Adaptive Spatially-varying Variance Inflation in an Ensemble Filter

Kevin Raeder, Jeff Anderson, Tim Hoar, Nancy Collins, Hui Liu
NCAR Data Assimilation Research Section (DAReS)
Assimilation System
Data Assimilation Research Testbed
20-member ensemble filter
Sondes, ACARS, Drift Winds
0.2 Gaspari-Cohn Localization

Forecast Model
Community Atmosphere Model
T85L26
Climatological I.C.s
January, 2003
Six-hour Forecast RMS Error and Spread: No Inflation

Observation Space for 500 hPa North America ACARS Temperatures.

Spread (dashed) gets too small.

Confidence in prior too great.

Observations given too little weight.

RMS error (solid) grows after a few days.
**Sources of insufficient prior spread.**

1. **Model error.**
   - Prior model spread doesn’t account for errors (including representativeness).
   - Result can be treated as lack of spread. Models have too little error growth, too.

2. **Ensemble sampling error.**
   - Expected value of |sample correlation| vs. true correlation is too large for small ensemble sizes.
   - Observations systematically reduce spread of state variables too much.
One Solution: Prior State Space Inflation

Prior State Ensemble  \[\rightarrow\]  Mean  \[\rightarrow\]  Inflated Prior Ensemble

Linearly expand around mean.
Spread is increased.
**Mean** is unchanged.
Correlation with other inflated state variables unchanged.

Applied independently to each state variable.
Spatially-varying temporally-adaptive inflation: Hierarchical Bayesian

1. Inflate Prior
2. Apply Forward Obs. Operator
3. Model Advance
4. Observed Value and Likelihood from Instrument
5. Prior Obs. Space Estimate

Is spread smaller than expected?
Yes: increase inflation.
No: decrease inflation.

Use joint prior of obs. and each state variable to regress inflation increment.

Each state variable has its own inflation distribution (Gaussian).

Observation Space for 500 hPa North America ACARS Temperatures.

- Fewer observations rejected.
  (If prior error is more than 3 times expected, obs. are rejected).
- RMS Error is Reduced.
- RMS does not increase with time.
- Spread is increased.
RMS Error Reduced in General: Number of Obs. Rejected Reduced

Obs. Space North America ACARS Temperatures vertical profile.
Spatial and Temporal Structure of Adaptive Inflation
Zonal Wind Inflation, 266 hPa
Why Does Inflation Get so BIG?

Correct Evolution

Model Evolution

Assimilating Dense Obs. Leaves Tiny Spread.

t₁

t₂

t₃

Observed Value

Spread still small after advance. Big inflation needed to handle model error.

Expected RMS
Conclusions (PLUS SEE OUR POSTER: A31B-06)

1. Ensemble Assimilations have too little variance.
2. Inflation can correct for this.
3. Hierarchical Bayesian algorithm automatically gives good inflation.
4. Applied to CAM, WRF, other large models without tuning.
5. Error is reduced, spread increased.
6. Available as part of DART from www.image.ucar.edu/DARes/.
After 6 hours.

NCEP

DART/CAM

CAM starts with climatology!
Nearly zonal.

Difference.
NCEP

After 1 day.

DART/CAM

Difference.
After 3 days.

NCEP

DART/CAM

Difference.

CAM gains zonal structure.
After 7 days.

NH converged.
SH poorly observed.
6-Hour Forecast and Analysis Observation Space Temperature RMS

RMS Error: Tropics

RMS Error: Northern Hemisphere

Tropics

Northern Hemisphere
6-Hour Forecast and Analysis Observation Space Wind RMS

RMS Error: Tropics

RMS Error: Northern Hemisphere

NCEP Opnl. Analysis
NCEP Opnl. Guess
T85 CAM Analysis
T85 CAM Guess

Tropics

Northern Hemisphere
Adding new models, new observations is simple.

1. Incorporating existing model requires handful of interfaces.
   A. No need for linear tangents or adjoints.
   B. Finite Volume CAM added in 1-month by postdoc.

2. Adding observations also straightforward.
   A. Only need forward operator (map state to expected observation).
   B. No linear tangents or adjoints.
   C. Several different GPS operators added in weeks.

GPS and other novel observations may help detect climate model bias.

GPS provides soundings of temperature and water world-wide.
Assimilating GPS Radio Occultation Observation
Assimilated as refractivity along beam path.
Complicated function of T, Q, P and ionospheric electric field.

Get a sounding as GPS satellite sets relative to low earth satellite.
DART compliant models (largest set ever with assim system?)

1. Many low-order models (Lorenz63, L84, L96, L2004,...).
2. Global 2-level PE model (from NOAA/CDC).
3. CGD’s CAM 2.0, 3.0, 3.1 (global spectral model)
3a. CGD’s CAM 3.1 FV (global finite volume model) with chemistry.
4. GFDL AM GCM (global grid point model).
5. MIT GCM (from Jim Hansen; configured for annulus).
6. WRF model (regional prediction grid point).
6a. WRF column physics model.
7. NCEP GFS (operational global spectral; assisted by NOAA/CDC).
8. GFDL MOM3/4 (global grid point ocean model).
9. ACD’s ROSE model (upper atmosphere with chemistry).

Also models from outside geophysics.
This allows for a hierarchical approach to filter development.
DART compliant Forward Operators and Datasets

Many linear and non-linear forward operators for low-order models.

U, V, T, Ps, Q, dewpoint for realistic models.

Radar reflectivity, doppler velocity, GPS refractivity for realistic models.

Mopitt CO retrievals.

Can ingest observations from reanalysis or operational BUFR files.

Can create synthetic (perfect model) observations for any of these.
Additional enhancements available for Earth System Analysis

1. **Smother**: uses observations in past and future to estimate state.

   Filter algorithms are naturally scalable.
   Run many copies of models.
   Impact of observations on model variables can be done in parallel.
   Algorithm has natural reformulations for different platforms.

3. Parameter estimation in large models.
   Use data to constrain parameters.
   Implementation is trivial.
   Interpretation is still VERY tricky.
   Did this for gravity wave drag efficiency in CAM.
Climate Model Parameter Estimation via Ensemble Data Assimilation

T21 CAM assimilation of gravity wave drag efficiency parameter.

Oceanic values are noise (should be 0).

$0 < \text{efficiency} < \sim 4$ suggested by modelers.

Positive values over NH land expected.
Problem: large negative values over tropical land near convection.
May reduce wind bias in tropical troposphere, but for ‘Wrong Reason’.

Assimilation tries to use free parameter to fix ALL model problems