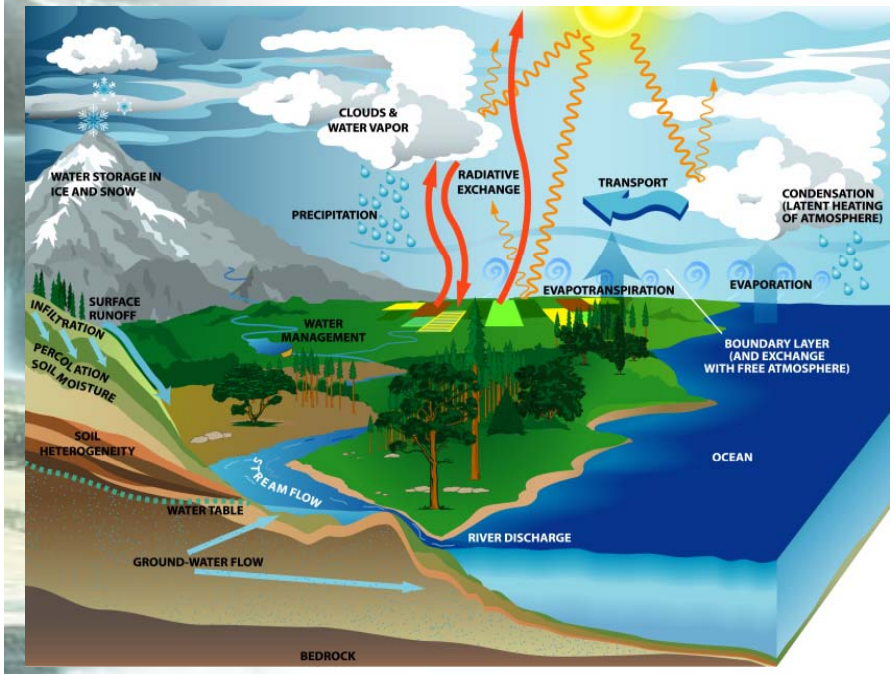
Water Cycle Data Integration, Assimilation & Utilization

Paul R. Houser (GMU & CREW)

The Water and Energy Cycle



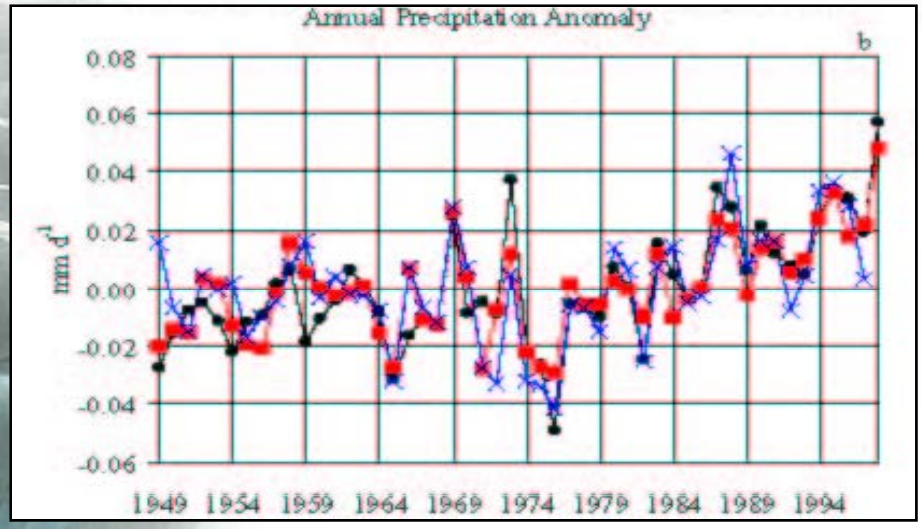
Water in the climate system functions on <u>all</u> time scales: From hours to centuries

Why study the water & Energy cycle?

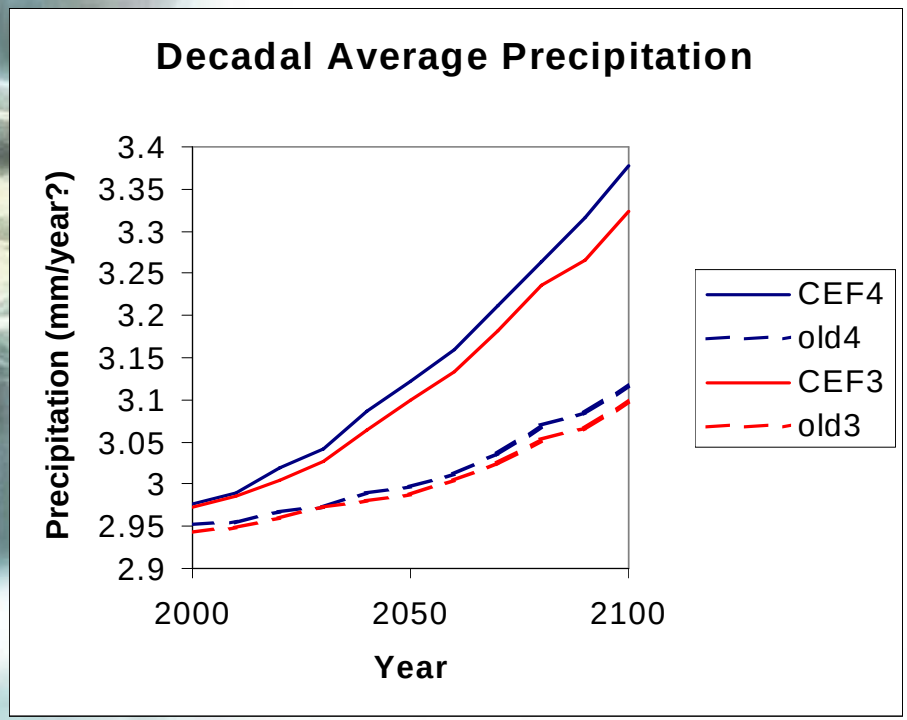
1. *Water exists in all three phases in the climate system and the phase transitions are a significant factor in the regulation of the global and regional energy balances*
2. *Water vapor in the atmosphere is the principal greenhouse gas and clouds at various levels and composition in the atmosphere represent both positive and negative feedback in climate system response*
3. *Water is the ultimate solvent and global biogeochemical and element cycles are mediated by the dynamics of the water cycle*
4. *Water is the element of the Earth system that most directly impacts and constraint human society and its well-being.*

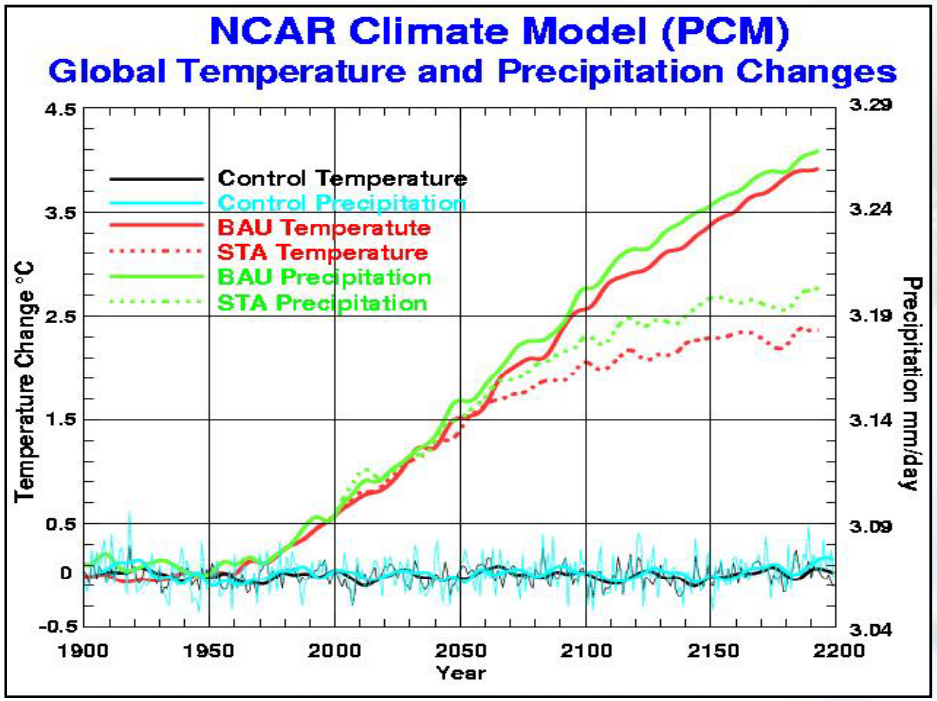
NASA ENERGY AND WATER CYCLE STUDY
NEWS

Global Water Studies: Motivation and Methodology



Obs variability: ~0.05 mm/d
Model variability ~0.005 mm/d
Model 100yr trend: ~0.1 mm/d





Motivation

Integration of the water cycle is expected to offer benefits to the local environment, community and economy (Schneider et al., 2006)

Following CCSP guidelines, *an interagency "integrating" priority project* will be implemented (contingent on funds) under the water cycle element over the next decade. The purpose of this project is to address significant uncertainties associated with the water cycle through a study that comprehensively addresses the water budget.

