implementation of CESM components (underway for OLCF Titan)
  - New atmospheric/ocean dycores: focus on refinement, scalability
  - Implementation of stochastic parameterizations in atmosphere/ocean
  - AMR techniques for land ice

3. Current HPC Usage (see slide notes)

- **Machines currently running CESM:**
  - Major Facilities: NERSC, NCCS / OLCF, ALCF, NCSA, NCAR
  - Architectures: Cray XE/XT, IBM Power Series, IBM Blue Gene, Linux cluster

- **Hours used in 2012:**
  - IMPACTS/CLIMES: XX / YY M Core-hours
  - Other users: O(XXX) Core-hours at NCCS, ~40M for IPCC, ~140M @ NCAR

- **Typical parallel concurrency and run time, number of runs per year:**
  - Timing data: http://www.cesm.ucar.edu/models/cesm1.0/timing/

  - Hopper cores: 2064 (for 1-degree resolution)
  - Hopper core-hours: 2063 core-hours per year of simulation
  - Hopper throughput: 24.01 simulation years per wall clock day
  - Number of years/year 4000 to 10000 simulation years / calendar year

3. Current HPC Usage (see slide notes)

- **Data read/written per run and data resources used**
  - In IPCC AR5 production runs, 56.2 GB/sim. month and 675 GB/sim. Year
  - Storage system: HPSS, 3.46M SRUs = 725 TB

- **Memory used per (node | core | globally)**
  - 135 GB (1850 carbon/nitrogen compset, (*Intel Benchmark, for HPC Advisory Council)*

- **Necessary software, services or infrastructure**
  - UNIX like operating system (LINUX, AIX, OSX)
  - csh, sh, perl, and xml scripting languages
  - subversion client version 1.6.11 or greater
  - Fortran 90 and C compilers. pgi, intel, and xlf are recommended options.
  - MPI (although CESM does not absolutely require it for running on one processor only)
  - netcdf 3.6.2 or greater
  - Earth System Modeling Framework (ESMF) (optional) 5.2.0p1
  - pnetcdf (optional) 1.1.1 or newer

4. HPC Requirements for 2017

(Key point is to directly link NERSC requirements to science goals)

- **Compute hours needed (in units of Hopper hours)**
  - 2012 ERCAP request = 11.9M (CLIMES) + 26.3M (IMPACTS) = 38.2M
  - Estimate for 2017 = (8/2) * 38.2M = 150M

- **Changes to parallel concurrency, run time, number of runs per year**
  - Parallelism : 2K -> 20K processors per integration
  - In principle, 100K processors per integration feasible

- **Changes to data read/written**
  - Data per sim. Year = 8^(2/3) * 56Gb/sim. year = 200 GB/sim. Year
  - Total volume = 4 * 725TB /calendar year = 3 PB / calendar year

- **Changes to memory needed per (core | node | globally)**
  - Memory required = 4 * 135 GB = 540 GB for entire simulation

5. Strategies for New Architectures

- **Our strategy for many-core architectures is collaboration with FASTMATH and SUPER Institutes via SciDAC climate apps:**
  - Transport and advection (led by ORNL)
  - Land ice (LANL)
  - Multiscale physics and integration w/new dycores (LBNL)

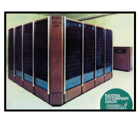
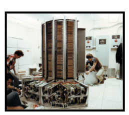
- To date we have prepared for many core by implementing capabilities for arbitrary hybrid MPI / OpenMP parallelism and end-to-end MPMD architecture.

- The CESM project is committed to porting to MIC machines, including the TACC Stampede machine based on Knights Corner.

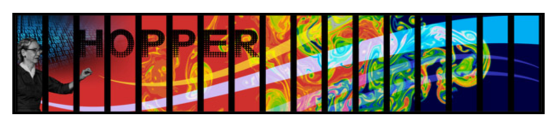
5. Summary

- What new science results might be afforded by improvements in NERSC computing hardware, software and services?
  - **Better understanding of land-ice dynamics of Antarctica and implications for SLR**
  - **Exploratory studies of multiscale climate dynamics from cloud/eddy to global scales**
  - **Model/data fusion for studies of the carbon cycle and carbon-climate feedbacks**
  - **Linking robust national energy strategies to needed advances in climate science**

- Recommendations on NERSC architecture, system configuration and the associated service requirements needed for your science
  - **"Leading without bleeding" procurement strategy of NERSC works well for our apps, since we can leverage substantial DOE and NSF investments in performance portability.**



5. Summary

- NERSC generally refreshes systems to provide on average a 2X performance increase every year. What significant scientific progress could you achieve over the next 5 years with access to 32X your current NERSC allocation?
  - **Advances towards global eddy-resolving projections of sea-level rise**
  - **Initial exploration non-hydrostatic cloud-system-resolving climate models**
  - **Development of regional-to-global carbon/water/climate analyses of the Earth system**
  - **Integrated climate/energy scenarios for the Sixth IPCC report and national assessments**

- What "expanded HPC resources" are important for your project?
  - **Integration of provenance tracking throughout software / project / data cycles.**
  - **“Rendering engines” (hardware and/or MPP software) for exabyte data sets**
  - **Multi-terabyte/second networks to key partners including LCFs, NCAR, etc.**