Coupled {model} Data Assimilation (CDA)

Abhishek Chatterjee

NASA Goddard Space Flight Center
Global Modeling and Assimilation Office

August 7th, 2015
Frontiers in Ensemble DA for Geoscience Applications
Contributors

- NCAR CGD – Alicia Karspeck, Gokhan Danabasoglu, Mitch Moncrieff, Rich Neale, Joe Tribbia
- NCAR IMAGe – Jeff Anderson, Nancy Collins, Tim Hoar, Kevin Raeder, and the DART team
- S. Akella (GMAO), C. Draper (GMAO), S. Karol (NCAR), P. Laloyax (ECMWF), C. Snyder (NCAR), R. Tardif (U. Washington), R. Todling (GMAO), among others
- Funding sources:
  - NOAA Climate and Global Change Postdoctoral Program
What we have talked about…

- **atm-DA**
  - atmosphere component
  - CAM-DART, Raeder et al. [2012], *J. Climate*

- **ocean-DA**
  - ocean component
  - POP-DART, Karspeck et al. [2013], *J. Climate*
What we’ll be talking about…

Coupler

atm

land

DA

DA

ocn

DA

rof

ice

observations
Outline

Coupled Data Assimilation – general issues

Whys, whats and hows of CDA?

Illustrative Examples

CESM-DART

Known unknowns + Unknown unknowns

Implementation -> Results

Scientific & Practical Challenges

Ocean

Atmosphere

Oceanic DA

Atmospheric DA

Ocean

Atmosphere

Science & Practice Challenges

Oceanic DA

Atmospheric DA

Ocean

Atmosphere
Introduction

What do we mean by ‘Coupled DA’?

A first order definition – Assimilation into a coupled model (ESM) where observations in one medium (i.e., atmosphere) are used to generate analysis increments in the other (i.e., ocean)

Why Coupled DA?

- Better and more-balanced (consistent) ocean-atmosphere states
- Better use of near-surface observational data
- Better representation of coupled phenomena
- Better initial conditions for S-I to decadal predictions

Physical processes governing air-sea exchange across the coupled boundary layers

Edson et al. [2007], BAMS
Weak vs. Strong CDA

- **Weak coupling**: background estimates produced by a coupled model, separate analysis updates for each component
- Thus, an observation in one model component cannot directly cause an analysis increment in the other components
  - increments calculated separately - so potentially unbalanced
  - always run coupled model – so background model fields are in balance

- **Strong coupling**: The analysis itself is coupled, so that any observation can affect analysis increments throughout the system
- Strong coupling requires –
  - coupled error covariance models
  - (possibly) well-tuned coupling parameters for mom., buoyancy fluxes
  - (possibly) strong observational constraint for all components
A word of caution about the **Terminology**

**Data Assimilation**

- Existing/Planned CDA Schemes

**Model**

- Uncoupled
  - Single Component
  - Multi Component
  - Cross Component

- Coupled
  - CESM-DART (Multi-component Coupled Model DA)

**Ideal Type**

- ECMWF, UK
- Met Office, UK
- NCEP CFSR, USA
- CMC, Canada
- BMRC, Australia
- JAMSTEC, Japan
- JMA-MRI, Japan
- NOAA-GFDL, USA
- NASA-GMAO, USA
- NCAR, USA
- NRL, USA

**Quasi-fully Coupled!**
Illustrative Examples: ESMs vs Low-order Models

- **ESMs**
  - CESM, GEOS-5, etc.
  - Fully simulates coupled atmosphere-ocean-land-cryosphere system
  - Requires enormous computational resources -> operational centers and/or big research organizations

- **Low-order coupled models**
  - Slab atm-ocean models, coupled Lorenz models, etc.
  - Cheap to run, allows multiple realizations/sensitivity experiments, availability of the ‘Truth’ to rigorously characterize errors in the system
  - Can never capture the full spectrum of dynamics, model biases
  - Suffers from inherent ‘scaling’ issues – proposed algorithms turn out to be incredibly expensive for ESMs
Illustrative Examples: Part 1 (Low-order Models)

