Coupled High-Resolution Modeling of the Earth System

Talk in support of NOAA-DOE Climate Computing proposal
Department of Energy, Germantown MD

V. Balaji

Princeton University

NOAA/GFDL

20 July 2007

Outline

1. NOAA-DOE mission goals

2. Why high resolution?
   - Regional scales are better represented
   - Fundamental new physics in atmosphere and ocean
   - Hurricanes and climate change

3. Proposed experiments on ORNL Leadership Computing Facility
   - NOAA models for these studies
   - Coupled model experiments and target resolutions
   - Schedule and requirements

4. Current status

Outline

1. NOAA-DOE mission goals

2. Why high resolution?
   - Regional scales are better represented
   - Fundamental new physics in atmosphere and ocean
   - Hurricanes and climate change

3. Proposed experiments on ORNL Leadership Computing Facility
   - NOAA models for these studies
   - Coupled model experiments and target resolutions
   - Schedule and requirements

4. Current status

NOAA-DOE mission goals

NOAA and DOE share goals in understanding and predicting climate change and its impact. This joint activity will

- support NOAA"s mission to understand and predict changes in the Earth's environment by
  - improving climate predictive capability from weeks to decades with an increased range of applicability for management and policy decisions
  - developing and contributing to routine state-of-the-science assessments for informed decision making
- support DOE's research and modeling efforts to
  - predict accurately any global and regional climate change induced by increasing atmospheric concentrations of aerosols and greenhouse gases
  - improve the scientific basis for assessing both the potential consequences of climatic changes, including the potential ecological, social, and economic implications of human-induced climatic changes caused by increases in greenhouse gases in the atmosphere and the benefits and costs of alternative response

Outline

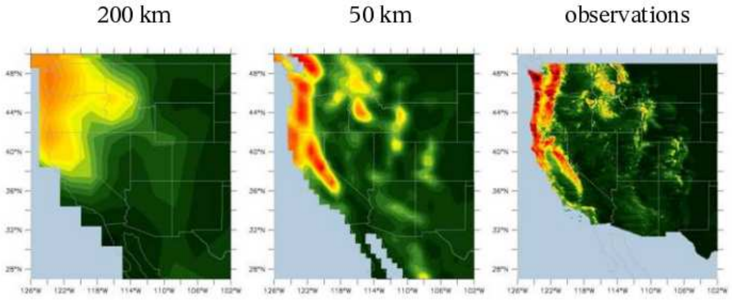
1. NOAA-DOE mission goals

2. Why high resolution?
   - Regional scales are better represented
   - Fundamental new physics in atmosphere and ocean
   - Hurricanes and climate change

3. Proposed experiments on ORNL Leadership Computing Facility
   - NOAA models for these studies
   - Coupled model experiments and target resolutions
   - Schedule and requirements

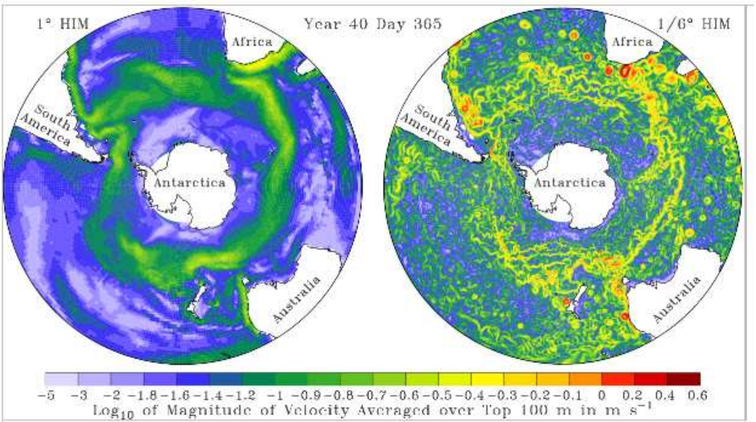
4. Current status

Regional scales are better represented



- There is a dramatic improvement in our ability to model regional scale climate response as we go to “high” (i.e beyond the IPCC AR4 norm) resolution.
- However, with current computing capacity, NOAA cannot perform the long control integrations needed for detection and attribution of climate change above the threshold of natural variability.

