Examples of Research Tools Enabled by CESM Atmospheric Models and DART

Kevin Raeder

Results provided by collaborators, as noted below
Strategy

A highlights tour of some of the techniques which have used DART and variations of CAM in the past few years. Scientific results have been presented in other venues; shown here for illustration.

1) Context in which CAM+DART works
2) Topics which I want to be sure to cover
   A. Generating a better picture of the atmosphere
      WACCM (Pedatella)
      CAM-Chem (Arellano, Barre’, Gaubert)
   B. Analyses as initial conditions (ICs) for forecast experiments
      CAM-SE Hurricane Katrina (Zarzycki)
   C. Sensitivity Studies
      Explosive Cyclogenesis (Chang)
3) A menu of topics to cover as time permits and you (students) express interest.
   A. Finding bugs
      Finite Volume dynamical core noise (Lauritzen)
   B. Analyses as ICs and finding bugs
      Cloud scheme fix (Kay)
   C. ‘Data atmosphere ensemble' forcing of other models (Raeder)
   D. Examining model bias (DART)
The CESM Environment

**Community Earth System Model:**
model all of the geophysics and fluid dynamics which determine the state of a planetary atmosphere, including inputs from biology.

Focused on climate forecasts (decades to centuries).

Each physical medium (gas, water, land, ...) is a separate ‘component’ in the model. They communicate through the ‘coupler’.

- Complicated software environment.
  - Coupler advances each component’s ensemble (‘multi-instance’).
  - State(s) are passed to DART.
  - Into each component DART assimilates the appropriate observations.
  - Modified state(s) are passed back.
  - Coupler spreads influence of obs between components.
  - Future work; ALL observations assimilated into each component.
A ‘CAM’ Assimilation

Community Atmosphere Model = *active* atmosphere, land, seaice, but *data* ocean (SSTs). Run it in ‘multi-instance’ (ensemble) mode. Only atmospheric observations assimilated.

CAM Variants:
1) CAM-dycore `low top' model
2) WACCM `high top'
3) CAM-Chem; extra chemistry equations

- a) Eulerian
- b) Finite Volume
- c) Spectral Element

Theme of the Year 2015
A Better Picture of the Atmosphere: Sudden Stratospheric Warming in WACCM

The Whole Atmosphere Community Climate Model is identical to CAM, except:
- higher model top
- additional chemical, dynamical, and physical processes

Jan-Feb 2009 Sudden Stratospheric Warming; WACCM has had trouble generating these.

Assimilate in the lower atmosphere:
- radiosonde+aircraft temperatures+winds,
- satellite drift winds
- COSMIC radio occultation (GPS)

Assimilate in the middle+upper atmosphere:
- temperature retrievals from
  - TIMED/SABER
  - Aura MLS

New CAM+DART feature:
- Pedatella implemented assimilation using scale height vertical coordinate instead of pressure.

picture from www.ac-ilsestante.it
A Better Picture of the Atmosphere: Sudden Stratospheric Warming in WACCM

Averaged over all longitudes, 70-80N latitude, and all 40 members.

Ensemble mean, zonal mean ozone at 2 hPa

Pedatella, et al. 2013
CAM-Chem + DART: CO

DART
EAKF based analyses
Space and time adaptive inflation
Same framework for all real or simulated (OSSE) observations

Observations
Meteorological observations
Satellite retrievals
  ➢ TERRA/MOPITT-CO
  ➢ METOP/IASI-CO
  ➢ Simulation of future observations

Ensemble of optimized initial conditions
Ensemble of forecast

CESM CAM-CHEM
Online global atmospheric chemistry and dynamic model

Independent observations for evaluation
Satellite retrievals, in-situ aircraft and surface measurement, ground-based infra-red spectroscopy

Arellano started it. Barre’, Gaubert have joined the work
CO Data assimilation of multi-instrument MOPITT and IASI (July 2008)

‘compset’ = F_2000_MOZMAM_CN (CAM5, Mozart chemistry, MAM aerosols)

MOPITT assim – Control Run  IASI assim – Control Run  MOPITT and IASI assim – Control Run

• Assimilation corrects a negative CO bias and improves CO variability: depending on instrument coverage/revisit and vertical sensitivity (Barré et al. 2015a)
Decadal reanalysis of MOPITT-CO retrieval

Gaubert is using compset F_2000_MOZSOA (CAM4, bulk aerosols)

Figures show all assimilation steps (6-hour, 2002-2006)

‘bias’: some groups use this definition. DART software uses the opposite: 

bias = model_state - observations

~93% of CO observations are assimilated throughout. Variations reflect instrument sensitivity (seasonal variation).