- Lorenz (1984, 1990) wave—mean-flow model: **fast chaotic atmosphere**
- Stommel (1961) 3-box model of overturning ocean: **low-frequency AMOC variability** (i.e. no wind-driven gyre)

**Coupling:**
- upper ocean temperature affects mean flow & eddies (**ocean -> atmosphere**)
- hydrological cycle affects upper ocean salinity (**atmosphere -> ocean**)

State vector: 10 variables!

See Tardif et al. [2014], Climate Dynamics
Illustrative Examples: Part 2 (ESMs)

NASA GMAO GEOS-5 (Atm-Land CDA & Atm-Ocean CDA)

Weakly-coupled AODAS Configuration

Atmospheric Model

Skin [Temperature, Salinity] = Bulk [T, S] + [ΔX] + A-ANA INC

Bulk [Temperature, Salinity] + O-ANA INC

NASA GMAO GEOS-5 (Atm-Land CDA & Atm-Ocean CDA)
**Illustrative Examples: Part 2 (ESMs)**

**Coupled ECMWF ReAnalyses (CERA)**

**Incremental variational approach:**
- A common 24-hour assimilation window
- Coupled model to compute observation misfits
- Increments computed separately and in parallel
- Separate background-error covariance model

**Sea Surface Temperature:**
- SST relaxation scheme towards a daily SST analysis product

**Model resolution:**
- Atmosphere: 1.125° horizontal grid with 137 levels
- Ocean: 1° horizontal grid with 42 levels (first layer of 10 meters)
- Wave: 1.5° horizontal grid

**Experiments:**
- Run successfully on short recent periods
**Summary**

- CDA is expected to produce self-consistent state estimates as well as optimal initialization for coupled model predictions
  - growing field of DA application
  - still ironing out consistent terminology, analyses frameworks, etc.

- For ESMs - (currently) CDA falls under two categories – always run coupled model in the background but do DA in a single component or in multiple components

- For low-order models – growing research on the need for CDA, how to specify coupled covariances, what types of observational constraints we need, etc.
Useful References (good starting point…)

- **Presentation by Michele Rienecker** (WMO CAS 2010 Workshop): Good discussion of coupled DA, practical issues

- **Presentation by Keith Haines** (ECMWF Seminar on DA for atmosphere and ocean, 2011): Review of coupled DA implementations and plans from different centers (Met Office, GFDL, JAMSTEC, BMRC, NCEP, Canada)

- **White Paper by Vitart et al.** on ‘Sub-seasonal to Seasonal Prediction: linking weather and climate’ (WWOSC Montreal, 2014): Argument for coupled DA, utility for initializing S2S prediction
Outline

Coupled Data Assimilation – general issues

Whys, whats and hows of CDA?

Illustrative Examples

CESM-DART

Implementation -> Results

Scientific & Practical Challenges

Known unknowns + Unknown unknowns

State of the art: weakly coupled DA

Ocean

Atmosphere

Forecast model

Observations

Data assimilation

(State ← Obs.)

Timescale: 6-12 h

Spatial scale: 1000km

Solver: 4DVAR

Atmospheric DA

Oceanic DA

Timescale: 1-10 days

Spatial scale: 100km

Solver: 3DVAR

Pressure

Temperature

Wind

Moisture

Salinity

Temperature

Currents

Pressure

Ocean
Community Earth System Model (CESM) Components

- All active components (B COMPSET)
- Horizontal Res: Nominal ~1°
- Vertical Discretization:
  - CAM5 – 30 levels (~2 hPa)
  - POP – 60 levels with 10 m resolution in the upper 200 m, gradually expanding to 250 m resolution below 3000 m depth
- 6-hr ocean-atm coupling

CESM Components – High Level Diagram
The coupler is in the middle and communicates with all other components
(adapted from - https://summerofhpc.prace-ri.eu)
Cooperator exchanges fluxes and other necessary information between component models at equal or higher frequency than assimilation update.

