# Climate in a Box to Support NASA Earth System Science

NASA Headquarters/Tsengdar Lee

Presented at AGU

December 17, 2009

# Climate-in-a-Box Team

**NASA-Headquarters**
Tsengdar Lee

**NASA-Goddard**
Mike Seablom
Gail McConaughy
William Putman
Greg Shirah

**AMTI, Inc.**
Rahman Syed
Hamid Oloso

**Northrop Grumman Corporation**
Eric Kemp
Joe Greenseid
Ryan Smith
Rob Burns
Shawn Freeman
Gary Wojcik

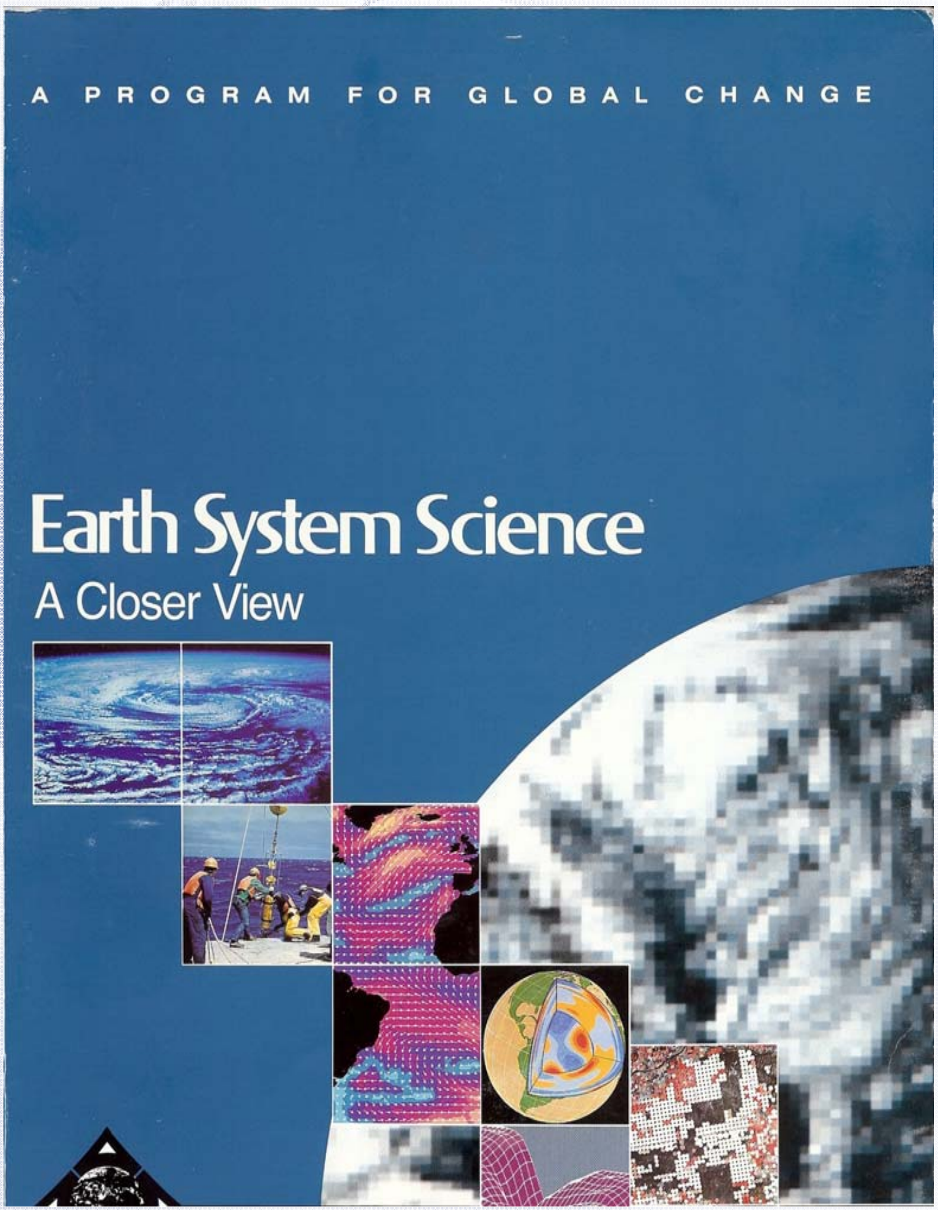
**GST, Inc.**
Ramon Linan
John Evans

# Acknowledgements

- GSFC's Office of the Chief Technologist-Internal Research and Development program
- NASA's Earth Science Technology Office-Advanced Information Systems Technology program
- GSFC's Codes 610, 581, and 583

# The Bretherton Report:

## Earth System Science's Founding Text

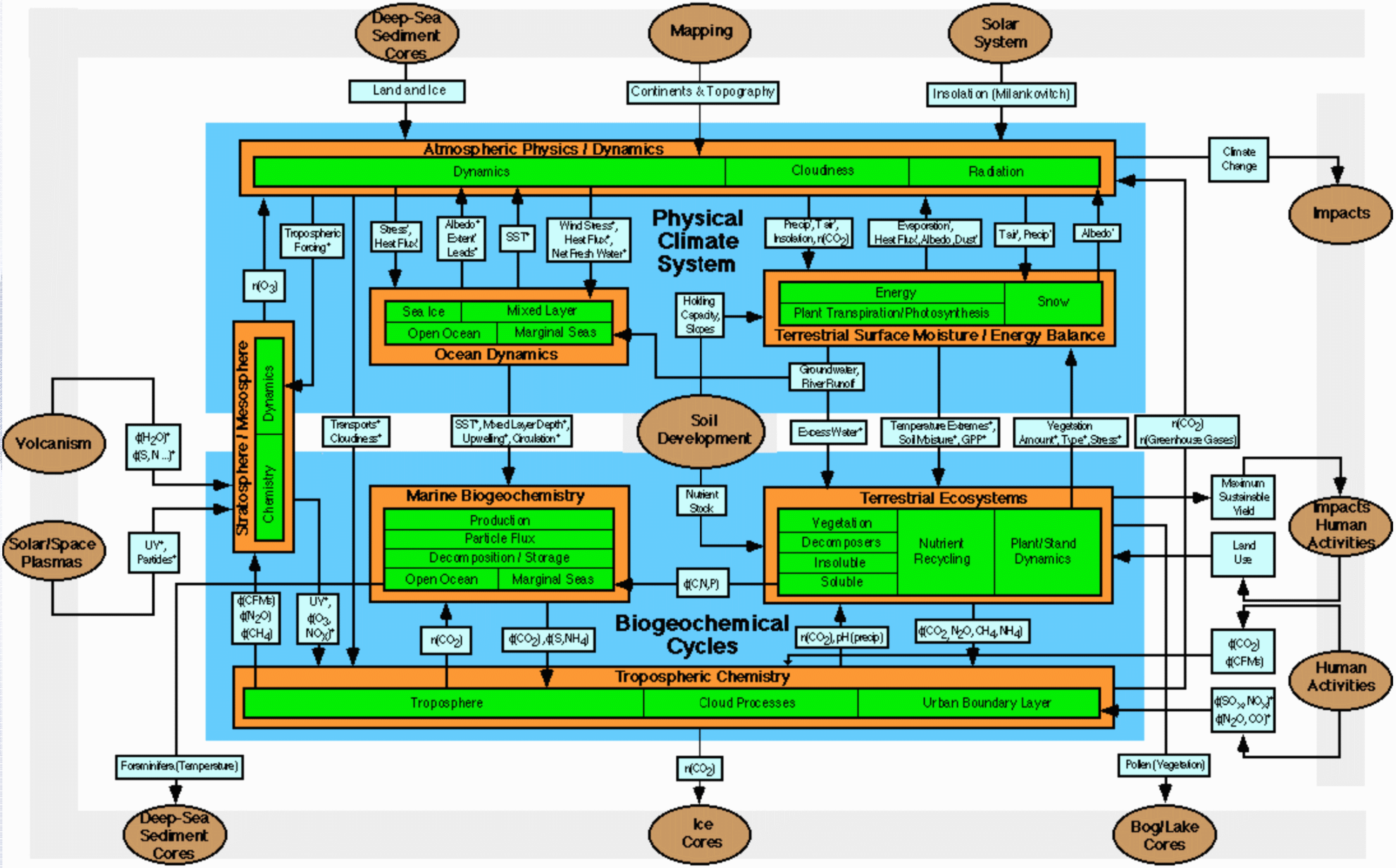



- The Goal of Earth System Science: To obtain a scientific understanding of the entire Earth System on a global scale by describing how its component parts and their interactions have evolved, how they function, and how they may be expected to continue to evolve on all timescales.

- The Challenge of Earth System Science: To develop the capability to predict those changes that will occur in the next decade to century, both naturally and in response to human activity.