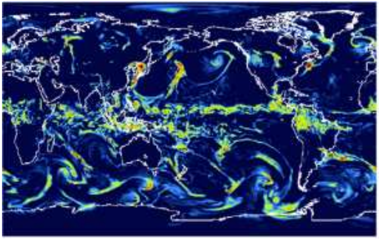
Ocean eddies

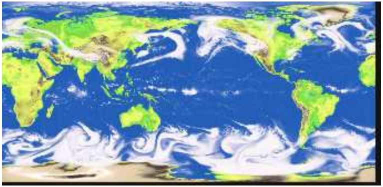


Shaded is the speed of the surface currents in the ocean model at two resolutions.

At 15 km resolution, turbulent eddies in the ocean are directly simulated, and show fundamentally different heat transport on climate scales.

Tropical storm systems; mid-latitude fronts; clouds





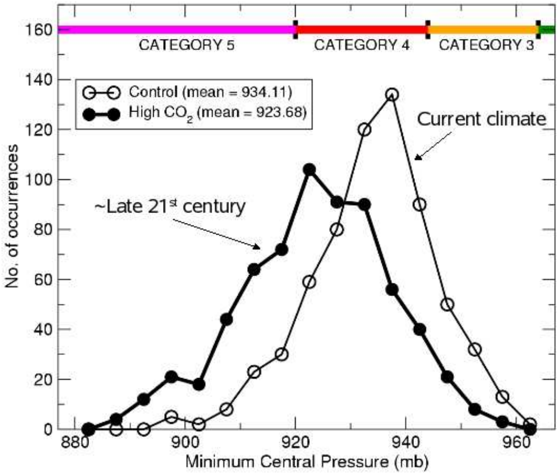
25 km

At 25 km resolution, we can study the statistics of tropical cyclones under climate change.

5 km

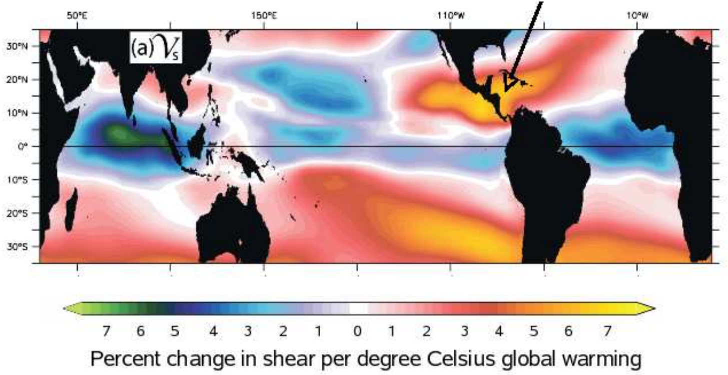
At 5 km resolution, we arrive at the long-awaited “global cloud-resolving model”, and the potential to dispense with paramaterizing deep convection in climate models.

Hurricane intensities are projected to increase. . .



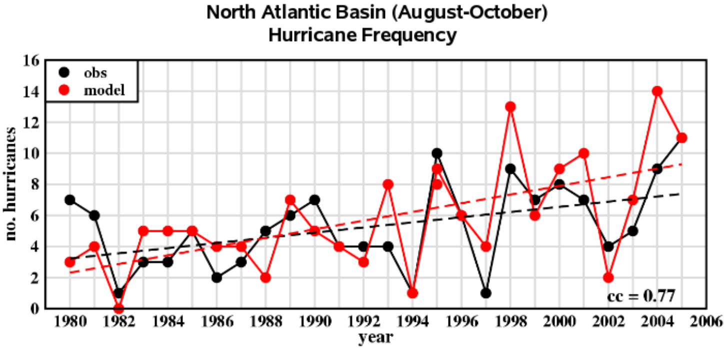
Hurricane models forced by CMIP data project increased intensities in a warmer world (Knutson and Tuleya 2004).