It is understood that WCRP and GEWEX goals require *effective data integration and information fusion* with minimum time delay and at minimum cost is central to achieving climate research goals (CEOP).

What is integration?

$\int_a^b f(x)\,dx$ The $\int$ sign represents **integration**, *a* and *b* are the endpoints of the interval, f(x) is the function we are integrating known as the integrand, and dx is a notation for the variable of **integration**.

Antonym: **disintegration** transformation from a more to a less massive configuration.

- Integration means different things to different people and as a consequence appears to only partially deliver on promised outcomes. The objective of water cycle integration is to provide better outcomes than would be expected by managing the parts independently.
- *Data Integrator* automates data extraction, transformation, and movement from diverse data sources. It's enterprise-class data integration that's easy to use and scale.
- *Informatica* provides comprehensive data integration solutions for today's complex enterprise environments, enabling them to access, discover, cleanse, integrate and deliver their data - where needed and when needed - so they can realize its full value.

Begin with discovery; end with enlightenment.

- Data integration has been recently recognized as a research topic of huge practical importance. The availability of integrated data from multiple independent, heterogenous data sources is crucial for many applications. Data integration requires combining and matching information in different sources, and resolving a variety of conflicts.

Water Cycle Integration – How to achieve the desired integrated products?

What is integration?

- Integrating observations to establish a more complete system description
- Integrating model components to build a earth modeling system
- Integrating research results to establish end-user solutions

Data Integration: Spatial and temporal rectification to allow intercomparison and quality evaluation of disparate model and observation data;
Data-Model Integration: Physical rectification or constraint of data and its error using four dimensional data assimilation and modeling techniques.
Model Integration: Using component models to build a system model.
Solution Integration: Integrating components (research results) to develop solutions
Interpersonal Integration: Interconnection of disparate water cycle research teams.

Science Integration:

- Data integration
- Coordinate energy and water process modeling
- Water & energy cycle trend and variability assessments
- End-user decision support & solution network connections

Water Cycle Integration – How to achieve the desired integrated products?

Goal : provide a synthesis of available global water cycle information in a form that is accurate, unbiased, policy-relevant and global-scale to a broad range of stakeholders.

Tools: prediction models, Geographic Information Systems (GIS), data assimilation systems, observational simulation experiments and data mining.

Products: merged products, water balance information, process understanding, observationally-based model diagnostics, improved model performance, improved decision support.

Define three aspects of data integration:

1. *Spatial and temporal rectification:* Simply interpret relevant data to a common time and space domain for intercomparison and visualization. This will enable assessment of data set error and bias, variability, uncertainty and predictability. Water-balance analyses will further interrelate the data, highlighting gaps in our knowledge.
2. *Physical rectification or constraint:* Data assimilation techniques merge a range of diverse data fields with a model prediction to provide the best estimate of the current state of the natural environment. This is an integration process that uses many data sources, and resolutions.
3. *Communication:* Because of the highly specialized nature of earth science research, it is common for disparate research groups and stakeholders to be unaware of each other's activities. Therefore, enabling the sharing of data, ideas and techniques can be an important integrating function.

State of the Water & Energy Cycle

Evaluate the research community's current ability to detect, analyze, understand and explain global water cycle change, variability, prediction and predictability.

Water and Energy Cycle Data Integration

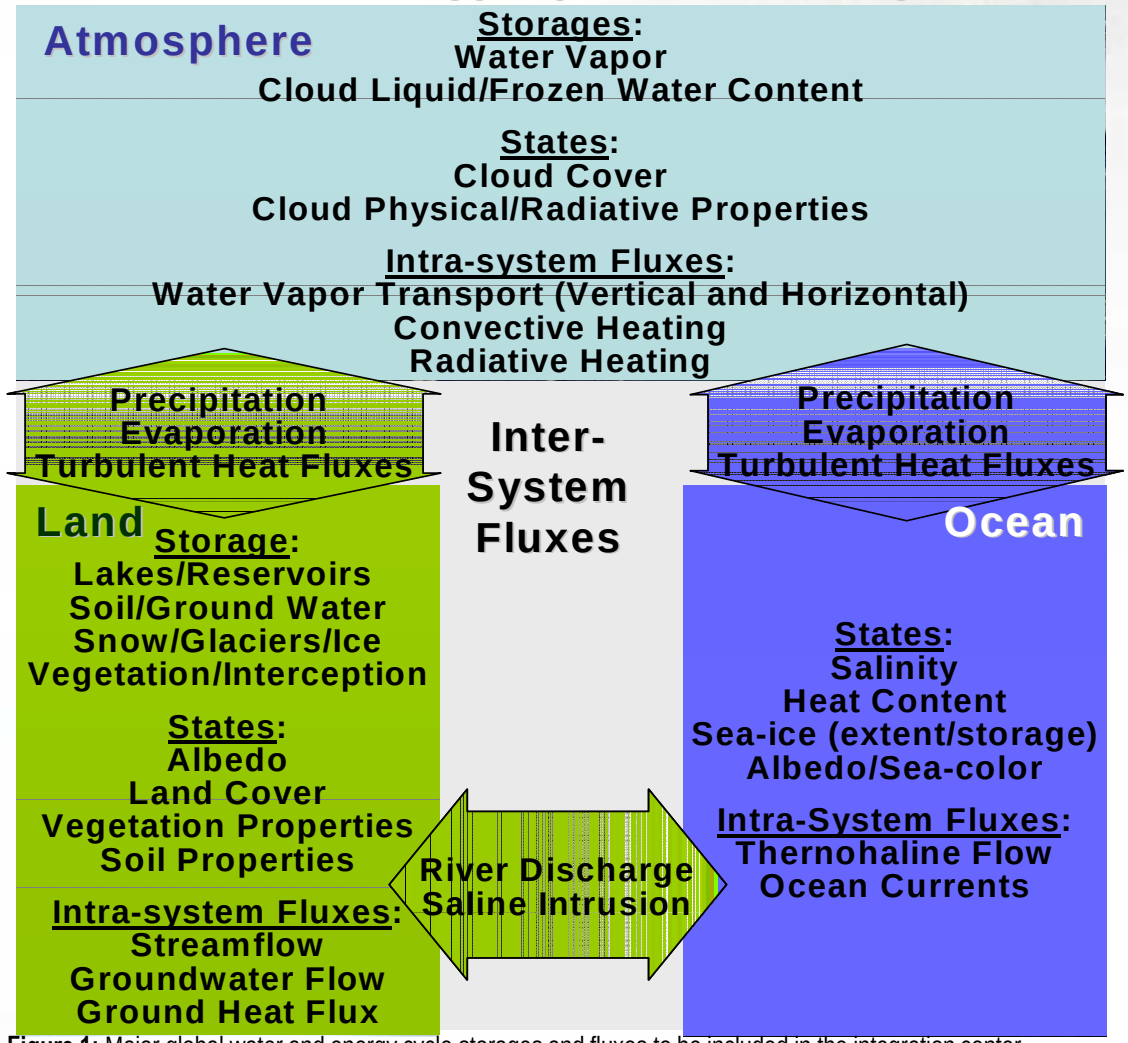


Figure 1: Major global water and energy cycle storages and fluxes to be included in the integration center.

