Regionally Enhanced Global (REG) Data Assimilation

Szunyogh István

Texas A&M University
Department of Atmospheric Sciences

OMSz, Budapest, július 6, 2017
**General idea:** Assume a forecast center that prepares both global and limited area analyses. Let’s prepare the global and limited area analyses of the center valid at the same time by a single computational process (Merkova et al. 2011; Yoon et al. 2012; Kretschmer et al. 2015)

**Illustration for 3 limited area domains:**

![Map of COAMPS Regions with three highlighted areas: Europe (EUR), Northeast Pacific (NEPAC), and North America (CON).]
Potential benefits of the approach

- Reduced development, maintenance, and computational cost
- The limited area analysis is informed about the large scale flow and may benefit from observations outside of the limited area domain
- The limited area analysis may benefit from the direct use of satellite radiance observations
- The global analysis may benefit from the availability of higher resolution model information in the interior of the limited area domain

**REG DA** is the latest algorithmic implementation of the general idea from my research group (Michael Herrera and Adam Brainard) and our collaborators (Craig Bishop and Dave Kuhl from NRL)
“The sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work.” – John von Neumann
The Model of Data Assimilation

Sequential DA–Extended Kalman Filter (EKF):

\[ \mathbf{x}^a = \mathbf{x}^b + \delta \mathbf{x}^a, \quad \delta \mathbf{y} = \mathbf{y}^o - \mathcal{H} \left( \mathbf{x}^b \right), \]

The analysis increment \( \delta \mathbf{x}^a \) is computed by \( \delta \mathbf{x}^a = K \delta \mathbf{y} \), where

\[ K = P^b \mathbf{H}^T \left( \mathbf{H} P^b \mathbf{H}^T + \mathbf{R} \right)^{-1} \]

or a 4D-Var
A (flow-dependent) background bias $b (x^b)$ leads to a biased prediction $\mathcal{H} (x^b)$ of the observations, which in turn leads to a large innovation $\delta y = y^o - \mathcal{H} (x^b)$.

There are two alternative strategies to deal with this situation:

1. **Strategy 1**: Pushing $x^a$ closer to $y^o$, (further away from $x^b$)
2. **Strategy 2**: Keeping $x^a$ closer to $x^b$, (further away from $y^o$)

**Strategy 1** works when the resulting large analysis increment $x^a$ does not lead to a strong adjustment process (in my experience, a real but rare situation)

**Strategy 2** is advantageous when the original increment $x^a$ would lead to a strong adjustment process (in my experience, the typical situation)
Two Ways to State the Problem

1. The background \( \mathbf{x}^b \) is a state on (or very near to) the model attractor. The observation \( \mathbf{y}^o \) observes a state on the attractor of the true atmosphere. When there is a substantial difference between the two attractors, forcing the analysis near to the true attractor leads to a “jump” of the state to the model attractor (strong adjustment process).

2. By definition, the observation function should satisfy

\[
\mathbf{y}^o - \mathcal{H}(\mathbf{x}) = \mathbf{\varepsilon}^o,
\]

for any value of \( \mathbf{x} \), where the observation error \( \mathbf{\varepsilon}^o \) is a random variable with mean 0 and covariance matrix \( \mathbf{R} \)

\( \mathcal{H}(\mathbf{x}) \) should include all corrections necessary to satisfy Eq. (1)
Schematics Illustration for a 2-dimensional State Space

- Observations
- Real attractor
- Model attractor
- Short term model integration
- Unknown true analyzed state
- Analysis
- No bias correction
- Unknown true forecast state
- Model attractor
- Forecast model integration
- Unknown true forecast state
- Also $\mathcal{M}(x^b)$ for no bias correction
- Bias correction
- Observation $e^0$
- Short term model integration