# Bretherton et al.'s Diagram

CONCEPTUAL MODEL of Earth System process operating on timescales of decades to centuries

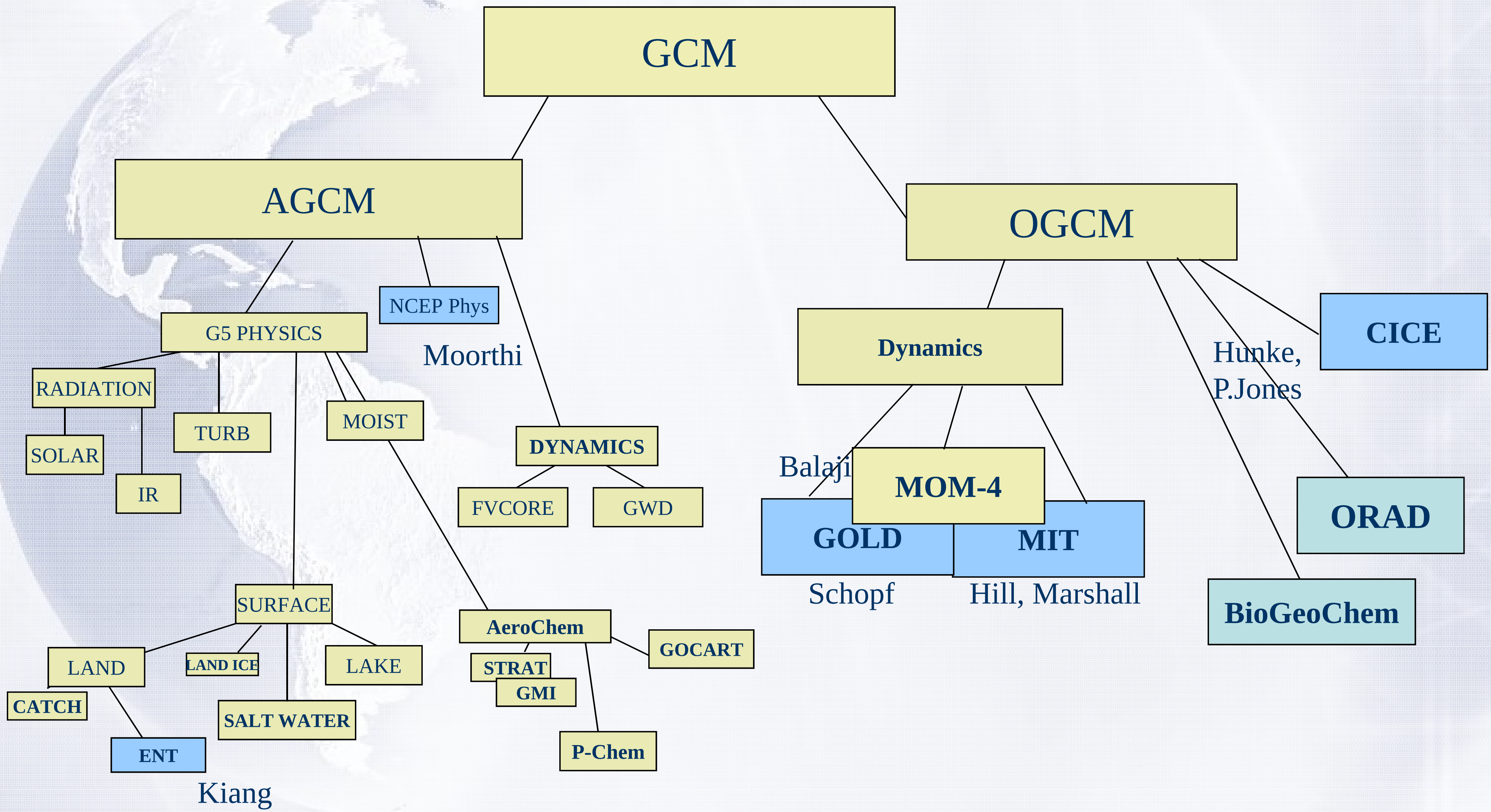



' = on timescale of hours to days * = on timescale of months to seasons ϕ = flux n = concentration

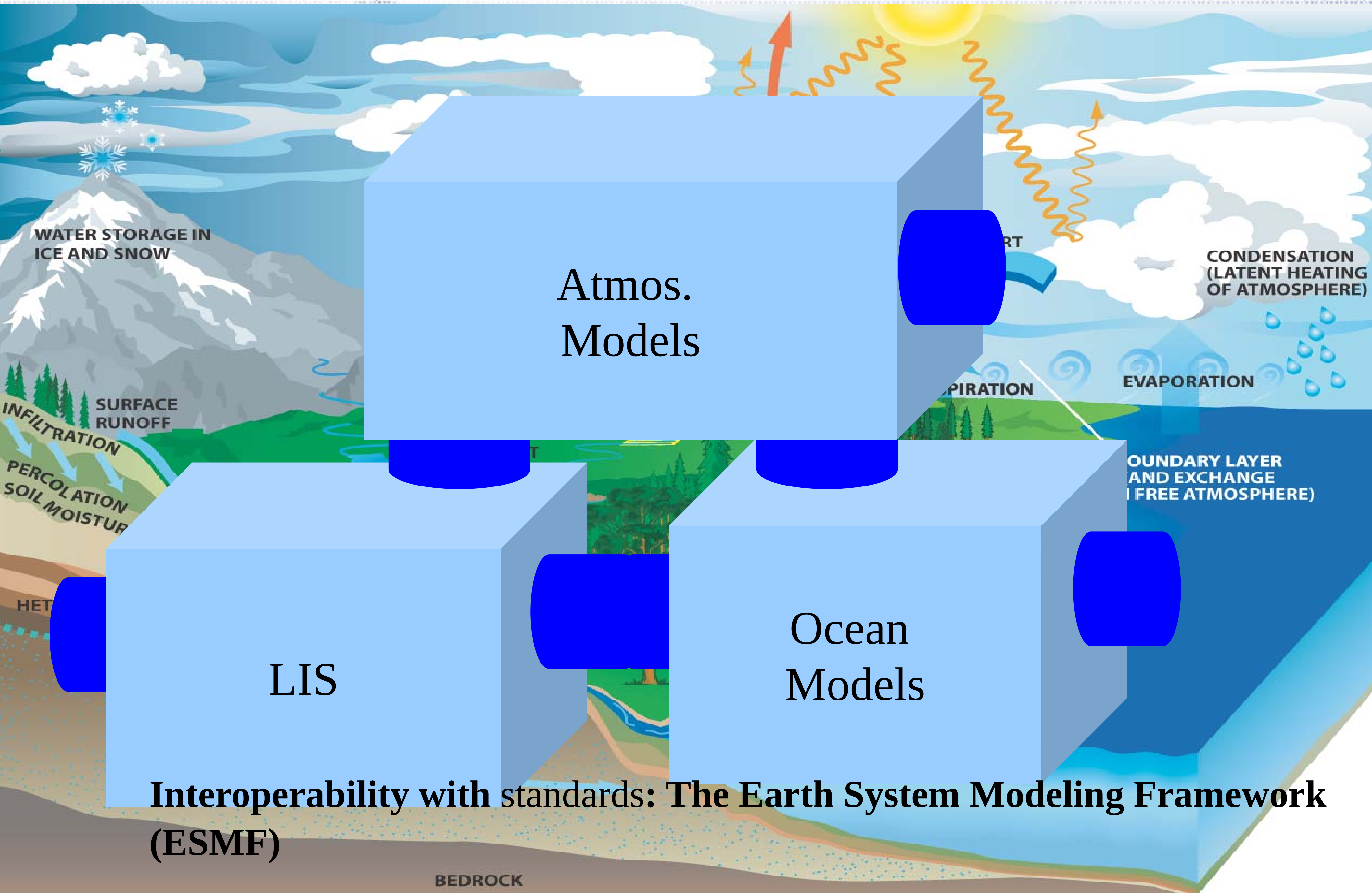
*(From: Earth System Science: A Closer View (1988))*