State of the Water and Energy Cycle

Variable ↓ / Sphere →	Ocean	Terrestrial	Atmosphere
Internal or State Variable	upper ocean currents (I/S) sea surface temperature (I/S) sea level/surface topography (I/S) sea surface salinity (I/S) sea ice (I/S) wave characteristics (I/S) mid- and deep-ocean currents (I) subsurface thermal structure (I) subsurface salinity structure (I) ocean biomass/phytoplankton (I/S) subsurface carbon(I), nutrients(I) subsurface chemical tracers(I)	topography/elevation (I/S) land cover (I/S) leaf area index (I) soil moisture/wetness (I/S) soil structure/type (I/S) permafrost (I) vegetation/biomass vigor (I/S) water runoff (I/S) surface temperature (I/S) snow/ice cover (I/S) subsurface temperature (I/S) subsurface moisture (I/S) soil carbon, nitrogen, phosphorus, nutrients (I)	wind (I/S) upper air temperature (I/S) surface air temperature (I/S) sea level pressure (I) upper air water vapor (I/S) surface air humidity (I/S) precipitation (I/S) clouds (I/S) liquid water content (I/S)
Forcing or Feedback Variable	ocean surface wind & stress (I/S) incoming SW radiation (I/S) incoming LW radiation (I/S) surface air temperature (I/S) surface air humidity (I/S) precipitation (I/S) evaporation (I/S) fresh water flux (I/S) air-sea CO_2 flux (I) geothermal heat flux (I) organic & inorganic effluents (I/S) biomass and standing stock (I/S) biodiversity (I) human impacts-fishing (I)	incoming SW radiation (I/S) incoming LW radiation (I/S) PAR radiation surface winds (I) surface air temperature (I/S) surface humidity (I/S) albedo (I/S) evapotranspiration (I/S) precipitation (I/S) land use (I/S) deforestation (I/S) land degradation (I/S) sediment transport (I/S) air-land CO_2 flux (I)	sea surface temperature (II/S) surface soil moisture (I/S) surface soil temperature (I/S) surface topography (I/S) land surface vegetation (I/S) CO_2 & other greenhouse gases, ozone & chemistry, aerosols (I/S) evapotranspiration (I/S) snow/ice cover (I/S) SW and LW surface radiation budget (I/S) solar irradiance (S)

BLUE=Water Cycle Variable; RED=Energy Cycle Variable; GREEN=Carbon/Chemistry Variable; BLACK=Boundary condition

$$\frac{d\langle Q\rangle}{dt} = \langle E\rangle - \langle P\rangle$$

$$R = P - E \pm \Delta G$$

$$P_o = E_o - D_o + D_l = E_o - R$$
$$P_l = E_l + D_o - D_l = E_l + R$$

$$\frac{\partial S}{\partial t} = -\nabla_H \cdot \vec{R}_o - (E - P).$$

$$P + R_o + \Delta O + G_{do} = E$$

NASA ENERGY AND WATER CYCLE STUDY NEWS

Water & Energy Cycle Data Integration

Water Cycle Data Integration: compile, diagnose and disseminate water and related energy cycle observations and predictions.

1: Enable the quantification of global water cycling rates, which requires global data integration for vertical water fluxes, land water storages, and lateral land water fluxes.

2: WEC variability (extremes) studies.

3: Diagnostic trend studies, transient variability and predictability; model validation; and initialization.

Key components of the WEC data integration include:

- **Provide a "one stop", streamlined access to coordinated, geolocated, and integrated water & energy cycle data and visualizations from all sources.**
- **Identify and acquire global water and energy cycle observations and model predictions from all relevant sources, over the longest available period.**
- **Provide a framework for sharing IT resources.**
- **Establish the products in consistent formats and access protocols.**
- **Develop an integrated re-processing plan.**
- **Assess the physical consistency of products.**
- **Establish a data management strategy to organize and link to local and distributed data resources. Whenever possible, virtual and meta data links will be established to existing data archives, rather than explicitly downloading them.**

State of the Water & Energy Cycle

Evaluate our ability to detect, analyze, and understand global water cycle change, variability, prediction and predictability.

Water and Energy Cycle Data Integration

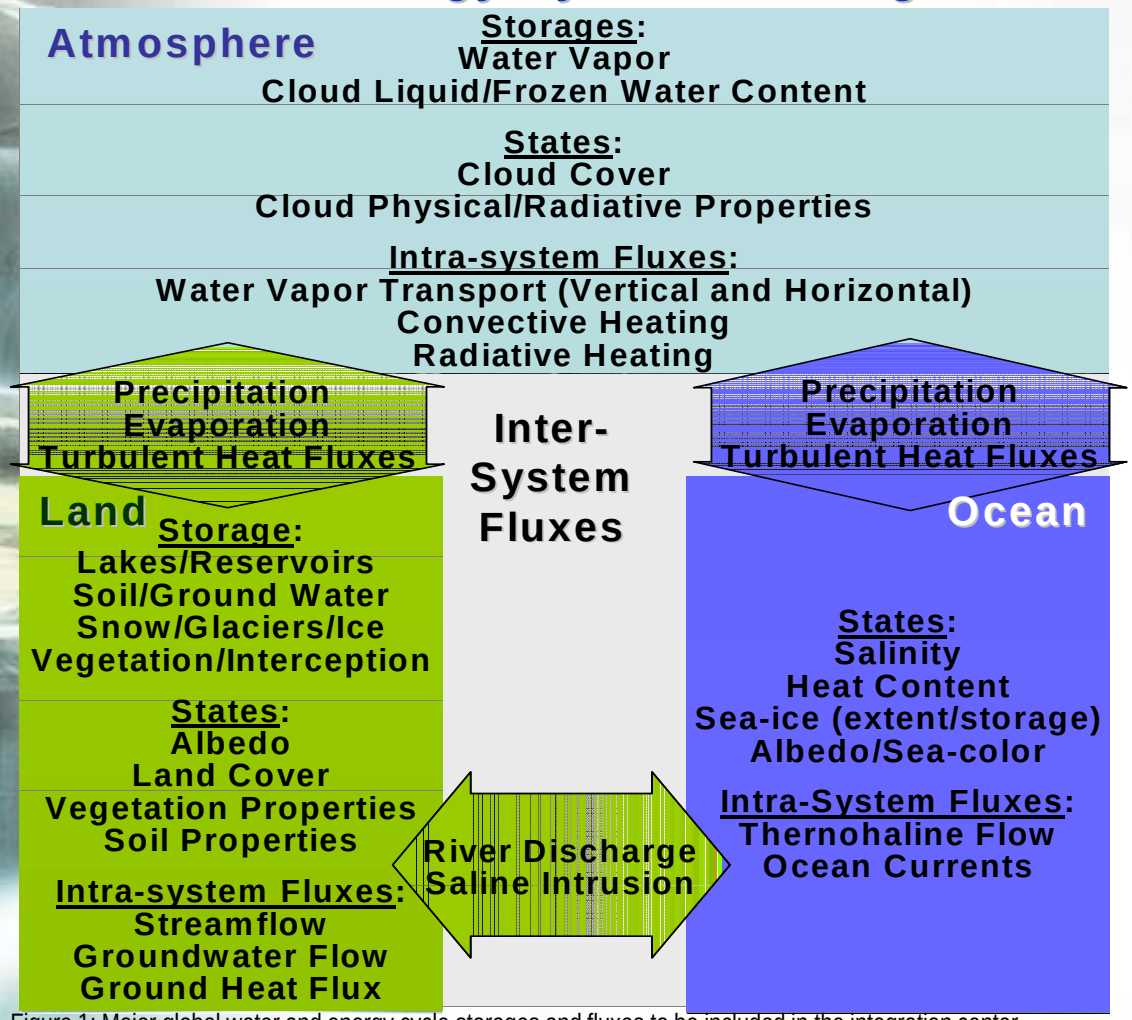


Figure 1: Major global water and energy cycle storages and fluxes to be included in the integration center.