Overall good performance: low bias, improvement against independent observations.

Gaubert et al. 2015 in prep
Forecast Studies; Hurricane Katrina

Studying hurricanes -> resolution ≤ 1/4°. Global forecasts at that resolution are expensive and/or slow.

The Spectral Element dynamical core solves the dynamical equations on each element of a ‘cubed sphere’ pattern (shown below). Each element is a 4x4 array of grid points. Scales well on thousands of tasks (short jobs). Allows variable resolution where users actually want it (less expensive).

This refined grid (or variable resolution) is ~1° globally, except for the nested grids, which go down to ~1/4° over the Caribbean.
Forecast Studies; Hurricane Katrina

Studying Hurricane Katrina -> need initial conditions that actually gave rise to Katrina. We can generate those using CAM-SE + DART on the native grid.
“Ensemble sensitivity analyses make use of the different evolution of the forecasts among the different ensemble members and employ correlation and regression between the chosen forecast metric and initial condition state vectors from the ensemble members to derive the sensitivity between any forecast metric and initial conditions.” Chang, et al. 2012

\[ J_M = \text{the ensemble of estimates of the forecast metric, e.g. cyclone minimum pressure at 2010-1-4 06 GMT} \]

\[ x_{iM} = \text{the ensemble of state variable } x_i, \text{ e.g. T at (150W, 50N,850 hPa) at 2010-1-1 06 GMT} \]

\[ "sensitivity" = \frac{\text{cov}(J_M, x_{iM})}{\sqrt{\text{var}(x_{iM})} \sqrt{\text{var}(J_M)}} \]

This sensitivity is dimensionless, and is simply the correlation between \( J_M \) and \( x_{iM} \). This allows sensitivities of different forecast metrics to be directly compared.
Sensitivity Studies

Sensitivity (colors) of the cyclone minimum pressure to sea-level pressure (black contours) at the forecast time and 2 days before.

Some sensitivities can be traced back up to a week before the time of the forecast metric.
Sensitivity Studies

These sensitivities are linear. How well do they represent the actual evolution?

1) Choose a forecast lead time.
2) Perturb the ensemble based on the sensitivity pattern at that time (details later).
3) Run the ensemble forecast to the time of J.
4) Calculate the actual change in J in each member.
5) Compare with the change in J predicted by the sensitivity.

Changes in central pressure (J) due to initial time perturbations of sea-level pressure (x_i).

- mean, 1 standard deviation, and 95% confidence, for one perturbation amplitude.
\( /\) = perfect agreement.

**time: -2.5 days**

![Graph a) showing linear relationship between predicted and actual changes in central pressure.](image)

**time: -5.5 days**

![Graph c) showing linear relationship between predicted and actual changes in central pressure.](image)

Deviations from:
- nonlinearity
- forecast length
- unbalanced perts
- sampling errors

Theme of the Year 2015
Sensitivity Studies

An initial condition perturbation is derived by regressing the forecast metric ($J$) with the initial condition ensemble, following a procedure outlined in Appendix A of Torn and Hakim (2009).

$$\Delta_J = \frac{\text{cov}(J_M, x_{iM})}{\text{var}(J_M)} \alpha$$

$x_{iM}$ = state variable ensemble at the initial time.  
$J_M$ = forecast metric at the forecast time  
$\alpha$ = perturbation amplitude

For the special case of $\alpha = 1/\text{var}(x_{iM})$, $\Delta_J$ = “sensitivity”. 
Optional Topics

A. Finding bugs 4 min
   Finite Volume dynamical core noise
   (Lauritzen)
B. Analyses as ICs and finding bugs 3 min
   Cloud scheme fix (Kay)
C. ‘Data atmosphere’ ensemble forcing of other models (Raeder) 7 min
D. Examining model bias (DART) 1 min
Seeing and Fixing Numerical Noise

Connecting CAM and DART enables looking at CAM(-FV) in new ways. Climate model developers typically look at time averages. Viewing instantaneous fields (a natural part of DA) showed
Model experts quickly suspected the transition from the algebraic filter to the Fourier.
Not a DA artifact; present in free runs
No one at multiple labs over many years reported seeing this.

Model state with a new scheme is almost indistinguishable, except where the noise was.
ICs from DART Help Detect Code Errors

CAM4’s cloud response to sea ice loss; July 2006 to 2007

24-hour forecasts started from DART/CAM analyses identified erroneous cloud response to disappearing sea ice. Jen Kay found that low clouds were only diagnosed over open water, not ice, and the low cloud scheme should have required a well mixed boundary layer.