# GEOS-5 GCM

## Structure & External Contributions

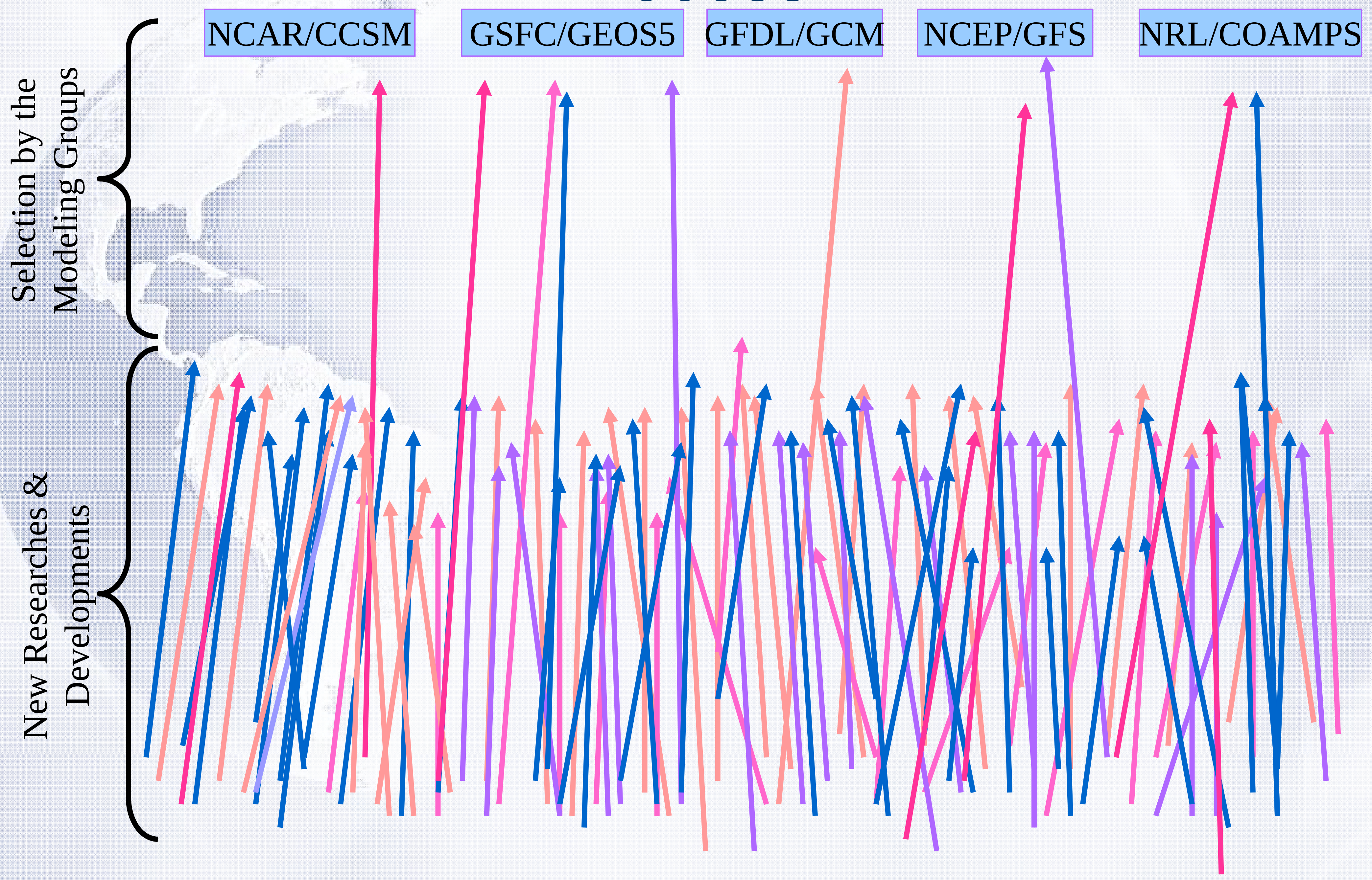


# Simplified View of Bretherton et al.'s Diagram




**Interoperability with** standards**: The Earth System Modeling Framework (ESMF)**

# Earth System Models Improvement Process



# The Transition Challenge

- The cost to transition is under appreciated by the R&D community.
  - Recent NSF cyber-infrastructure study estimated 1:10 ratio for R&D vs. transition costs
- Operating centers do not have enough money to transition many R&D results to the operation environment.
- R&D PIs are frustrated by not getting their innovation accepted.

# Challenges in Climate Model Development

- It's a natural tendency for climate models to become more and more "elaborate."
- It is difficult to verify and validate the complex models.
- Climate model code development is tightly controlled by selected few organizations.
- Hierarchical structure inhibits community inputs into the core model.
- As the community grows, the challenge becomes unmanageable.
- Need an **agile governance model** and a **reward system** that encourage community engagement and allow a community selection process.

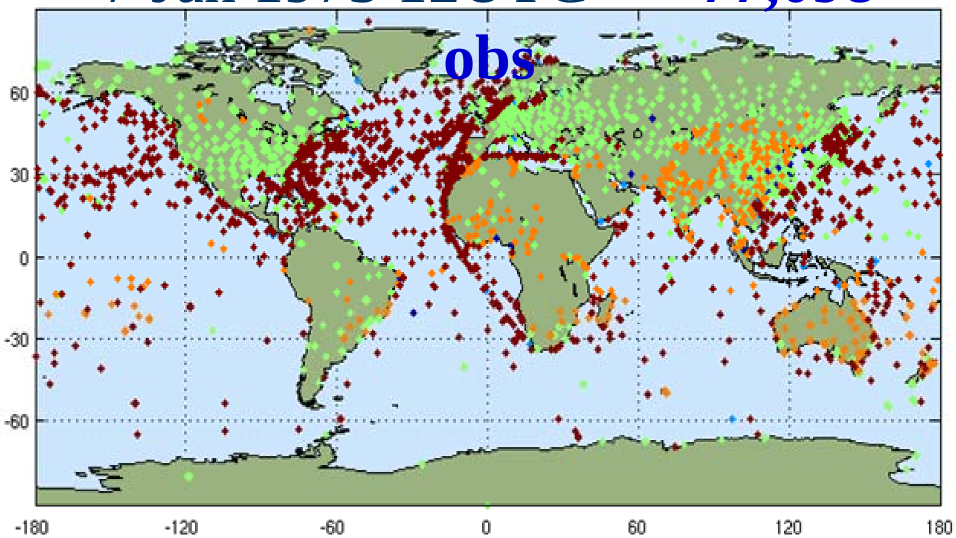
# IPCC 5th Assessment Report

- IPCC AR5 report due 2013. Model runs due end of 2010.
- Observations play a critical role in climate research
  - Process understanding
    - Exploratory data analysis
    - Hypothesis formulation
  - Parameterization and model development
    - Statistical description of sub-grid-scale processes
    - Hypothesis testing
  - Model evaluation (IPCC WG1)
    - Comparison of model output against observations
    - Weighting multi-model ensemble members (“scoring")
- Growing IPCC requirements in mitigation and adaptation

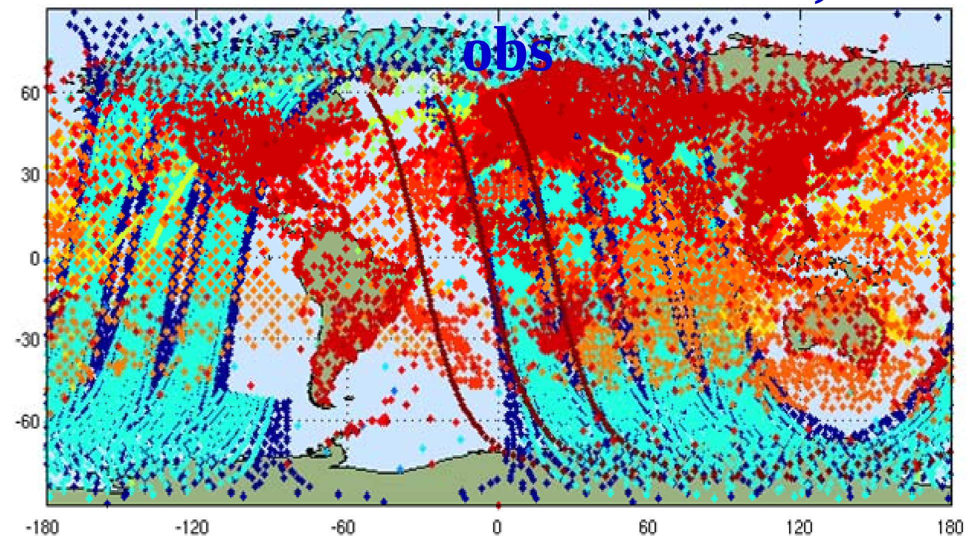
# The Changing Observing System…6hr snapshots through time