Global Satellite Water Balance Study:

Schlosser & Houser, 2007

$$\frac{d\langle Q\rangle}{dt} = \langle E\rangle - \langle P\rangle$$

Precipitation (1979-1999):

- Global Precipitation Climatology Project (GPCP): Adler et al., (2003)
- CPC Merged Analysis of Precipitation (CMAP): Xie and Arkin (1997)

Ocean Evaporation (1987-1999):

- Goddard Satellite-based Surface Turbulent Fluxes Version 2 (GSSTF2): Chou et al., (2003)
- Hamburg Ocean Atmosphere Parameters and Fluxes from Satellites: Bentamy et al. (2003)

Land Evaporation:

- Global Offline Land Dataset (GOLD) (1959-2002): Dirmeyer et al., (2005):
- *Global Soil Wetness Project Phase 2 (GSWP2): 1986-1995*

Precipitable Water: NASA Global Water Vapor Project (NVAP)

Model Output: Climate of 20th Century

Annual Mean Statistics (1988-1999)

NEWS (Schlosser&Houser)

Units in kg/yr	Precipitation	Evaporation	| P-E |
Land	1.05E+17 ± 0.02E+17	GOLD1: 0.64E+17	~4.0E+16
	1.02E+17 ± 0.02E+17	GOLD2: 0.62E+17	~4.2E+16
Ocean	3.80E+17 ± 0.06E+17	4.41E+17	6.5E+16
	3.72E+17 ± 0.04E+17	3.93E+17	1.7E+16
Global	GPCP 4.85E+17 ± 0.06E+17	GSSTF2+GOLD 5.03E+17	~ 2.4E+16
	CMAP 4.74E+17 ± 0.04E+17	HOAPS+GOLD 4.56E+17	

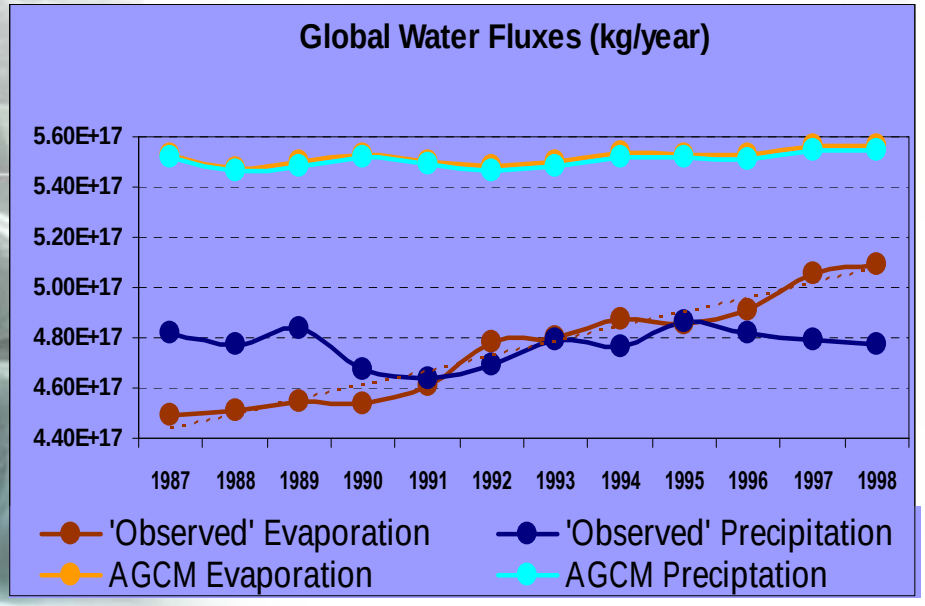
Note: Total atmospheric water storage ~ 10^{16} kg, annual change ~10^{14} kg

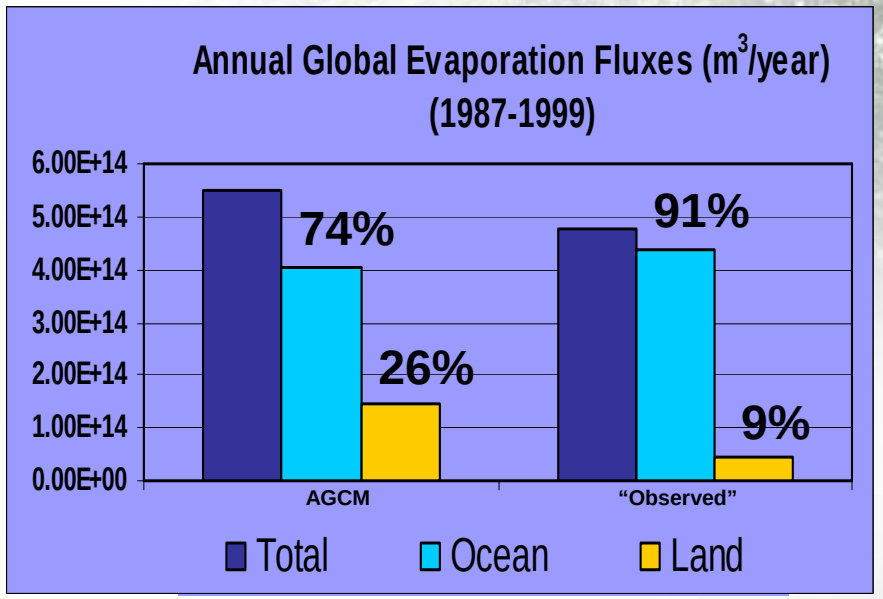
Adapted from Schlosser and Houser (2006,)

- Global annual mean precipitation and evaporation balance to ~5% or 24,000 (metric) gigatons of water.
- Imbalance exceeds global estimate of annual precipitation error.

NASA ENERGY AND WATER CYCLE STUDY NEWS

AGCM Precipitation and Evaporation Evaluation





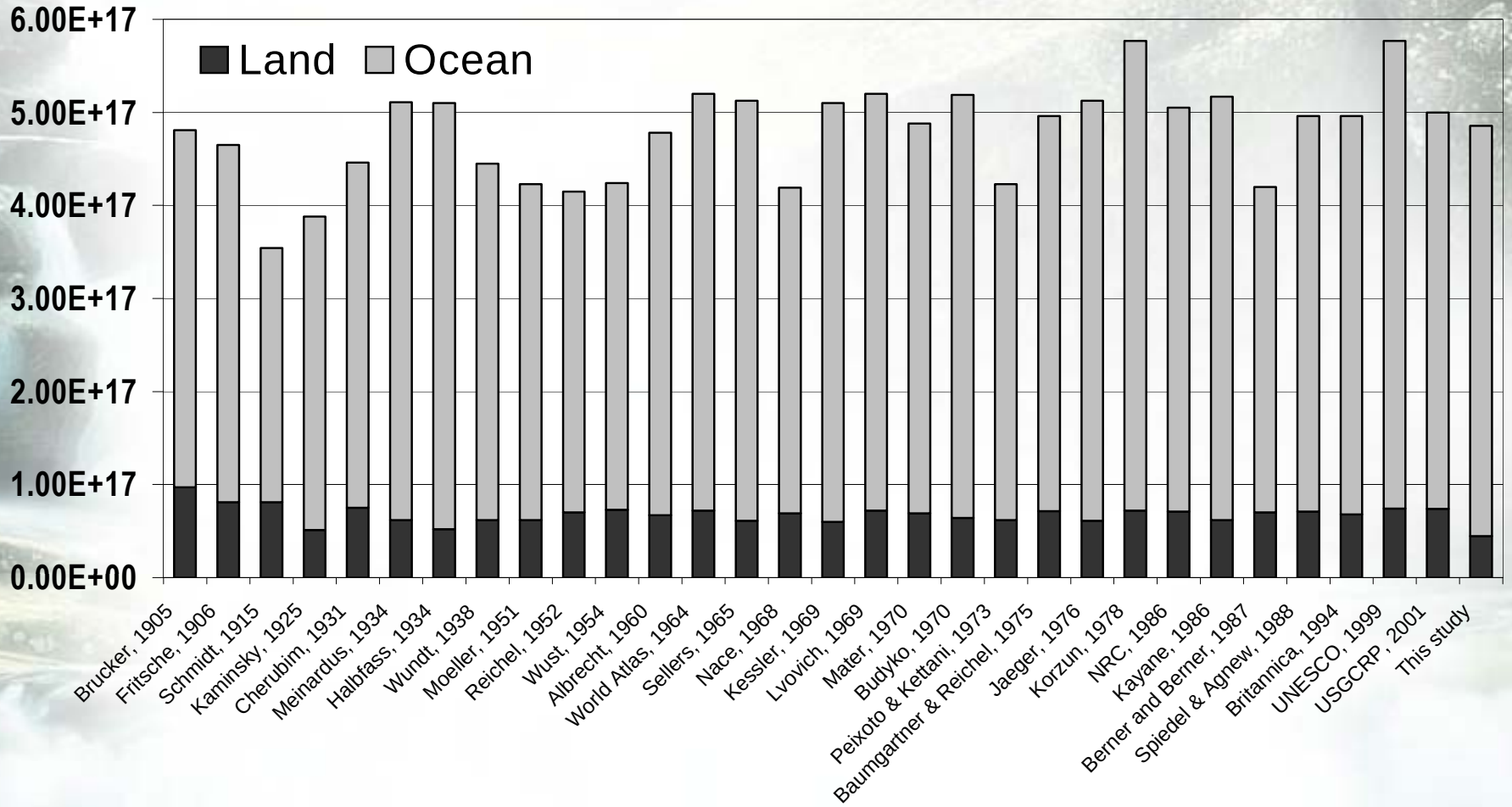
Schlosser

- **Observed averaged annual evaporation and precipitation mass flux balance to within 1%.**
  - **However, interannual global variations considerably uncorrelated.**