However, hurricane frequencies might decrease…



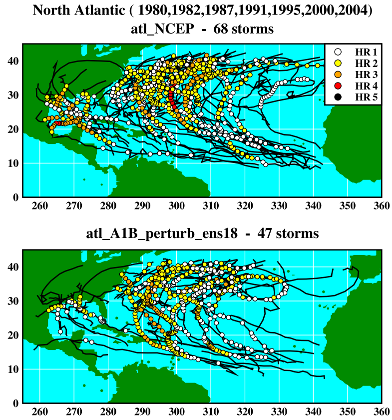
Vecchi and Soden (2007) show wind-shear increasing in a warming world, potentially leading to a decrease in Atlantic hurricane frequency (though not elsewhere…)

NOAA high-resolution models capture hurricane statistics



NOAA regional model ZETAC captures inter-annual variability in hurricane frequency when forced with historical data (Knutson et al 2007). This study is being repeated now for a warming world from IPCC AR4 data to confirm or refute Vecchi and Soden.

Vecchi and Soden results corroborated



Preliminary regional model results show reduced Atlantic hurricane frequency in the late 21st century. Forced regional model results need to be supported by global coupled models for a complete understanding of this key result. NOAA models are ready to make the leap given enough computing and analysis power.

Science in support of NOAA-DOE mission goals

- High-resolution models can be used to provide improved understanding of regional climate change.
- Understanding and analyzing the inherent predictability of models at high resolution will help design international modeling campaigns aimed at climate prediction.
- Understanding tropical storm statistics using high-resolution global models provides insight in planning for extreme events in a warming world.
- The science therefore fulfils the NOAA and DOE missions of providing a *predictive understanding* of climate change, and providing *credible and timely information for decision makers* in preparing for climate change.

Outline

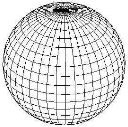
1. NOAA-DOE mission goals

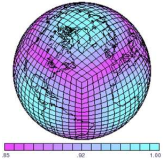
2. Why high resolution?
   - Regional scales are better represented
   - Fundamental new physics in atmosphere and ocean
   - Hurricanes and climate change

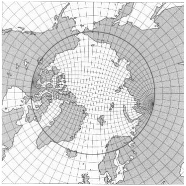
3. Proposed experiments on ORNL Leadership Computing Facility
   - NOAA models for these studies
   - Coupled model experiments and target resolutions
   - Schedule and requirements

4. Current status

NOAA Models and the FMS Mosaic infrastructure







- The Finite-Volume conventional grid dycore (`FVLL`) is used in CM2.4. The Fourier filter at the pole limits scalability on distributed memory.
- The Finite-Volume Cubed-Sphere dycore (`FVCS`) eliminates the pole and vastly increases scalability on distributed memory.
- The ocean model in these experiments is `MOM4`, running on a tripolar grid: also has no pole problem. The newer `GOLD` model also is tripolar.
- Parallelism in all the models is provided by the FMS `Mosaic` infrastructure, which handles parallel I/O and communication (`MPI`, `shmem`, threads).

Coupled model experiments and target resolutions

Experiments are designed to understand model behaviour and climate predictability at high resolution, and to study year-over-year changes in hurricane frequency and intensity in a warming world.

- Control integrations of high-resolution versions of `CM2.1`, the NOAA/GFDL flagship model. `CM2.4` will double atmospheric resolution to 1° and quadruple the ocean resolution to 0.25°. `CM2.5`: atmospheric resolution increased to 0.5°.
- $2\times CO_2$ control of `CM2.4` and `CM2.5`.
- AMIP runs of `AM2` model configured with next-generation atmospheric dynamical core ("Finite-Volume Cubed-Sphere") at 0.5° and 0.25°.
- "Time-slice" experiments of the high-resolution AM2 models above forced with data from the $1\times CO_2$ and $2\times CO_2$ `CM2.4` runs.
- Eventually, we plan to build coupled models using the cubed-sphere dycore.