Szunyogh
REG DA
Examples for the Correction of $\mathcal{H}(\mathbf{x})$

- **Observation bias correction** (e.g., for radiance observations) is a correction of $\mathcal{H}(\mathbf{x}^b)$ rather than $\mathbf{y}^o$
- For instance, adjusting the surface pressure to the model orography should be done in $\mathcal{H}(\mathbf{x})$ rather than correcting $\mathbf{y}^o$
- $\mathcal{H}(\mathbf{x}^b)$ can be biased because
  - $\mathcal{H}(\cdot)$ is biased, and/or
  - $\mathbf{x}^b$ is biased
- **REG DA** is an attempt to account for the bias in $\mathbf{x}^b$ in the computation of $\mathcal{H}(\mathbf{x}^b)$:
  \[
  \mathbf{x} = (c - 1)\mathbf{x}_g + (c)\mathbf{x}_\ell \quad 0 < c < 1.
  \]
  where $c$ is the **blending coefficient**, $\mathbf{x}_g$ is the **global model state** and $\mathbf{x}_\ell$ is the **limited area model state**, but the data assimilation updates the **global background** $\mathbf{x}_g^b$. 

Reg DA
Manuscripts: Herrera et al. 2017 and Brainard et al. 2017

- **Global model:** NAVGEM T119
- **Limited area model:** COAMPS with 32 km
- **Data Assimilation:** NAVDAS-AR with TLM at resolution T119
- **Blending Grid, \( y^o - \mathcal{H}(x) \):** T319 Gaussian Grid
- **Blending Coefficient:** \( \alpha = 0.3, 0.5, 1.0 \)
- **Number of limited area domains:** 3
- **Test Period:** from 0000 UTC, October 1, 2012 to 0000 UTC, November 1, 2012 (includes Hurricane Sandy)
- **Limited Area Analysis:** Interpolated global analysis (using the standard interpolation routines)
- **Verification:** ECMWF analyses (NAVGEM, COAMPS), RAOB (COAMPS)
Configuration: Limited Area Domains

COAMPS Regions

- Europe (EUR)
- Northeast Pacific (NEPAC)
- North America (CON)
Atmospheric Balance in NAVGEM

T119 Control

Blend Skip

50% Blend
Geopotential Height: NAVGEM Forecasts

Red: Improvement  Blue: Degradation

Geopotential Height

Northern Hemisphere  North America  Europe  Northeast Pacific

30% 50% 100%
Pressure (hPa)

Lead Time (hour)
Geopotential Height: COAMPS Forecasts

Red: Improvement  Blue: Degradation

Blend Skip - REG 4D Var, Geopotential Height

30%

50%

100%

Pressure (hPa)

200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000

00 12 24 36 48 60 72

Lead Time (hour)

00 12 24 36 48 60 72

00 12 24 36 48 60 72

00 12 24 36 48 60 72

00 12 24 36 48 60 72

3 2 1 0 [-1] [-2] [-3]

Szunyogh  REG DA
Geopotential Height: NAVGEM Forecasts

Red: Improvement  Blue: Degradation

Szunyogh  REG DA
v-wind: NAVGEM Forecasts

Red: Improvement  Blue: Degradation

Meridional Wind

Northern Hemisphere  North America  Europe  Northeast Pacific

Pressure (hPa)
30%  50%  100%

Lead Time (hour)

Szunyogh  REG DA
v-wind: COAMPS Forecasts

Red: Improvement  Blue: Degradation

Blend Skip - REG 4D Var, Meridional Wind

30%

50%

100%

Pressure (hPa)

Lead Time (hour)

Composite  North America  Europe  Northeast Pacific
Red: Improvement  Blue: Degradation

v-wind: COAMPS Forecasts (RAOB verification)
Air Temperature: COAMPS Forecasts (RAOB verification)

Red: Improvement  Blue: Degradation

RAOB Blend Skip - REG 4D Var, Temperature
Concluding Remarks

- **Results** with prototype REG 4D-Var system at reduced resolutions are promising

- **Ongoing and Future Efforts** (Max Gavryla’s thesis research):
  - Implementation of a different approach for the preparation of the limited area analyses
  - Analysis/forecast experiments at operational resolution
  - Replacement of 4D-Var by the Local Ensemble Transform Kalman Filter (LETKF)—Development of REG LETKF