Notes from 1st and 2nd SingleTrack DA Meetings

Notes from 1st DA WG meeting, 24 Jan 2018; 2nd Meeting 5 Feb 2018

Notes take the form of additions to and revisions of DA presentation from NCAR Atmospheric Modeling Workshop, Jan 2017

-- original presentation file is here. I've restored the original presentation (as current version) but you can see the changes made at the meeting by looking in the version history (version from 24 Jan 2018).

In attendance, Mtg 1: (FL) Ben Gaubert, Glen Romine, Helen Worden, Wen-Chau Lee, Luca Delle Monache, Rajesh Kumar, Jake Liu, Soyoung Ha, Nick Pedatella, Jenny Sun, Craig Schwartz, Chris Snyder, (ML) Joe Tribbia, Steve Goldhaber, Moha El Gharamti, Kevin Raeder, Jeff Anderson.

Science Applications

1. Coupled systems
   (atmos-chem, atmos-ocean, atmos-land, or lower-upper atmos)
2. Prediction, especially high-res, severe weather. Air-quality prediction too.
3. Source estimation for constituents, parameter estimation
4. Multiscale DA (analyses resolve multiple time, space scales) and prediction
5. Using DA info to improve atmospheric (and other) models, & understand bias
   a. Link to parameterizations, physics
6. Reanalyses and reforecasts, atmospheric and coupled

Additional Science Applications

1. OSSES, OSE, observation impact, FSO
2. Process studies using sensitivities from DA
    a. E.g. convective initiation: ensemble-based sensitivity calculations
    b. Targeted observations
3. prediction
    a. For climate
    b. For nwp, esp. For high res
    c. Weather to climate / “seamless”

4. Real-time experiments and field-campaign support

Key Requirements for Atmospheric Model

1. DA and ensemble-forecasting “friendly”
2. Coupled to land, chem/constituents, ocean, etc
   … agnostic on how coupling should occur, subject to (1)
3. Skillful, w/ minimal systematic error over many assimilation cycles … physical parameterizations are crucial

New ideas:

1. Robustness and reliability of model
2. Model capabilities span range of scales of interest
   a. emphasis on global
   b. LES is of interest
   c. Nesting, regional

3. Skillfull and small bias in stratosphere (impt for radiance DA, subseasonal

DA Friendliness

1. Clear definition of prognostic state
    a. Ability to specify prognostic state when configuring model (e.g. difft physics)
2. Stable interface: formats, scripting, etc modified infrequently
3. Can cheaply stop and restart from own output
4. Modeling system can compute forward operators at high frequency efficiently
5. Easy to control computation and output of diagnostic quantities (perhaps gridded)
6. Scalability
7. TLM and adjoint (forward operators, and model)

\+ Includes forward operators (calculation of observed variables at specified times and locations) as diagnostic calculation during model integration

\+ Has tangent-linear and adjoint versions

Ensemble-Forecasting Friendliness

1. Capable of embarrassingly parallel runs of ensemble
   … scripting of multiple jobs OK; in-core ensemble also OK

+ If in-core ensemble, capability is default
+ Supports stochastic approaches to represent model error (also use same tools for ICs?)
+ Diagnostic routines work with ensembles
+ Run multiple of configurations of model w/in ensemble (resolution, physics)

Atmospheric Model Requirements: Forecasting

1. Can make coupled forecasts with suite of Earth System Models (including global ocean, regional ocean, land surface, chemical constituents, ice, …)
2. Useful forecast skill for research across wide range of space and time scales.
   a. Global to cloud scale
   b. LES (desirable)
   c. Depths from troposphere, stratopshere, to exobase (500km)
3. Capable of supporting stochastic forecasts (physics, noise addition)

Atmospheric Model Requirements: Computing

1. Efficient/scalable ensemble forecasts on available high performance computing including large numbers of tracers
2. Can perform a sequence of short integrations with minimal computational overhead relative to a single long integration
3. Can stop and restart exactly
4. Ensemble members can have different resolution/physical configurations
5. Easy-to-use interface to compute/output diagnostic quantities
6. Configurable to be efficient for range of space/time/depth scales including limited-area

Atm. Model Requirements: DA Specific

1. Precise, accessible definition of prognostic state
2. Easy to invoke a range of damping and smoothing operators so that model is stable when using DA increments
3. Incremental analysis update capability
4. Digital filter initialization
5. Can compute ensemble forward operators efficiently at high frequency
6. Tangent linear capability and adjoint model capability
7. Tangent linear and adjoint for model computed forward operators (if 5 on this slide is implemented)

Atmospheric Model Requirements: Support

1. Ongoing support for DA users
2. Stable interface that is updated infrequently, updates supported
3. DA users should be consulted before support is discontinued for legacy versions of model

Use case for ensemble DA

Key feature: DA implementation controls algorithm (required unless noted)

Cycle:

1. Obtain metadata from model describing state variables
2. Ask model to make ensemble of forecasts to a set of specified lead-times
3. Optional: Ask model to compute forward operators during forecasts
4. Ask model to interpolate state variables to arbitrary space-time location

Use case for variational DA

Variational requires everything from ensemble use case, plus an additional forecast by the model (potentially at different resolution from the ensemble).

For 4DVar or for computation of sensitivity of forecast to observations (not DA, but useful diagnostic), DA also asks for integrations of TLM and adjoint of model.

Use case for variational DA

1. Ask the model to provide state at specified times.   Read the background (or the first guess) fields (could be at multiple times for FGAT mode) from the model restart file (prefered) or model output file and write analysis in restart file or output file (shall replace analyzed fields in the background file instead of rewrite the whole file).
2. Background error modeling or ensemble covariance localization could use model integration, e.g., via diffusion operator technique
3. 4DVAR specific: need non-linear model integration to obtain the forecast trajectory (store in memory or in disk) and compute OMF within a time window, need TL/AD model integration to compute gradient of the cost function in each inner iteration
4. Adjoint-based forecast error sensitivity to observations needs the backward AD model integration

Notes

“prognostic state” = variables necessary at t to advance to $t + dt$ (and not static across all t’s)

“Cheap restarts” = cost of single, long integration ~ cost of integration over same time interval, but with many stops, restarts

Case control system must support both DA-calls-model (e.g. via scripts as in WRF, MPAS, etc) and model-calls-DA scenarios

I/O is potentially major issue in future, but details unclear