- **AGCM mean "rate" of annual global water cycle exceeds observed (~15%).**
- **AGCM interannual variability of annual global precip/evap ~50%/35% lower than observed.**
- **Relative contributions of land and ocean fluxes differ considerably.**
  - **What are the sources of these discrepancies (both in the models and "observations")?**

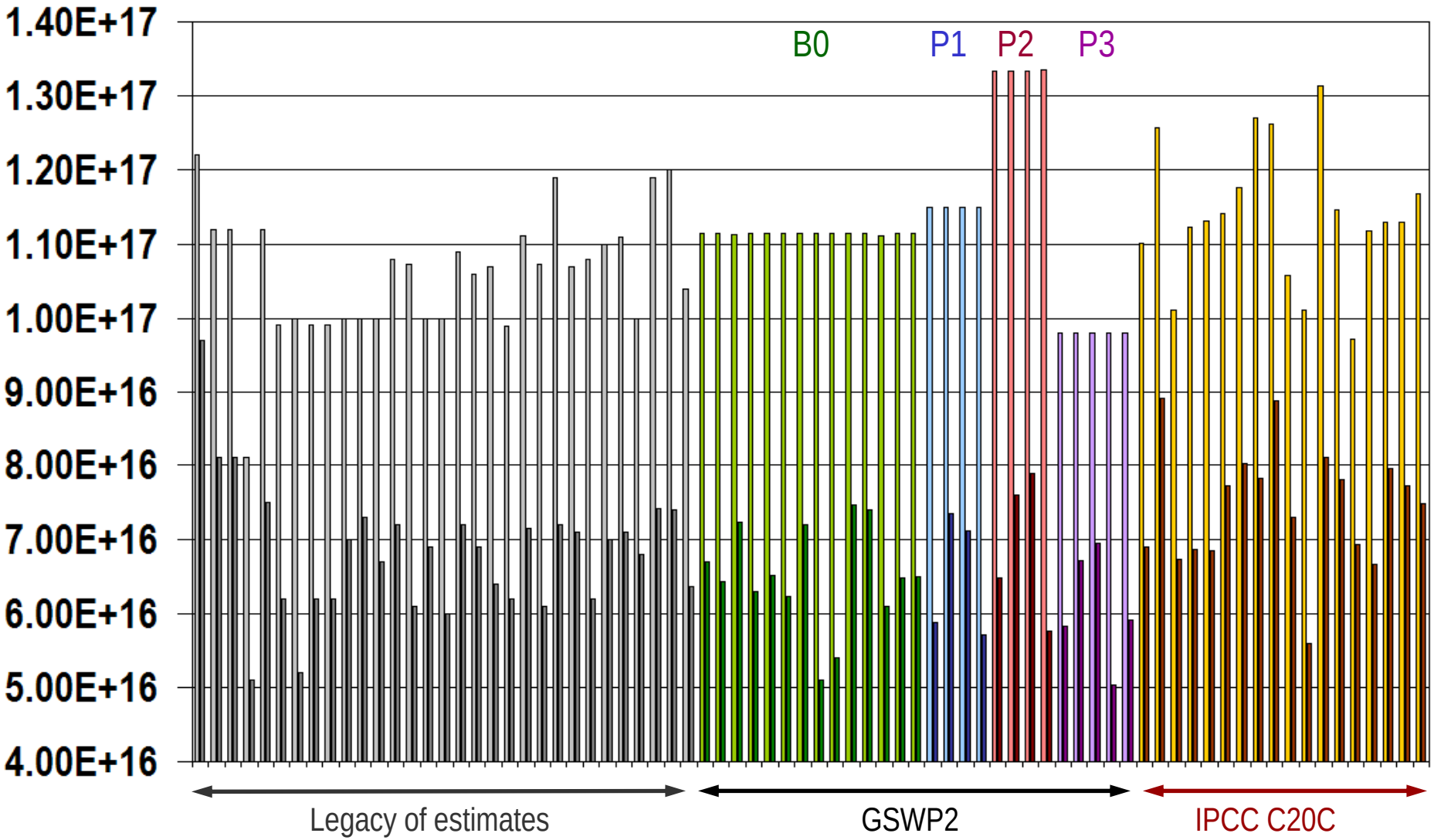
- **Trend in "observed" global evaporation (~1 %/year), but no trend in precipitation.**
- **Trend in AGCM global water-cycle rate during 1987-1999 and order of magnitude smaller.**
  - **Source of modeled trend from prescribed SSTs, is the response accurate?**
  - **Observations insufficient to detect AGCM trend (e.g. Ziegler et al., 2002).**

Comparison of Global Evaporation Fluxes to Previous Estimates



- Global fluxes of precipitation and evaporation are comparable to previous century of estimates.
- No discernable trend is seen in both compilations of the flux estimates.
- The notable disparity with this study is the lower values of both precipitation (not shown) and evaporation flux estimates over land.

Mean Annual Global Land Precipitation and Evaporation (kg/yr)

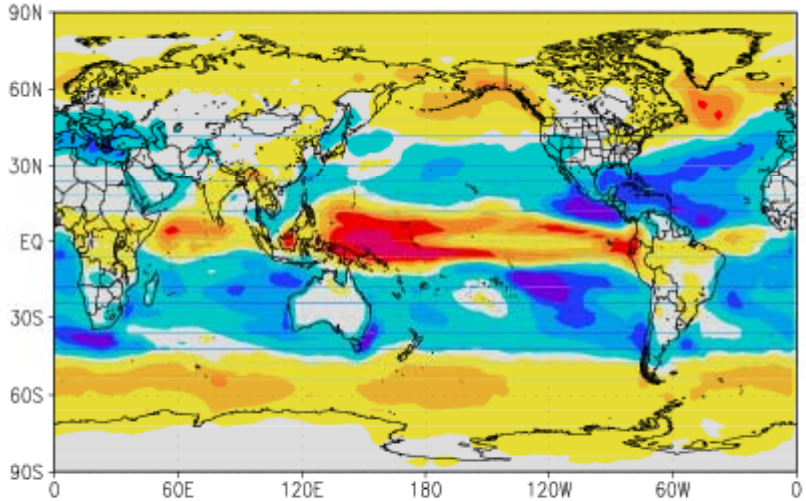


- Model-based (offline and coupled) scatter of estimates marginally higher than compilation of “modern” observationally-based estimates.

Multi-model ensemble mean change from IPCC GCMs

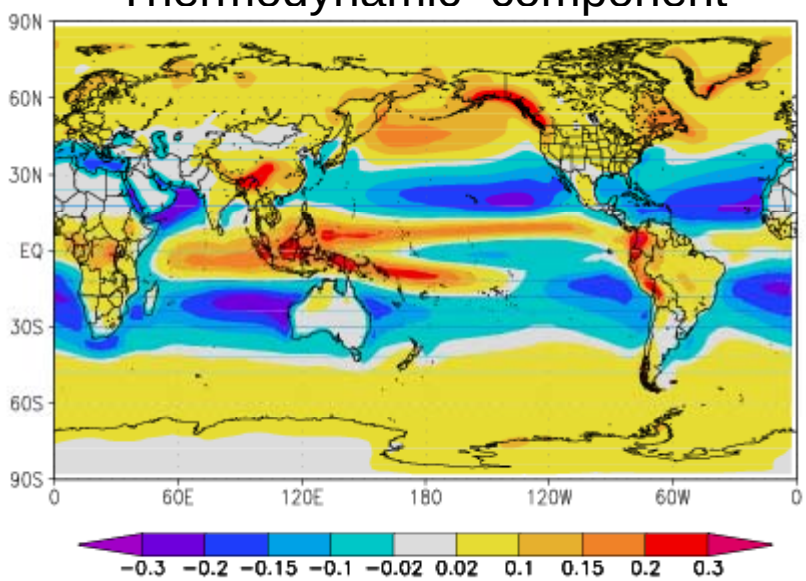
Change in (P-E) for 2100 minus 2000

'Dry regions get drier, wet regions get wetter"

