Challenges in Developing Coupled Earth System Model Data Assimilation

Jeffrey Anderson, Tim Hoar, Nancy Collins, Kevin Raeder, Alicia Karspeck, Peter Lauritzen: National Center for Atmospheric Research
Yongfei Zhang, University of Texas Austin
We are building an ensemble data assimilation system for NCAR’s Community Earth System Model.

- **DART**: Data Assimilation Research Testbed
- **CESM**: Community Earth System Model
- **CAM**: Atmospheric Component
- **POP**: Ocean Component
- **CLM**: Land Component
Fully coupled assimilation will need data from all models at the same time.
Current state of Atmospheric (CAM) Assimilation with CESM ensembles.
Atmospheric Ensemble Reanalysis, 1998-2010

Assimilation uses 80 members of $2^\circ$ FV CAM forced by a single ocean (Hadley+ NCEP-OI2) and produces a very competitive reanalysis.

O(1 million) atmospheric obs are assimilated every day.
Challenge for Earth System Model DA

- Atmospheric components of earth system models may not be as mature as NWP models.
- Model systematic or algorithmic errors may be large.
- Can lead to reduced quality analyses.
- But, DA can help to detect and correct errors!
Example: Gridpoint noise in CAM/DART analysis

Ensemble Mean V at 266 hPa at 6 hours

CAM FV core - 80 member mean - 00Z 25 September 2006
Suspicions turned to the polar filter (DPF)

CAM FV core - 80 member mean - 00Z 25 September 2006
Continuous polar filter (alt-pft) eliminated noise.
Differences mostly in transition region of default filter.
Diagnosing and Correcting Errors with DA

• DA diagnosed a problem that had been unrecognized (or at least undocumented).

• The problem can be seen in ‘free runs’ - it is not a data assimilation artifact.

• Without assimilation, can’t get reproducing occurrences to diagnose.

• Could have an important effect on any physics in which meridional mixing is important.
Current Ocean (POP) Assimilation

- **Obs**
- **DART**
- **3D state**
- **3D restart**
- **POP**
- **Coupler**
- **DATM**
- **2D forcing from CAM assimilation**

**2D forcing**

**3D forcing**

**CAM assimilation**
World Ocean Database T,S observation counts

These counts are for 1998 & 1999 and are representative.

- FLOAT_SALINITY: 68200
- FLOAT TEMPERATURE: 395032
- DRIFTER_TEMPERATURE: 33963
- MOORING_SALINITY: 27476
- MOORING_TEMPERATURE: 623967
- BOTTLE_SALINITY: 79855
- BOTTLE_TEMPERATURE: 81488
- CTD_SALINITY: 328812
- CTD_TEMPERATURE: 368715
- STD_SALINITY: 674
- STD_TEMPERATURE: 677
- XCTD_SALINITY: 3328
- XCTD_TEMPERATURE: 5790
- MBT_TEMPERATURE: 58206
- XBT_TEMPERATURE: 1093330
- APB_TEMPERATURE: 580111

- temperature observation error standard deviation == 0.5 K.
- salinity observation error standard deviation == 0.5 msu.
Physical Space: 1998/1999 SST Anomaly from HadOI-SST

Ensemble Assimilation
48 POP oceans
Forced by 48 CAM reanalyses
Challenges where ocean model is unable, or unwilling, to simulate reality.

Example: cross section along Kuroshio; model separates too far north.
Challenges in correcting position of Kuroshio.

60-day assimilation starting from model climatology on 1 January 2004.

Initially warm water goes too far north.

Many observations are rejected (red), but others (blue) move temperature gradient south.

Adaptive inflation increases ensemble spread as assimilation struggles to push model towards obs.
Challenges in correcting position of Kuroshio.

60-day assimilation starting from model climatology on 1 January 2004.

Model forecasts continue to quickly move warm water further north. Inflation continues to increase spread.

Model forecasts finally fail due to numerical issues. Black dashes show original model state from 10 January.
Challenges in correcting position of Kuroshio.

60-day assimilation starting from model climatology on 1 January 2004.

Assimilation cannot force model to fit observations. Use of adaptive inflation leads to eventual model failure. Reduced adaptive inflation can lead to compromise between observations and model.
Current Land (CLM) Assimilation.

Implementing This is EASY!!!
The HARD part is: *What do we do when only SOME (or none!) of the ensembles have [snow, leaves, ...] and the observations indicate otherwise?*

- Corn Snow?
- New Snow?
- Sugar Snow?
- Dry Snow?
- Wet Snow?
- “Champagne Powder”?
- Slushy Snow?
- Dry Snow?
- “Old Snow”?
- Dirty Snow?
- Early Season Snow?
- “Packed Snow”?
- Snow Density?
- “Crusty Snow”?
- “Snow Albedo”?

The ensemble **must** have some uncertainty, it cannot use the same value for all. The model expert must provide guidance. It’s even worse for the hundreds of carbon-based quantities!
If we restrict ourselves to the simple cases ... here is the early result of an assimilation of MODIS snowcover fraction on total snow water equivalent in CLM.

The model state is changing in reasonable places, by reasonable amounts. At this point, that’s all we’re looking for.
Summary

DART ensemble DA exists for:

- CAM: Multiyear ensemble reanalysis available,
- POP: Ensemble analyses used for decadal prediction initial conditions,
- CLM: Ensemble snowcover analyses.
Plans and Challenges

- Build fully coupled earth system model ensemble assimilation system by end of 2012.
- Methods for dealing with land surface variables.
- Methods for dealing with strongly biased models.