Short forecasts with a climate model from analyses, compared against observations, point to model improvements.
Some researchers are more interested in the ocean, land, chemistry, ... than the atmosphere(!). CESM’s atmospheric models, with DART DA, can provide forcing for those CESM component models with:

- realistic mean and variability,
- very low computational cost,
- higher resolution than they might otherwise use,
- the consistency of one directional forcing (no influence of X on the atmosphere).

Demonstrated with a 2° CAM4-FV assimilation from 2000-2011 (plus bonus earlier years). Plans for a 1° CAM5.4-FV assimilation (using a 1/4° x daily data ocean) are being formed.

Some uses to which this data set can be put:
1) Multi-year spin-up of slowly evolving models to a realistic state.
2) Injecting sufficient variability into models which lack enough to do good DA.
3) Providing the “off-line” dynamics and thermodynamics for “specified dynamics” chemistry models.
Ensemble “Data Atmosphere” Forcing

Ensemble Temperature Differences

Member 1, level 30

Member 2, level 30

Each member is an equally likely representation of the atmosphere, given the observational network and errors, and the model uncertainty.
Ensemble “Data Atmosphere” Forcing
Members 2-1, Temperature
The Parallel Ocean Program 2 is the ocean component of CESM. Data ATMosphere = fluxes from the atmosphere generated by an independent CAM+DART assimilation. Every 6 or 24 hours, as needed.

"Loosely coupled" assimilation
CAM forcing of POP2

POP “Assim”:
- Initial ensemble = 48 Jan 1st states from a multi-decadal POP run forced by historical atmospheric states -> climatological spread.
- Forced by 2° CAM analyses’ fluxes.
- 2009 World Ocean Database temperature and salinity from the start.
- ARGO floats (subsurface) increasing from 2001 through 2006
- No vertical localization; shallow obs felt by deep ocean.
- No inflation; spread maintained by atmospheric ensemble forcing.

“NoAssim”:
- ensemble forecast identical to Assim, but no assimilation.

Subsequent work (better representativeness error estimates) enabled the use of vertical localization, preventing much of the collapse in spread at depth.
The Community Land Model component of CESM is fundamentally different from CAM and POP; no free-flowing fluid, so no turbulence and the associated perturbation growth. Assimilation with CLM relies heavily on the variability of forcing from the atmosphere.

Lack of mixing -> tremendous spatial heterogeneity on scales from continental down to meters.

Open question: what resolution of forcing from the atmosphere is required to force the land surface with the correct mean and adequate variability?

In these early days of CLM DA, the bar is low.

Ensemble forcing from the atmosphere has limits. The reach of the atmosphere into the earth, on time scales shorter than decades, is only a few meters, with the possible exception of precipitation. Fortunately (or not), there aren’t many observations down there.
CAM Biases

DART naturally provides a wealth of information in observation space:

- bias, rmse, (total) spread
- relative to each observation type,
- for prior and posterior states,
- number of obs used vs. available in the region in each layer = a measure of confidence.

CAM-SE, refined grid (Katrina) assimilation focuses on N Atlantic, where hurricanes form.
Show basics, then let audience choose topics?

① Goal
- CESM environment (schematic)
  1. CESM runs in multi-instance mode
  2. DART is called after each ensemble forecast
  3. Variants of CAM

② Generating a better picture of the atmosphere:
- 1. WACCM and Pedatella
- 2. CAM-Chem (Arellano, Barre', Gaubert)

③ Analyses as ICs for forecast experiments
- 1. CAM-SE Hurricane Katrina (Zarzycki)

④ Finding bugs
- 1. FV noise (Lauritzen)

⑤ Combining ④ and ⑤
- 1. Cloud scheme fix (Kay)

⑥ Sensitivity Studies
- 1. Explosive Cyclogenesis (Chang)

⑦ 'data atmosphere ensemble' forcing (Raeder)

⑧ Examining model bias

⑨ Examining model errors via NMF
- 1. Zagar

⑩ Other potential uses
- 1. Targeted observations for field programs.
- 2. Model parameter estimation
References

CESM: www.cesm.ucar.edu/models/cesm1.2/

N. Pedatella et al., 2013: Application of data assimilation in the Whole Atmosphere Community Climate Model to the study of day-to-day variability in the middle and upper atmosphere. *GRL* **40** pp. 4469—4474.


ICs from DART Help Detect Code Errors

Divergence field in free running CAM at model level 10 (around 200 hPa). Noise visible throughout the run.
1. Figure all assimilation steps (6-hour, 2002-2006)
2. Overall good performance: low bias, improvement against independent observations
3. Conservation of the number of assimilated: variations reflects instrument sensitivity (seasonal variation)

Gaubert et al. 2015 in prep