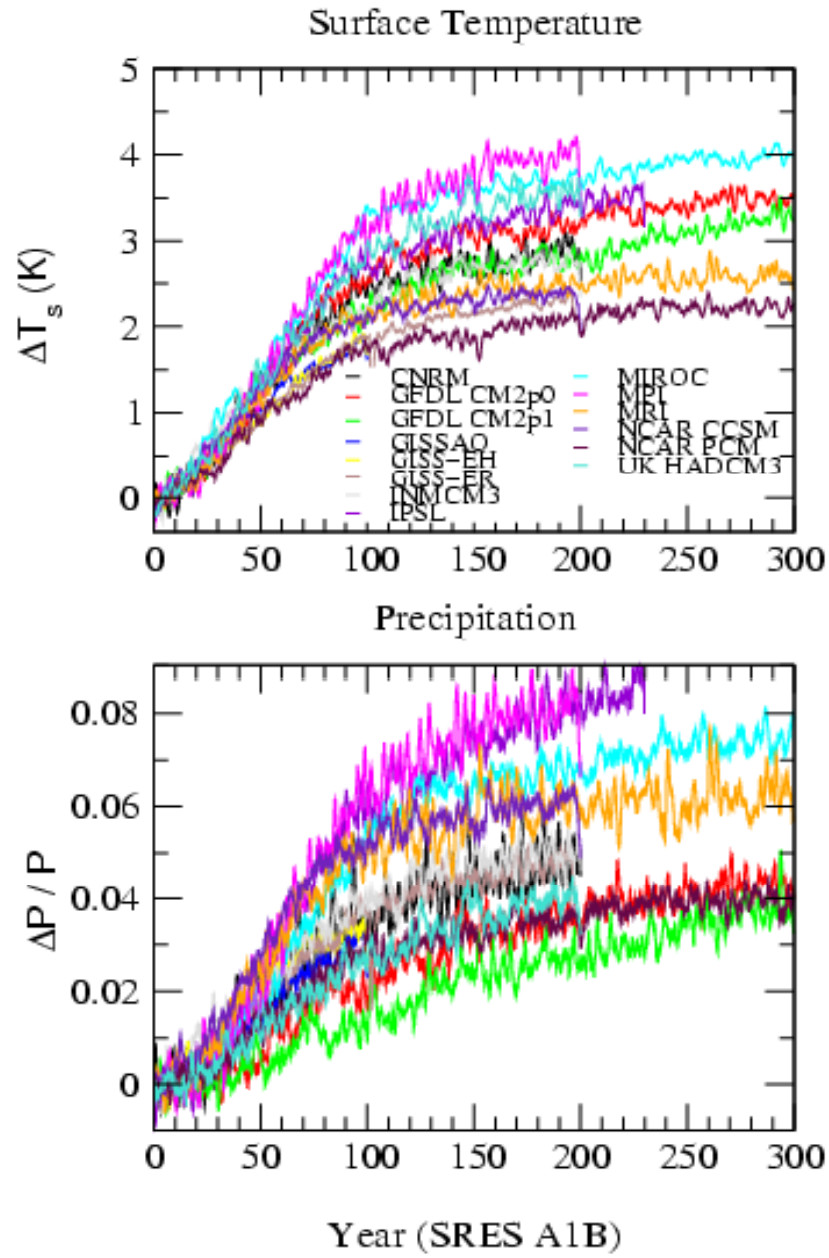


Held and Soden (2006)

"Thermodynamic" component



Δ(P-E) mm/day

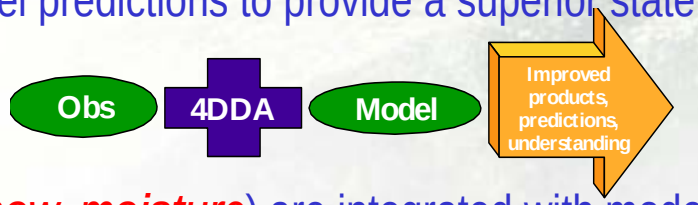


Vecchi and Soden (2007)

Data Assimilation

Data Assimilation merges observations & model predictions to provide a superior state estimate.

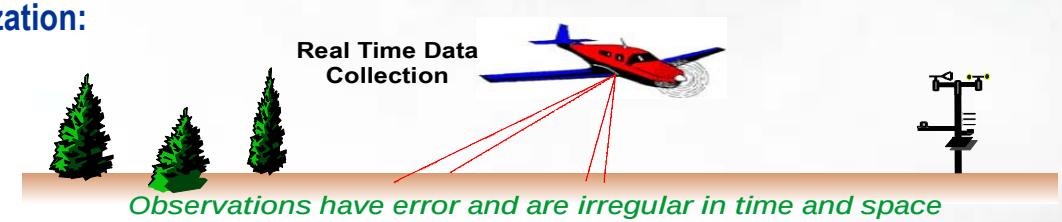
$$\partial x / \partial t = dynamics + physics + \Delta x$$

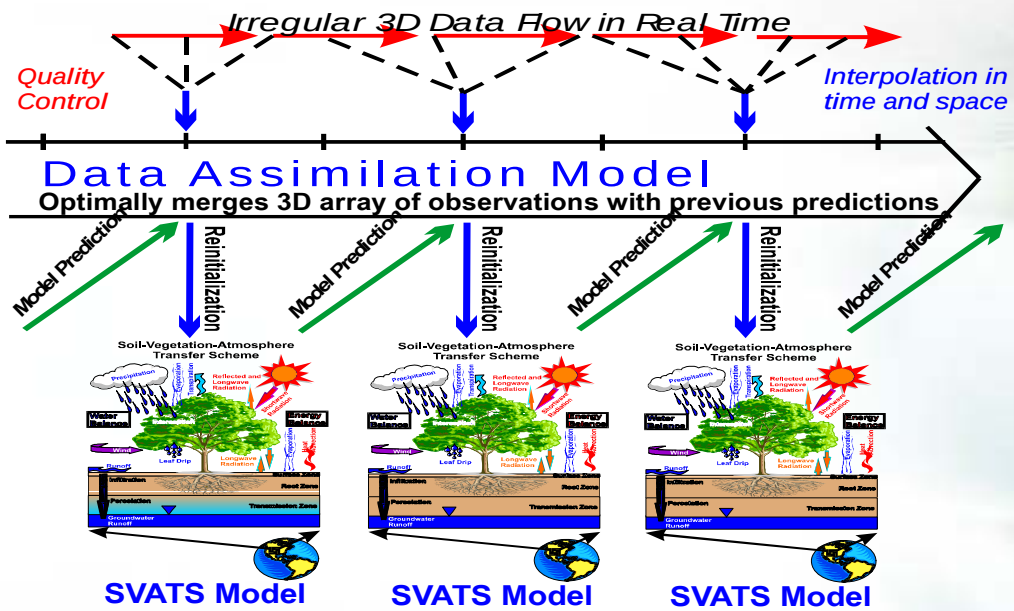


State or storage observations (*temperature, snow, moisture*) are integrated with models.

Data Assimilation Methods: Numerical tools to combine disparate information.

1. **Direct Insertion, Updating, or Dynamic Initialization:**
2. **Newtonian Nudging:**
3. **Optimal or Statistical Interpolation:**
4. **Kalman Filtering: EKF & EnKF**
5. **Variational Approaches - Adjoint:**





Model errors result from:

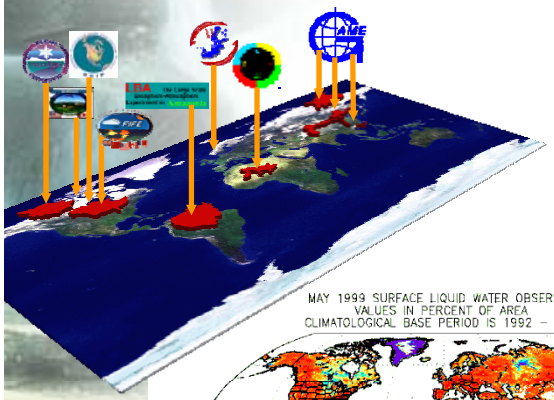
- Initialization error.
- Errors in atmospheric forcing data.
- Errors in LSM physics (model not perfect).
- Errors in representation (sub-grid processes).
- Errors in parameters (soil and vegetation).