Assimilation of conventional (surface, aircraft, etc.) observations independently in each component.
## Timescales Involved in CDA

<table>
<thead>
<tr>
<th>Timescales</th>
<th>CESM-DART</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(internal) Time step of model components</td>
<td>Different for individual models</td>
<td>Dynamics, physics time step (not much control for CDA)</td>
</tr>
<tr>
<td>Coupling frequency of model components to CESM coupler</td>
<td>For example, CAM – 30 minutes POP – every 6 hours</td>
<td>At what time do model components pass information to other components?**</td>
</tr>
<tr>
<td>DA time step for model components</td>
<td>For example, CAM – every 6 hours CLM, POP – every 24 hours</td>
<td>Take into account different timescales at which atm and ocn processes operate</td>
</tr>
<tr>
<td>CESM stop/start</td>
<td>6 hours</td>
<td>--</td>
</tr>
</tbody>
</table>

** "In summary, users should ensure that the following is true, ATM_NCPL = LND_NCPL = ICE_NCPL >= ROF_NCPL >= OCN_NCPL"
**Multi-Component Coupled Data Assimilation (MuC)**

- Coupler exchanges fluxes and other necessary information between component models at equal or higher frequency than assimilation update.
- Assimilation of conventional (surface, aircraft, etc.) observations independently in each component.
Ensemble analysis provides an estimate of analysis and forecast uncertainty:

- (Top Panel) evolution of prior and posterior RMS error
- (Bottom Panels) profile of time-averaged prior and posterior RMS error, total spread and bias relative to the actual radiosonde temp. observations
**Multi-Component Coupled Data Assimilation (MuC)**

- Coupler exchanges fluxes and other necessary information between component models at equal or higher frequency than assimilation update.
- Assimilation of conventional (surface, aircraft, etc.) observations independently in each component.

**Diagram Notes:**
- Initial Conditions
- Atmosphere (CAM)
- Land (CLM)
- Ocean (POP)
- Ensemble members
- Time markers: 00Z, 06Z, 12Z, 18Z, 24Z
- Assimilation of conventional observations
- Coupler exchanges information between models

**Image Source:** NCAR Institute for Mathematics Applied to Geosciences
Coupler exchanges fluxes and other necessary information between component models at equal or higher frequency than assimilation update.

Assimilation of conventional (surface, aircraft, etc.) observations independently in each component.
Ocean-Component Coupled Data Assimilation (Ocean-C)

- Coupler exchanges fluxes and other necessary information between component models at equal or higher frequency than assimilation update.
- Assimilation of conventional (surface, aircraft, etc.) observations independently in each component.
No-Assimilation Coupled Model Run (CESM Free Run)

- Coupler exchanges fluxes and other necessary information between component models at equal or higher frequency than assimilation update.
- Assimilation of conventional (surface, aircraft, etc.) observations independently in each component.
Experiment Configurations

- **MuC CDA**
- **Atmos-C CDA**
- **Ocean-C CDA**
- **CESM Free Run**

<table>
<thead>
<tr>
<th>Jan 11, 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>MuC CDA</td>
</tr>
<tr>
<td>Atmos-C CDA</td>
</tr>
<tr>
<td>Ocean-C CDA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Boreal Summer 2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boreal Winter 2004-2005</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- What are the **impacts on model biases** due to assimilation of observations in **multiple** CESM components?
- What are the **impacts on the modes of tropical intraseasonal variability**, for e.g., Madden-Julian Oscillation?
Reduction in SST Biases