## GEOS-5 Data Assimilation System

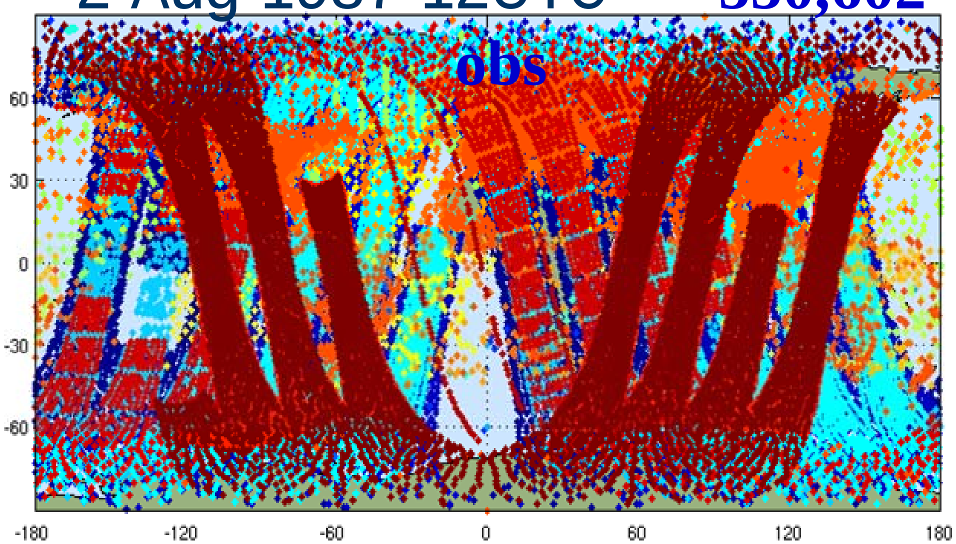
7 Jan 1973 12UTC **77,098**



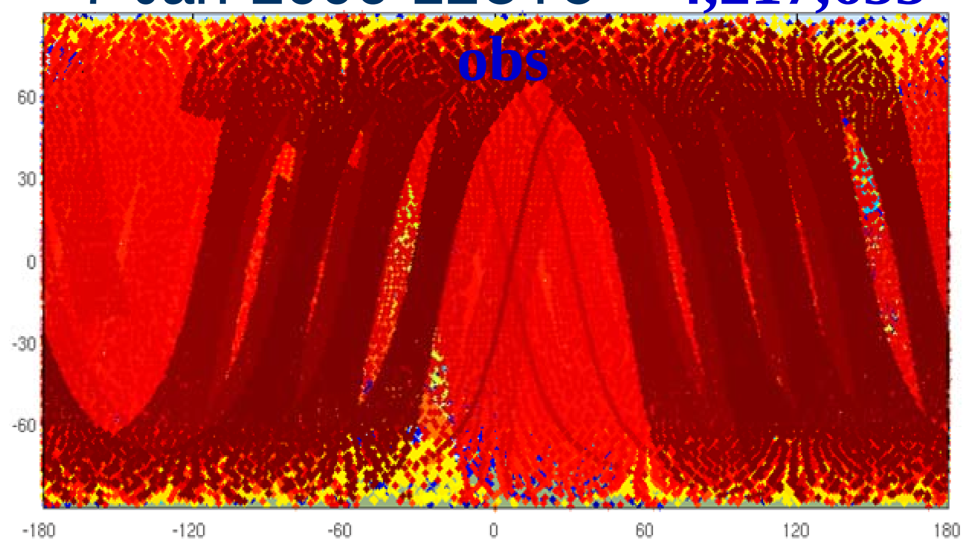

7 Jan 1979 12UTC **325,765**




2 Aug 1987 12UTC **550,602**



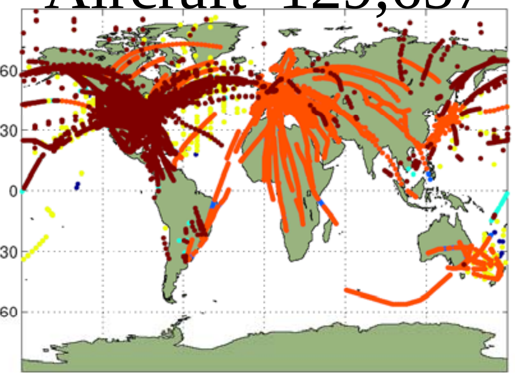

7 Jan 2006 12UTC **4,217,655**



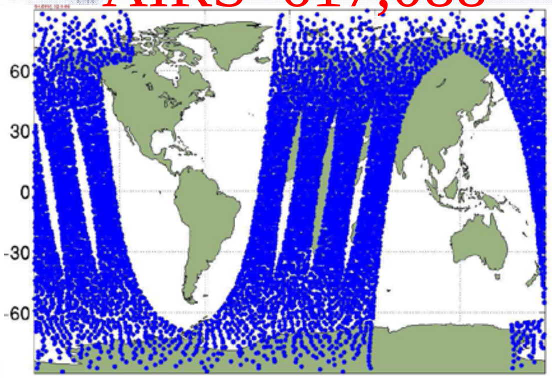
## Main Observing Systems Assimilated in GEOS-5
6-hr window centered at 00 UTC 11 Nov 2007