Projected performance and throughput

Model	PE Count	model y/d	CPU-h/m	Data out (GB/d)
CM2.4	480	2	345600	550
AM2C200	500??	4	360000??	350
AM2C400	1000??	2	720000??	700

- Projections are based so far on SGI Altix results scaled by processor speed ratio between Itanium and Opteron.
- Throughput numbers assume dedicated CPUs: no queue wait time.
- `CM2.4` results assume a successful transition to Compute Node Linux, and Cray implementation of remote addressing.
- `AM2C200` is expected to reach 50% scalability at 2560 cores; `AM2C400` at 10240 cores. Actual PE counts to be decided after further measurement.
- We plan to stream data back to GFDL for post-processing and analysis: minimum sustained bandwidth requirement of 150 Mbps not counting software stack (encryption, firewall, checksums, ...).

Aggregate costs of proposed runs

- If these projections hold up, one run each of CM2.4, AM2C200, AM2C400 would require 2000 dedicated cores (1.44M CPU-hours/month).
- Rule of thumb for "scientifically useful" runs is a minimum of 3-4 model-years/day (100 y/month). Current projected throughput is marginal and assumes dedicated CPUs: no queue wait time.
- Aggregate data bandwidth of 1.6 TB/day may be an issue.

Proposed schedule

- FY2007: 1M CPU-h/month.
  - Complete project initialization.
  - Performance studies of CM2.4 and CM2.5, and FVCS at C48, C90, C200, C400 and C2000 (dry only).
  - Production runs of CM2.4SP and CM2.4SQ (these are different projections used in the ocean: one has fewer gridpoints, and may be used in future if climates are similar). Timeline has a dependency on CNL installation schedule.
  - Production runs of FVCS AMIP runs at $2^\circ, 1^\circ, 0.5^\circ, 0.25^\circ$.
- FY2008: 1.44M CPU-h/month.
  - Production runs of CM2.4 and CM2.5 for $1\times CO_2$ and $2\times CO_2$ continue.
  - Time-slice experiments of AM2C200 and AM2C400 forced by output from runs above.
  - Configuration and performance studies of coupled models using FVCS and next-generation ocean model GOLD (C-grid hybrid vertical coordinate model, currently in validation against MOM4 in climate mode at GFDL).

FY2008 and beyond

- Decadal (and longer) predictability studies and hurricane statistics under climate change both require ensemble runs. We could easily use as much as is made available given other constraints such as data volume. Perhaps 2000 dedicated CPUs plus more time in the general queues. . . ?
- Public dissemination of data from high-resolution models: this would be a valuable community resource in planning the next cycle of international modeling campaigns.
- And beyond. . .
    - Global non-hydrostatic FVCS (C2000) with full cloud physics enabled.

Outline

1. NOAA-DOE mission goals

2. Why high resolution?
   - Regional scales are better represented
   - Fundamental new physics in atmosphere and ocean
   - Hurricanes and climate change

3. Proposed experiments on ORNL Leadership Computing Facility
   - NOAA models for these studies
   - Coupled model experiments and target resolutions
   - Schedule and requirements

4. Current status

Status report

- Project request filed; account requests underway. (One GFDL-er already has an account).
- `FVCS`, `FVLL`, `MOM4` and `GOLD` have been run on `jaguar`. (Recently ran `FVCS` dry dycore at `C2000`: 5 km global!)
- Met Cray engineers on 13 July 2007 to discuss CNL transition and implementation of on-node remote addressing across dual-core. We have also begun to work together with Cray engineers at Eagan and Oakridge on improving single-processor performance. Regular meetings have been scheduled.
- Data transfer plans need to crystallize: contact has been initiated between GFDL and ORNL networking teams. Preliminary tests on transfer using simple `ftp` show insufficient bandwidth.