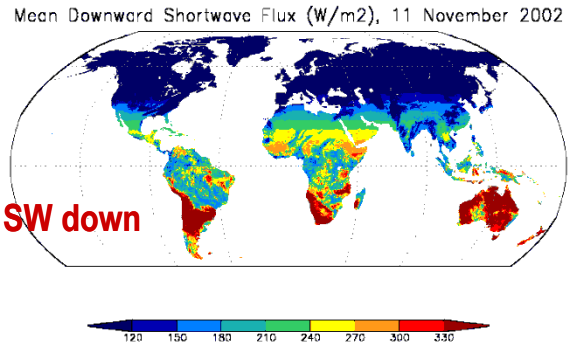
NASA ENERGY AND WATER CYCLE STUDY

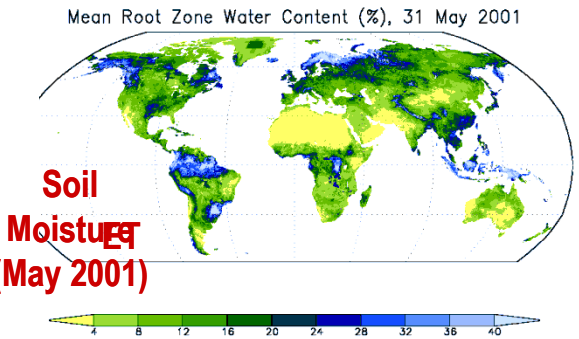
Global Land Data Assimilation System

Objective: A 1global land modeling and assimilation system that uses most relevant observed forcing, storages, and validation.

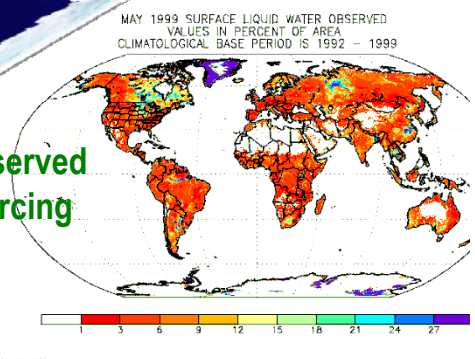
Consistent Global Intercomparison

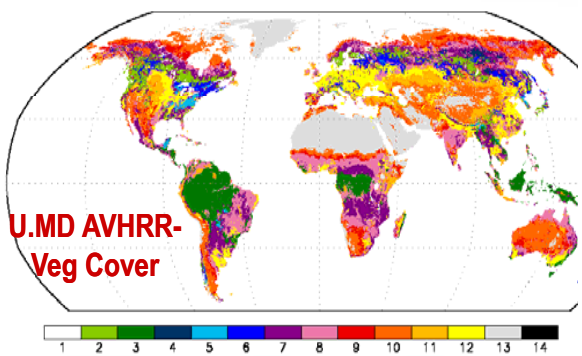


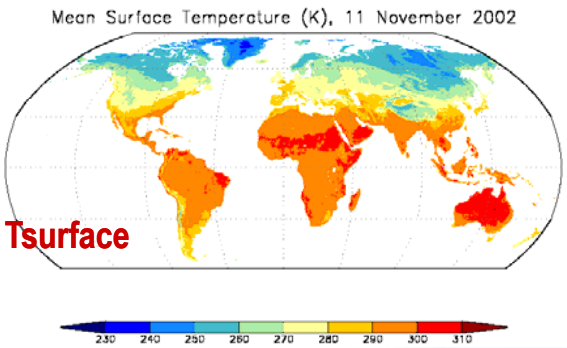




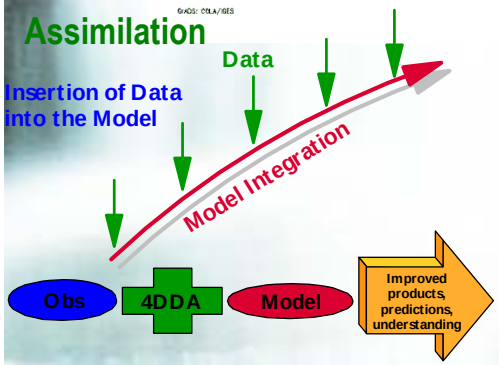
Observed Forcing

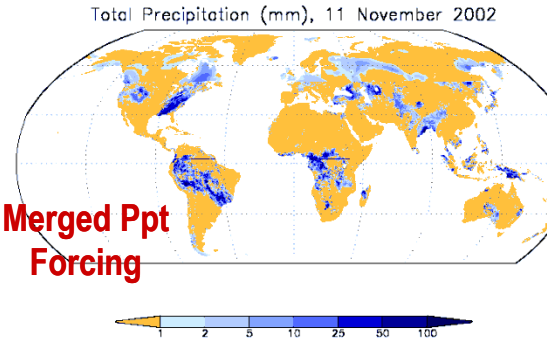


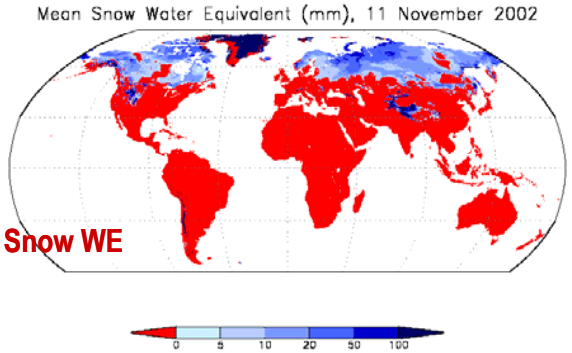




Land Data Assimilation

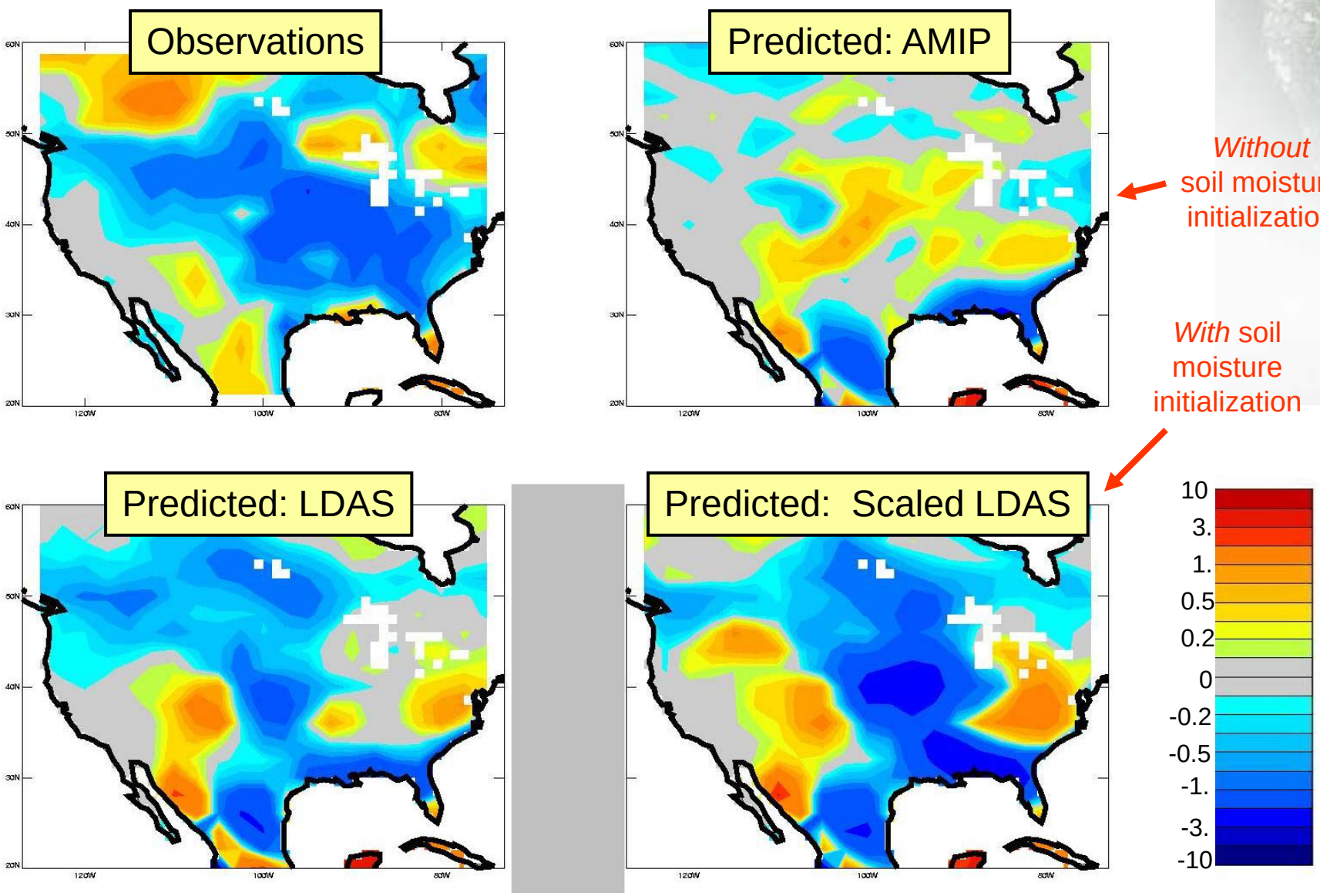






Coupled Model Forecast: 1988 Midwestern U.S. Drought

(JJA precipitation anomalies, in mm/day)