2004 Annual Mean SST
Reduction in SST Biases

2004 Annual Mean

- MuC (minus) Hurrell SST
- CESM Free Run (minus) Hurrell SST

FLOAT_TEMPERATURE
Tropical Eastern Pacific

- Surface

- Depth (m)
  - rmse

- Bias
Differences between key variables from MuC and the free run of CESM
10N-10S averages over DJFM 04-05
Madden-Julian Oscillation (MJO)

Key Features

- Dominant mode of Tropical variability at intra-seasonal time scales
- 30-60 day period
- Eastward propagation of large scale convective precipitation
- See Zhang [2013] BAMS for a thorough review
Precipitation Lead-Lag Correlation Patterns

- Lead-lag correlation coeff. of 20-90 day band-pass filtered precipitation
- 10N-10S averages over DJFM
- IO reference point (75-100 E)
MJO State during Boreal Winter 2004-2005

MuC CDA (Boreal Winter 04-05)

Corr. = 0.96
RMSE = 0.47
Amp. Err = 0.21
Phase Err = -1 day

CESM Free Run (Boreal Winter 04-05)

Corr. = 0.22
RMSE = 1.60
Amp. Err = -0.28
Phase Err = -10 days
MJO State during Boreal Winter 2004

Assimilation in the coupled model

- impacts atmospheric forcing (westerly wind-bursts)
- impacts air-sea coupling (SST – convection relationship)

→ improves simulation of the MJO
MJO Prediction Skill

- After 1 year of assimilation, a 3-week prediction is started
- Caveat: only one event
  - Only ens mean shown
  - MuC retains the MJO signal for ~3-5 days
  - Drift towards model climatology after day 6
What are the differences due to assimilation of observations in a single-component vs. multiple-components?
Impact of DA in a single-component vs. MuC

Atmos-C Experiment

- Comparable to MuC in terms of estimating atmospheric states
- Small reduction in SST bias and/or biases in other oceanic states

Ocean-C Experiment

- Comparable to MuC in terms of estimating oceanic states
- Poor job in simulating MJO or reducing biases in atmospheric states
Impact of DA in a single-component vs. MuC

- What are the differences due to assimilation of observations in a single-component vs. multiple-components?
  - single component assimilation limits the ‘full’ impact of observations across the air-sea interface, even though forecast step may be coupled
  - Ocean-C (Atmos-C) provide limited improvement in atmospheric (oceanic) states

<table>
<thead>
<tr>
<th></th>
<th>Atmosphere</th>
<th>Ocean</th>
</tr>
</thead>
<tbody>
<tr>
<td>MuC</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Atmos-C</td>
<td>✓</td>
<td>--</td>
</tr>
<tr>
<td>Ocean-C</td>
<td>--</td>
<td>✓</td>
</tr>
<tr>
<td>CESM Free Run (baseline)</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
Towards an Experimental Climate Reanalyses

- Primarily motivated by the need for a more self-consistent initial conditions for decadal prediction
- Uses a ‘similar’ setup as MuC CDA, no assimilation in land

**Main challenge:** computational time, for e.g., 2 sim-years per wallclock month on NCAR HPC
- CESM stops and starts every 6 hours,
- Lots of IO
  (CESM model components -> DART)
Ocean results: SST variability

1970-1979 Monthly SST correlation

cesm_6hr_1970b,Hadley-OI SST

Overall high +ve correlation with HADISST

Source: A. Karspeck, S. Karol (NCAR)

1972-73 El Nino event simulated
Atmosphere results: Tropical Cyclones

6hr snapshots of sea level pressure from CAM5

Source: A. Karspeck, S. Karol (NCAR)

A case for CDA for hurricane forecasting?
Summary (1)

- Implementation of CESM-DART
  - multi-component coupled model framework -- test-bed for transitioning to cross-component coupled model scheme

- What are the impacts due to assimilation of observations in multiple-components in CESM?
  - reductions in model biases, improvements in model fidelity and forecasting skill

- What are the impacts on the modes of tropical intraseasonal variability, for e.g., MJO?
  - MuC improves the simulation of MJO state in terms of the amplitude (larger), seasonality (stronger), phase speed (faster)
Summary (2)