Operational
Research (NASA)
Operational+Research

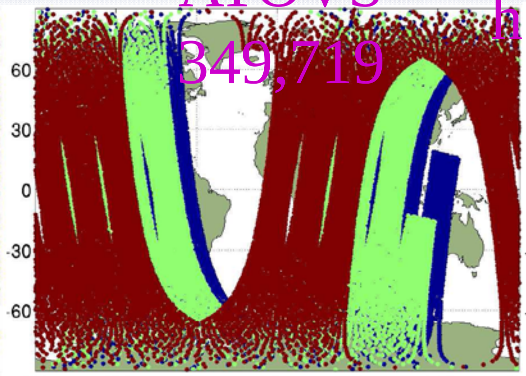
Aircraft 129,657



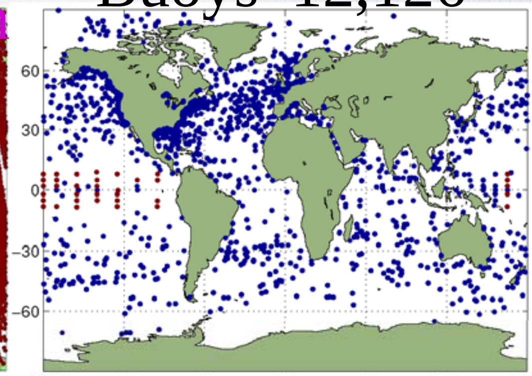
AIRS 617,088



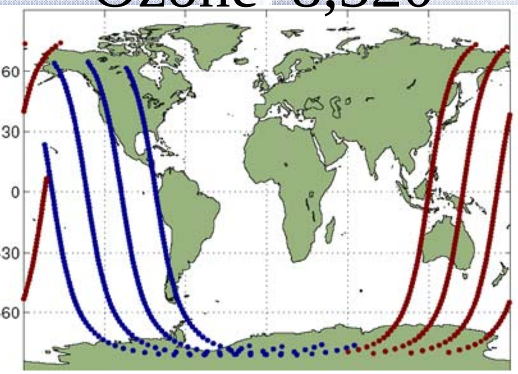
ATOVS 349,719



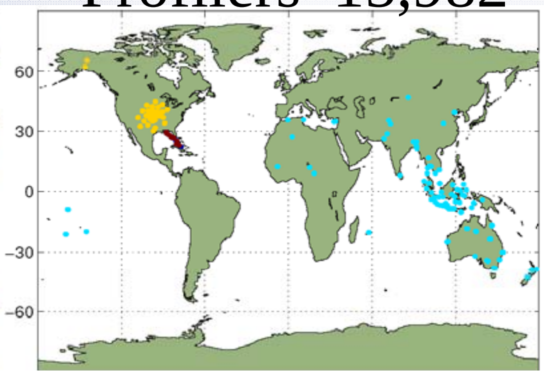
Buoys 12,126



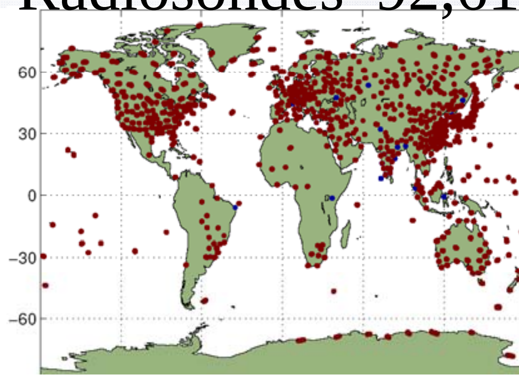
Ozone 8,320



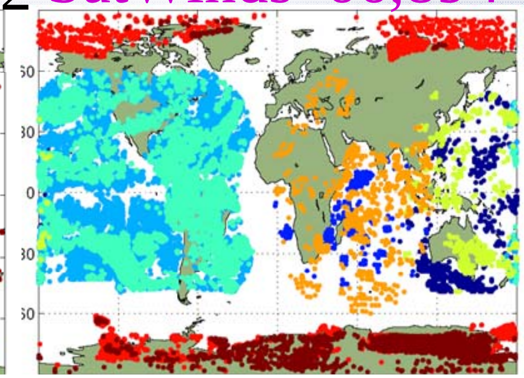
Profilers 15,982



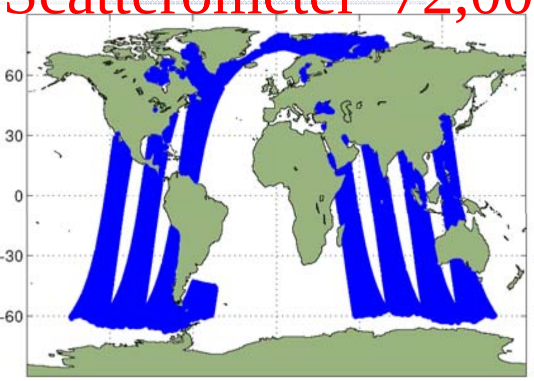
Radiosondes 92,612



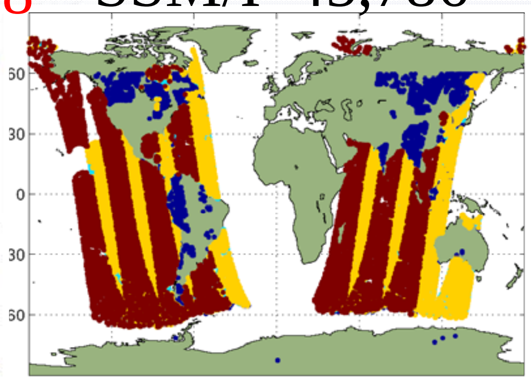
SatWinds 66,894



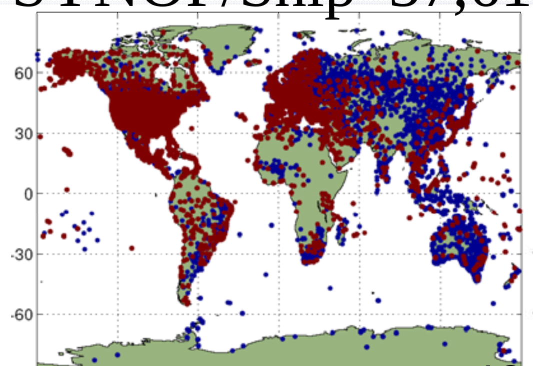
Scatterometer 72,008



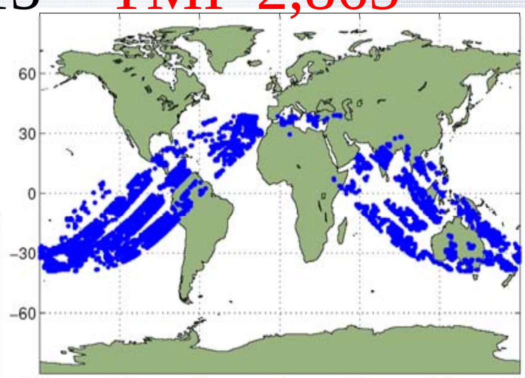
SSM/I 45,786



SYNOP/Ship 37,615

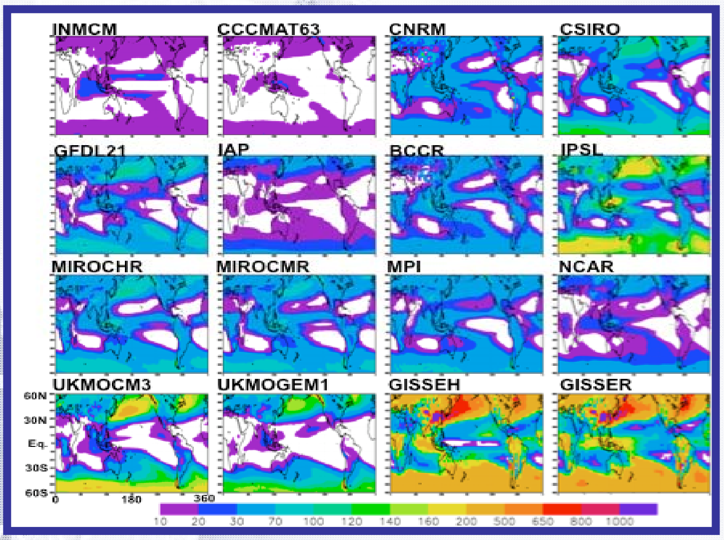


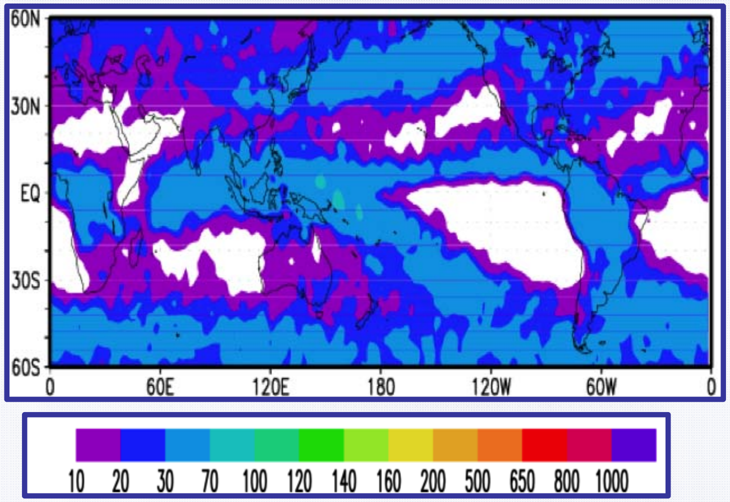
TMI 2,865



# Improvements in Climate Models with CloudSat: Ice concentration and precipitation frequency

Climate model estimates of cloud ice concentrations

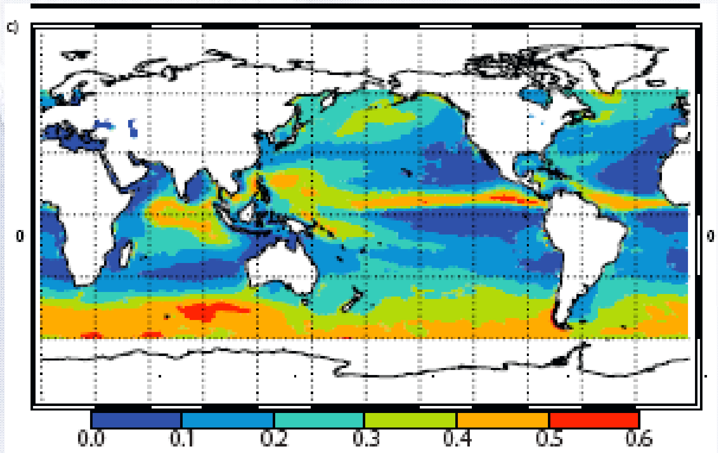


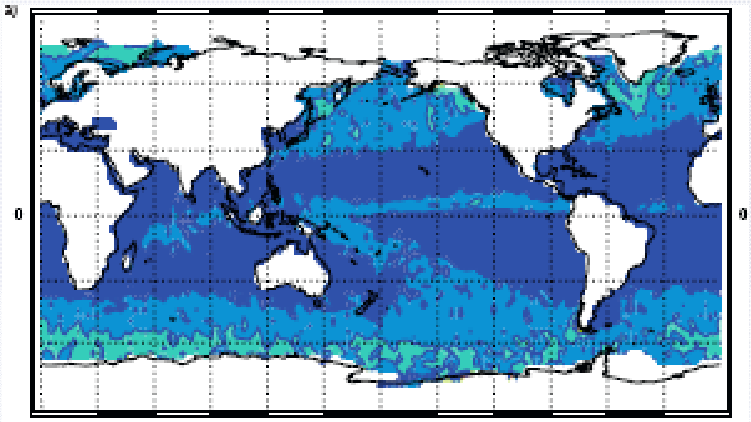




CloudSat- global ice concentrations in gm m-2

Waliser et al, 2008

UKMO– 24.4% average frequency of precipitation over the oceans



CloudSat - 11% average frequency of precipitation over the oceans

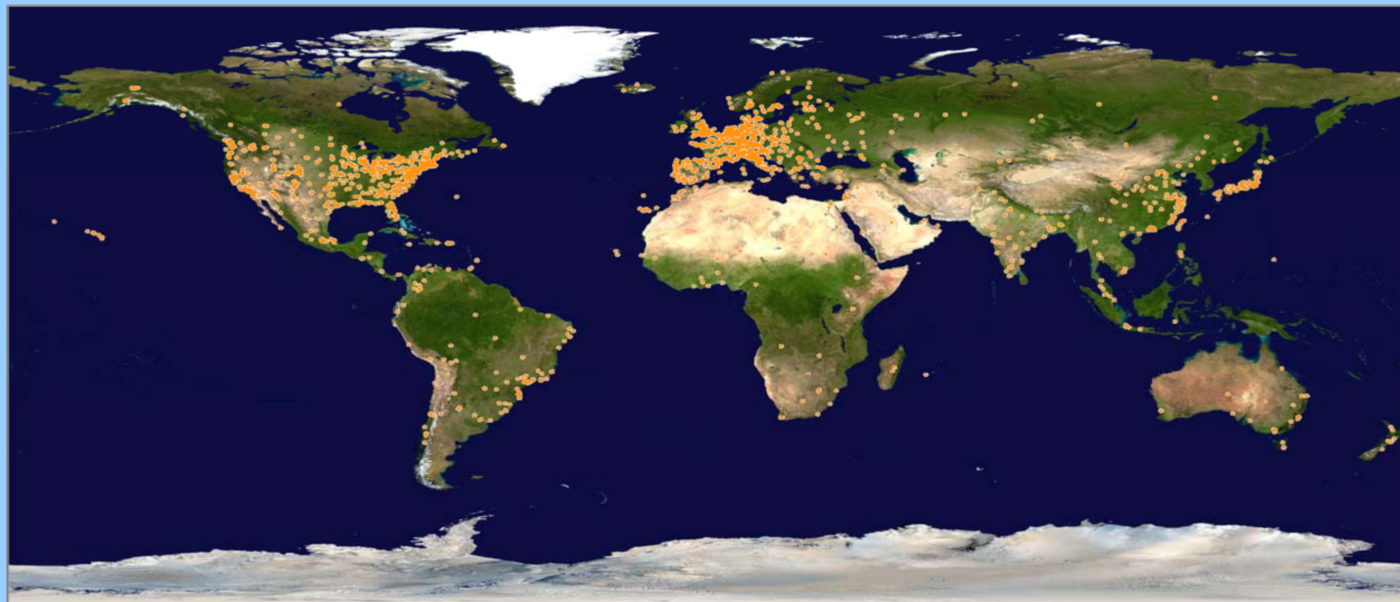
Cloudsat provides observations of global cloud ice (top) and rainfall occurrence (bottom) to constrain models

# Science Gateway

## Earth System Grid
Center for Enabling Technologies

**Agencies**

● Registered sites



*ESG-CET enables Scientific Discovery in Climate Science by providing an international community of over 16,000 registered users with climate simulation data, climate models, analysis and visualization tools, and enabling technologies for a distributed, global science enterprise*

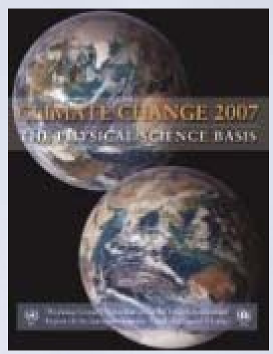
## ESG turns climate science data into community resources

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE

WMO UNEP

*Data warehouse, search and discovery, access, and reduction*

*Data used in hundreds of scientific papers* →




*Much of which provided a basis for the 4th Assessment Report of the IPCC* →