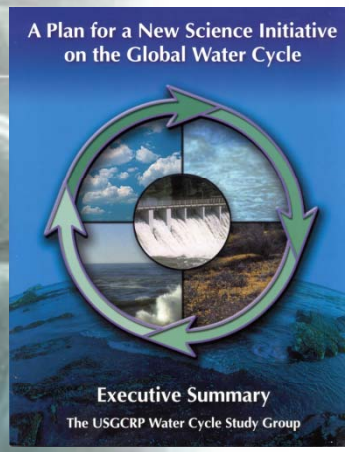
Without soil moisture initialization

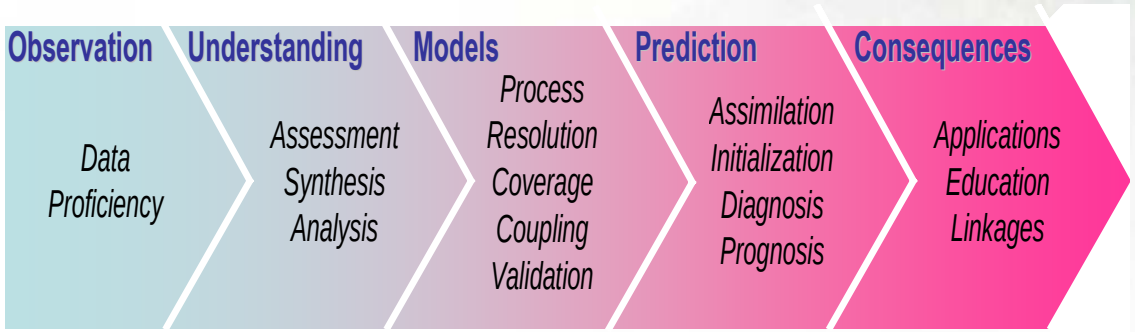
With soil moisture initialization

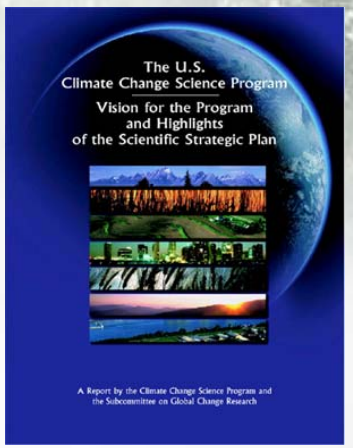
Koster et al., 2004

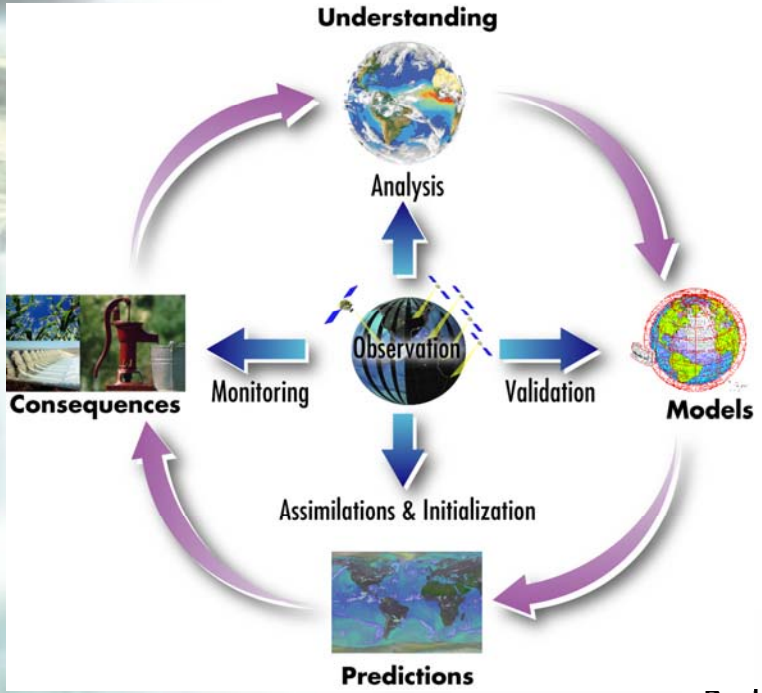
Program Integration

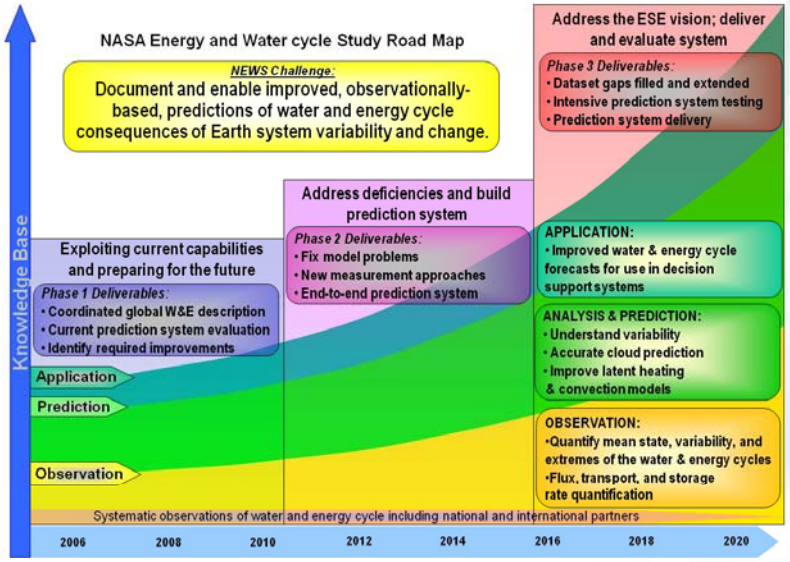
- Establish programs that encourage integration











NASA ENERGY AND WATER CYCLE STUDY NEWS

Water Cycle Science Synthesis

Synthesis products show the collective progress we are making towards our goals.

Candidate synthesis products:

- *Quantify the potential predictability of land and ocean states on seasonal precipitation.*
- *Quantify the sensitivity of seasonal climate predictions to land states.*
- *Quantify seasonal climate forecast skill & accuracy requirements for water applications.*
- *Develop operational hydrologic forecasts incorporating the use of climate forecasts.*
- *Develop and useful seasonal hydrologic forecasts systems for water applications.*
- Determine how global precipitation, evaporation and the cycling of are water changing.
- Determine the effects of clouds and surface hydrologic processes on Earth's climate.
- Quantify how variations in local weather, precipitation and water resources related to climate variation.
- Determine how weather forecast duration and reliability can be improved.
- Quantify how can predictions of climate variability and can change be improved.
- Predict how will water cycle dynamics change will change in the future.

Water and Energy Cycle Integration

Integration is a necessary, often overlooked and ineffective, scientific process for creating knowledge, products and prediction from huge arrays of disparate information.

Integration involves more than just data – it involves :

- Digging deep into multi-variate problems to quantify uncertanties and identify patterns/synergies
- Enabling free, open, and broad communication colleagues with diverse backgrounds
- Establishing deliberate planning, facilities and mechanisms that encourage integration

Integration baby steps:

- Adopt some common data grids, resolutions, formats
- Adopt some common data analysis tools and techniques
- Develop a common set of water cycle model and observation diagnostics
- Develop some water cycle integration projects with published outcomes.
    - Periodic State of the Water & Energy Cycle assessments.
    - WEC predictability study.
    - WEC model evaluation study.