- Ongoing reanalyses from 1970-onwards
  - Early results (1970-1980+) are promising
  - Interested in looking at preliminary results from ocean/atmosphere/land/ice components? – contact Alicia (aliciak@ucar.edu)
Outline

**Coupled Data Assimilation – general issues**

**CESM-DART**

**Known unknowns + Unknown unknowns**

**Scientific & Practical Challenges**

**Illustrative Examples**

**Whys, whats and hows of CDA?**

**Implementation -> Results**

**State of the art: weakly coupled DA**

*Ocean* [\(x_{\text{oce}}\); \(y_{\text{oce}}\)]

*Atmosphere* [\(x_{\text{atm}}\); \(y_{\text{atm}}\)]

**Data assimilation (State ← Obs.)**

- Pressure
- Temperature
- Wind
- Moisture
- Salinity
- Temperature
- Currents

**Oceanic DA**

- Timescale: 6-12 h
- Spatial scale: 1000km
- Solver: 4DVAR

**Atmospheric DA**

- Timescale: 1-10 days
- Spatial scale: 100km
- Solver: 3DVAR

**Known unknowns + Unknown unknowns**

**Illustrative Examples**
**Scientific & Practical Challenges**

- All the usual challenges with atm/ocn/land DA remain; for e.g., a few common ones -
  - Model imperfections, biases – DA corrects for random errors
  - Observations – inhomogeneous, sparse for certain components
  - Representativeness error
  - Specification of background error variances
  - Related to DA tool – ensemble (sampling error) vs. variational (adjoint, Jacobians, etc.)

- In addition, we have brought into play exchange across the boundaries – atm/ocean, atm/land \( \rightarrow \) increased complexity
Questions remain…

<table>
<thead>
<tr>
<th>“Known” Unknowns</th>
<th>“Unknown” Unknowns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1) Practical Considerations – especially for ESMs</strong></td>
<td><strong>(1) Benefit of CDA (vs. uncoupled DA) – contribution of CDA in generating accurate initial condition for near-term climate prediction remains to be established at a fundamental level</strong></td>
</tr>
<tr>
<td><strong>(2) Specifying Cross-Covariances across model components, e.g.</strong></td>
<td></td>
</tr>
<tr>
<td>Ocean -&gt; Atmosphere</td>
<td></td>
</tr>
<tr>
<td>• How strong is the influence of the ocean on the atmosphere? Model dependent, resolution dependent, others?</td>
<td></td>
</tr>
<tr>
<td>• How large of an ensemble size do you need to capture the signal from the atmosphere? With small ensemble size, sampling error dominates the ‘small’ signal from ocean to atm.</td>
<td></td>
</tr>
<tr>
<td>Atmosphere -&gt; Ocean</td>
<td></td>
</tr>
<tr>
<td>• Do we really need coupled covariances or is the coupling between model components enough to propagate the information?</td>
<td></td>
</tr>
<tr>
<td>• Interacting slow (ocean) &amp; fast (atmosphere) components – and the fast component is noisy! How to reconcile differences in timescales?</td>
<td></td>
</tr>
</tbody>
</table>
Summary

Coupled Data Assimilation – general issues

- CDA is expected to produce self-consistent estimates
- Big push, especially in operational centers towards cross-component coupled model DA

CESM-DART

- CESM-DART – currently setup as a multi-component coupled model DA
- Initial runs designed to test the feasibility of the framework, benefits for sub-seasonal prediction
- Ongoing runs designed to improve decadal prediction skill

Known unknowns + Unknown unknowns

- Usual challenges with atm/ocean/land DA are applicable
- Fundamental “mathematical”/”methodological” development required on specifying cross-component covariances
- Big big...big unknown – how much benefit will coupled DA really provide?
QUESTIONS?

abhishek.chatterjee@nasa.gov