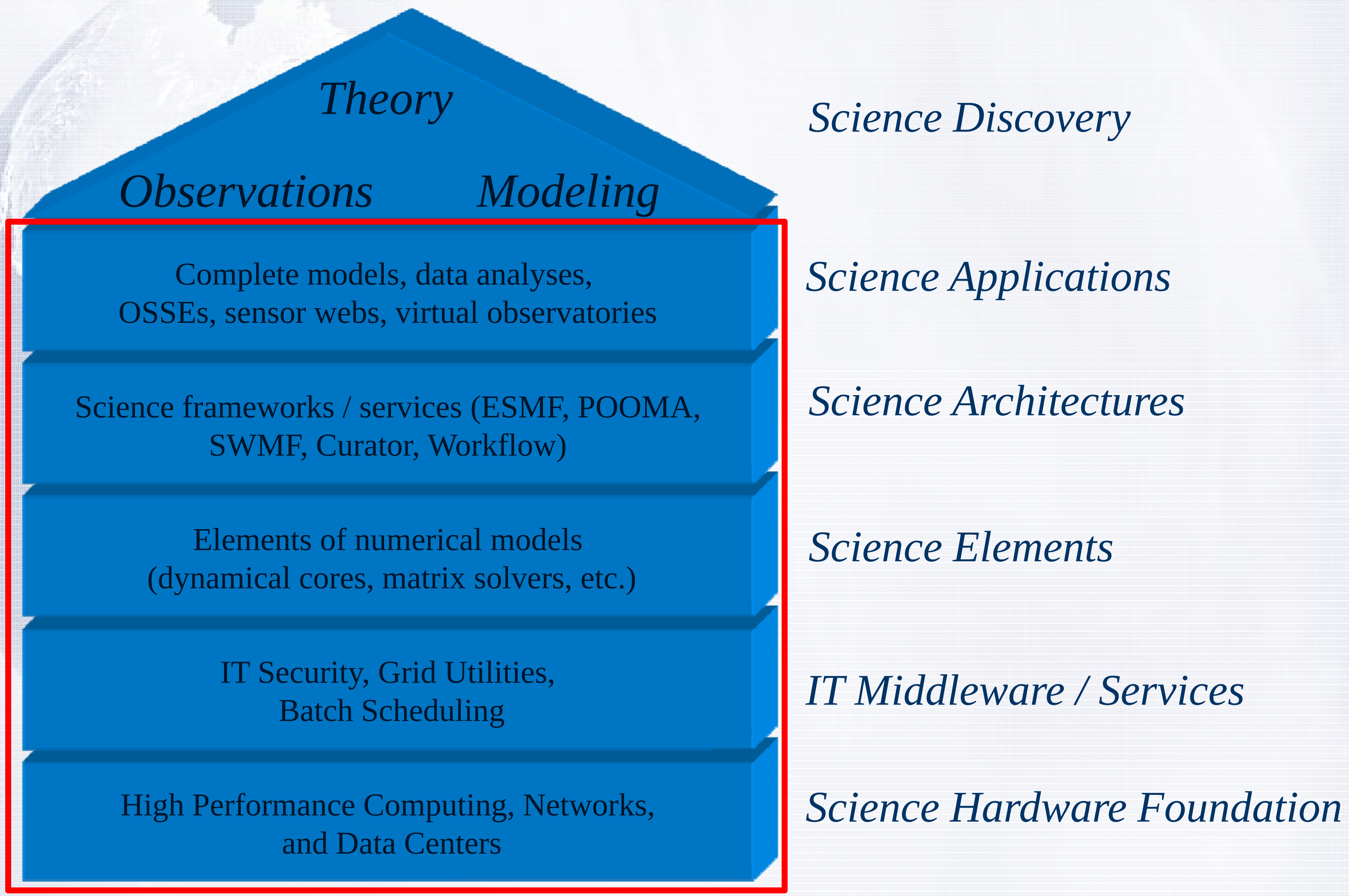
The Nobel Peace Prize 2007

Intergovernmental Panel on Climate Change (IPCC)

# Challenges in Earth System Science

- Earth System is a complex system
- Earth System Science requires significant system engineering discipline.
- Use of software framework is crucial in the success of future Earth System Science.
- Data will be used to constrain models.
- The staging of data has become a significant portion of Earth system science research.

# Science Information System Hierarchy



# Vision

- Climate in a Box (CIB) seeks to:
  - Open climate/Earth science model development and validation to a community beyond traditional domain scientists.
  - Develop/improve models through a more efficient “open” model development and validation process.
  - Involve a much bigger climate application user community.

# Climate in a Box Project Concept

- Climate models and model outputs will be used by application user communities and decision support communities
- Significant need to create a common framework to connect different communities
- Since model and application developers are good at creating sandboxes, Climate in a Box provides a playground for the sandboxes.

# Climate in a Box Goals

- Provide users with tools to assist them in their work
  - workflow tools
  - visualization and analysis tools
  - ancillary data system, validation data set, and test scenarios
  - data/security system
    - automate code updates and standardized testing
    - data transfer/storage
- Develop open/community model development structure
  - Use Web 2.0 technology to facility knowledge management/transfer

# Campaign Goals

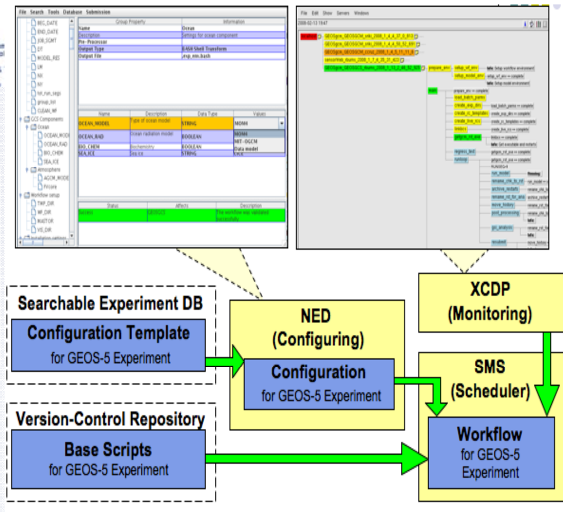
- Remove bottlenecks from climate model development life cycle.
- Evolve climate models using Darwinian natural selection processes.
- Build and maintain the climate modeling knowledge base.
- Broaden the base of climate modeling developer and user community.
- Advance science and science applications of satellite data assimilation and computational modeling for climate, weather, water and carbon cycles.

# Solution

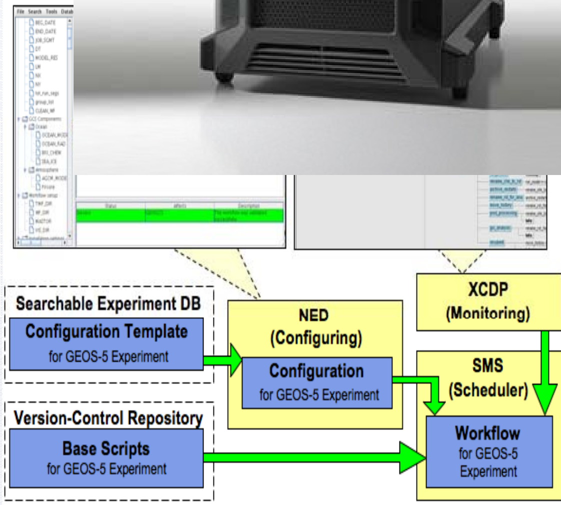
- Lower the bar for entry
  - Build and distribute low cost and turn key system packaged with HW, SW, data, scenarios, and productivity tools.
- Establish common baseline and boundary condition
  - Provide a development framework and a consistent climate modeling software architecture.
  - Provide startup models, configurations, and data analysis system
  - Establish standardized tests.
- Build a social network
  - Climate modelers
  - Model users
  - Data providers
  - Application users (water managers, energy & insurance sector, agriculture sector)
- Create incentives for modelers and users to participate and to volunteer the knowledge
- Create reward systems for long-term sustainment

# Lower the Bar for Entry – Model Developers

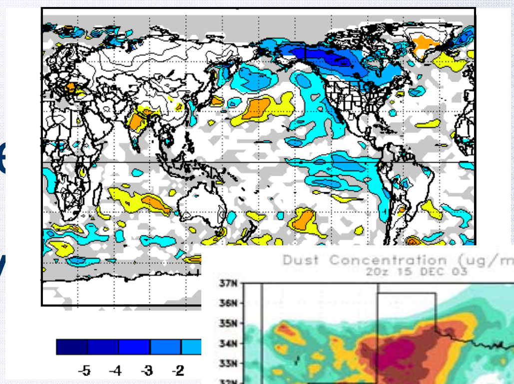

- Low cost computing platform (e.g. Cray CX1 w/ MS HPC 2008 or Redhat Linux)
- Atmosphere/Ocean models (Model-E, WRF, GEOS-5, CCSM, GFS)
- Earth System Modeling Framework & MAP Library (ESMF/MAPL)
  - Componentized architecture to reduce software engineering complexity
- Data (MERRA, SST, NCEP/NCAR reanalysis)
- Modeling Workflow
  - Model configuration, experiment design, and input/output data management
  - Tracking of experiments
  - Share experiment designs
- Development environment with compiler and debugger (e.g. eclipse, MS Visual Studio)
- Visualization Software (e.g. IDL, MatLab)
- Startup AMIP, CMIP, weather, and S/I runs
- MERRA scout run, ECMWF nature run
- Scenarios, OSSE, OSE

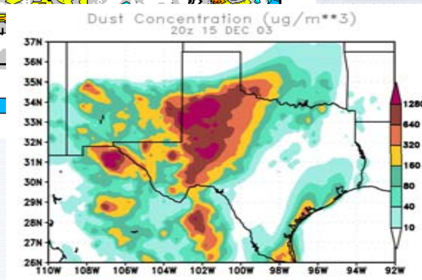



# Lower the Bar for Entry- Application Users




- Low cost computing platform (e.g. Cray CX1 w/ MS HPC 2008 or Redhat Linux)
- Pre-configured Model Interfaces
  - Global Climate to Regional to Local If-Then Applications
  - Chesapeake Bay Is Single Use Case (other examples: energy, agriculture insurance, transportation, etc.)
- Pre- Configured Modeling Workflow Hides Model Complexity
  - Interfaces Validated for Applications
  - Collaborative Hind-Casting Testing versus MERRA data
  - Share experiment designs
- Visualization Products via WMS/WCS/GIS Type Interfaces
- Remote Link to Large Scale Ensembles Runs on Large Scale Computing Facilities (e.g. Larger Numbers of CPUs needed for Ensemble Global Runs, Driving Cray WRF/Regional Models)
- Demonstration Project Taken to Applications Community to Identify Additional Specific Use Cases

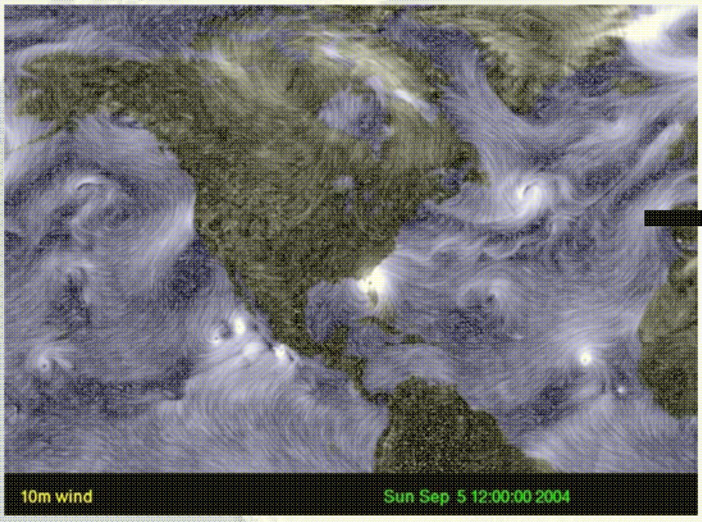




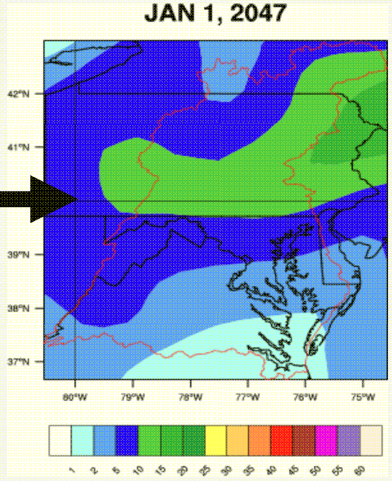
# Climate-In-A-Box: Application Users

Dynamic Downscaling: Scales That Matter to Decisions

GLOBAL MODEL




REGIONAL MODEL




REGIONAL DECISIONS

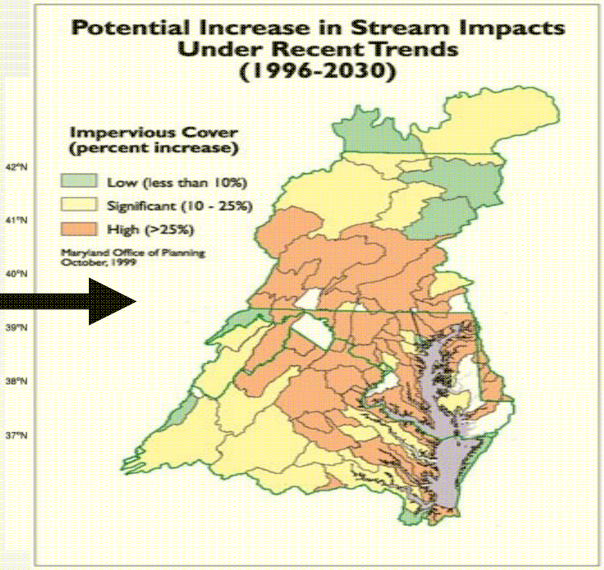
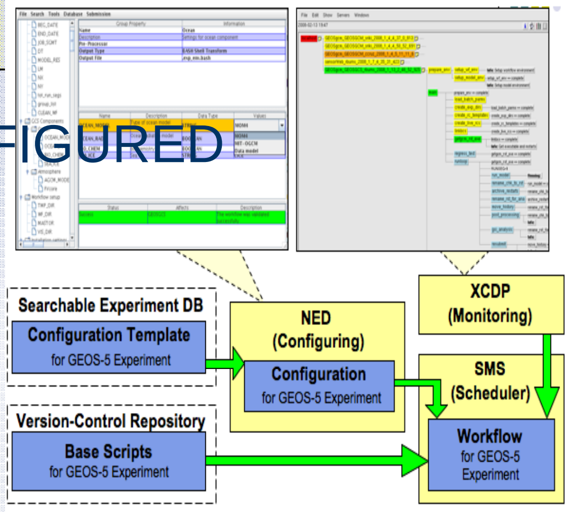



*Figure 5-8. Great stretches of the Cheapeake Bay watershed will likely see more areas covered by impervious surfaces—roads, highways, driveways, rooftops, and parking lots. The areas most acutely affected (see map) will experience increases of 25 percent or more in impervious cover, if recent trends persist.*

PRE-CONFIGURED




# Cray CX1 as an Example

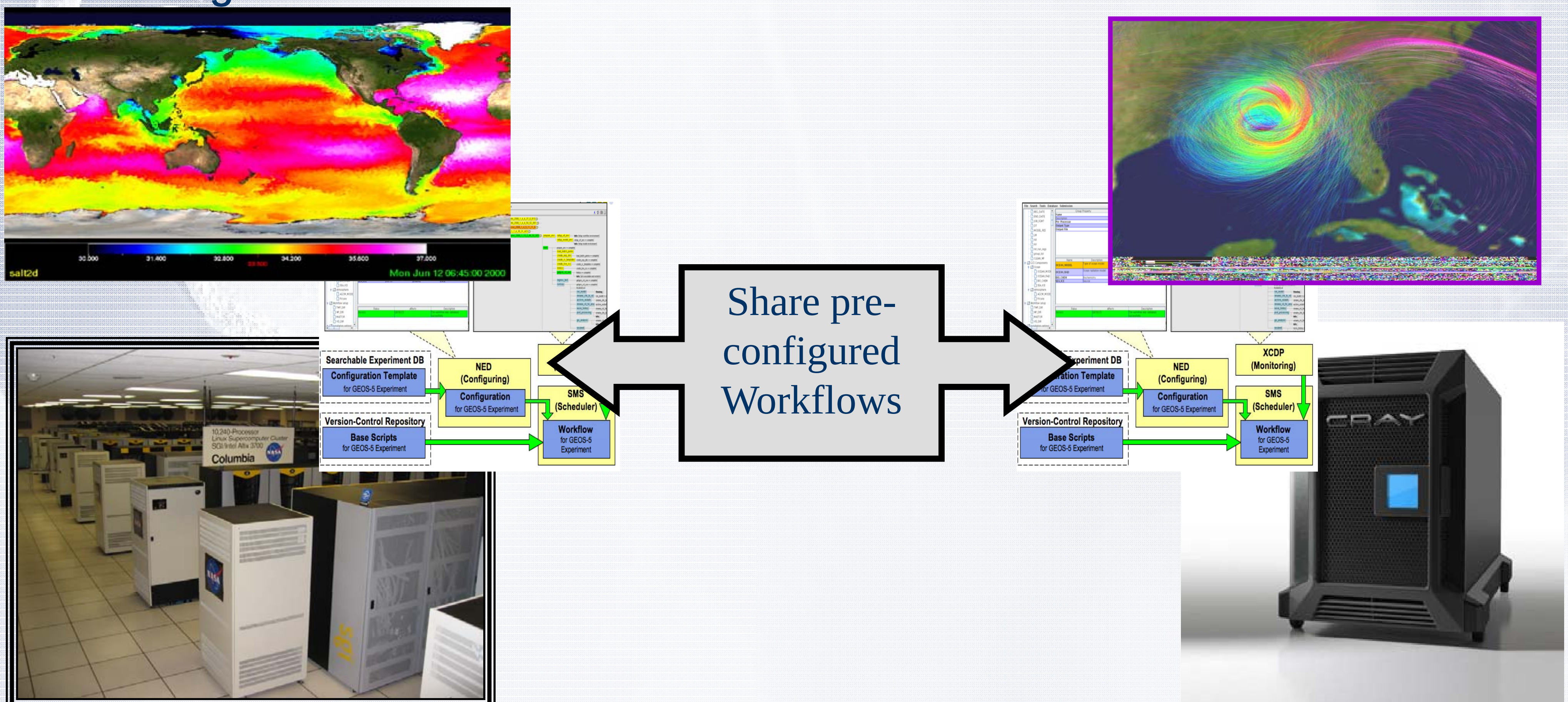
- Personal “turn-key” supercomputer
- Plug to the wall - No additional power and cooling required
- Starting $25K
  - 4 socket, 16 compute cores
- Fully populated at $90K
  - Up to 8 compute nodes
  - Up to 64 compute cores
  - 16 gigabytes of memory
  - 4 terabytes of disks
- ¼ x ¼ degree global atmosphere model run for hurricane forecast will fit in this machine
  - 5 day hurricane forecast may be done in two hours


http://www.cray.com/products/CX1.aspx

# Advanced Concept

*Collaborating Computing via "Sharing Workflow"*

- Shift from Local Model Runs to Larger Facility (Cloud or Grid computing) to Obtain Greater Numbers of CPUs, or
- Collaborate among Ensemble, or
- Use Global Model Output to Drive Regional Models
- Using Pre-Configured "Workflow Sharing" Support Global and Regional Modeling

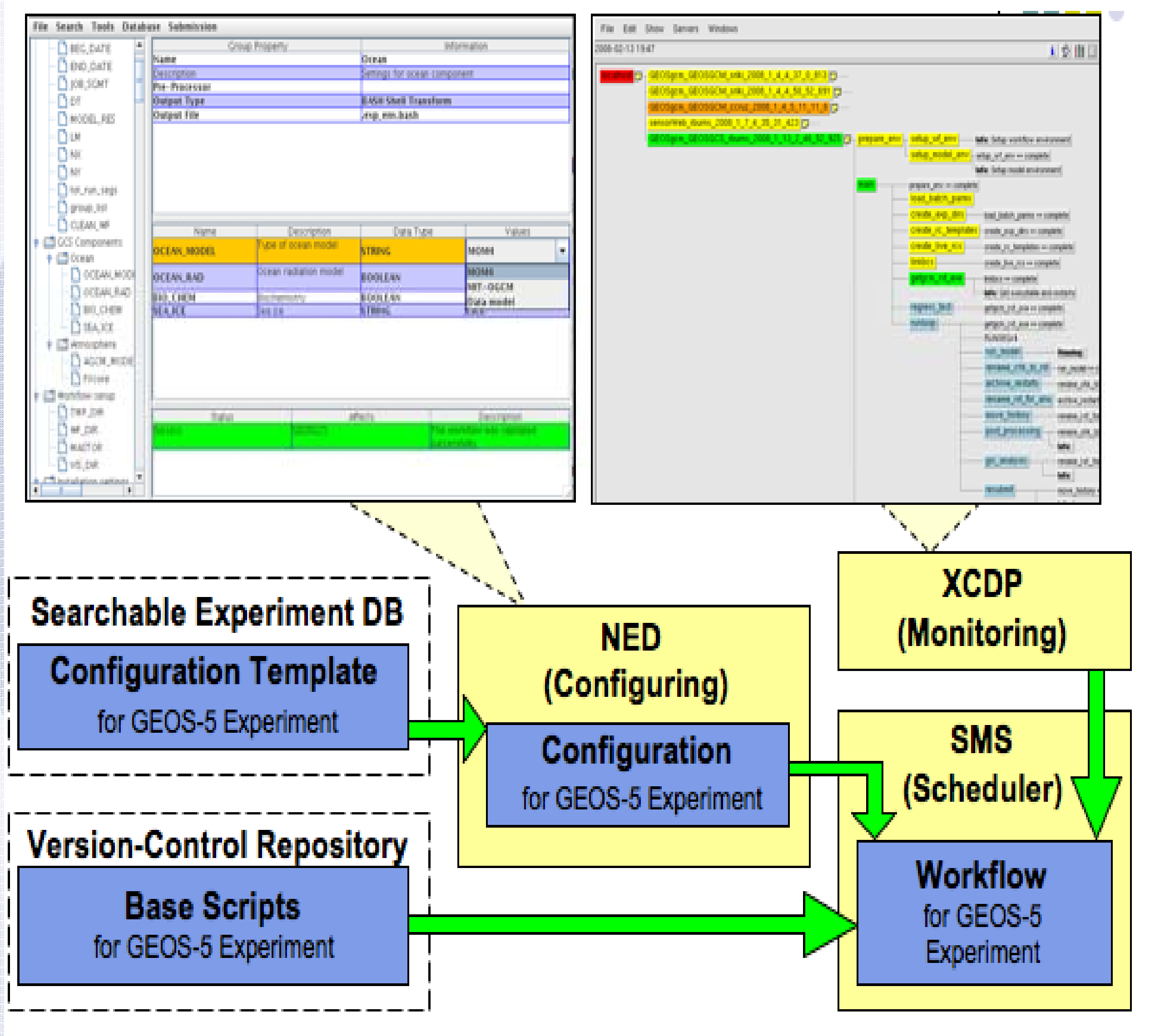


# Modeling Guru Social Network

- Web 2.0 based modeler's social networking site
- Knowledge management tool
- Exchange of model components and blog about modeling experiences
- Ranking by natural selection

http://modelingguru.nasa.gov

# Workflow & NASA Experiment Designer (NED)




- Model configuration, experiment design, and input/output data management
- Tracking of experiments
- Tracking and maintaining of I/O data
- Version control
- Repeatable experiments
- Sharing experiment designs

# Summary

- NASA Earth System Science continues to integrate models and observations to answer societal challenges.
- Climate in a Box is a toolbox for model developers, climate information, climate application, and decision support users.

Thank you!