Model error representation in WRF/DART

So-Young Ha, Judith Berner, Chris Snyder

MMM/NCAR
Why model uncertainty representations

- All current operational ensemble systems are underdispersive; The rms error grows faster than the spread. => the best estimate of the true atmospheric state is on average more often outside the range of predicted states than statistically expected. (Buizza et al. 2005)

- Small uncertainties in the initial state and NWP model lead to forecast errors and flow-dependent predictability.
Model uncertainties in short-range weather prediction

- Forecast error = IC error + Model error + LBC error

- Model errors represented by multi-model, multi-physics, multi-parameter, and stochastic schemes

- Retrospective case studies using the AFWA’s mesoscale ensemble prediction system (Hacker et al. 2011; Berner et al. 2011) showed that
  
  ⇒ Including a model-error representation leads to ensemble systems that produce significantly better probabilistic forecasts than a control physics ensemble that uses the same physics schemes for all ensemble members.

  ⇒ In overall, the stochastic kinetic-energy backscatter scheme is comparable or superior to the multi-physics ensemble.

  ⇒ The best performing ensemble system is obtained by combining the multi-physics scheme with the stochastic kinetic-energy backscatter scheme.
Model uncertainties in WRF/DART cycling

• Control-physics (CP) ensemble: each ensemble member uses the same physics configuration, but ensemble prior spread is adaptively inflated based on the observation likelihood and the prior PDF right before the analysis step.
• Multi-physics (MP) ensemble: each ensemble member uses a different set of physics schemes.
• Stochastic kinetic-energy backscatter (BS) ensemble: each ensemble member is perturbed by a stochastic forcing term that represents the statistical fluctuations in the subgrid-scale fluxes.
Multi-Physics ensemble configuration

- AFWA’s Mesoscale Ensemble Prediction System (MEPS)

<table>
<thead>
<tr>
<th>Member (JME mem)</th>
<th>Physical parameterizations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Surface</td>
</tr>
<tr>
<td>1</td>
<td>Thermal</td>
</tr>
<tr>
<td>2</td>
<td>Thermal</td>
</tr>
<tr>
<td>3</td>
<td>Noah</td>
</tr>
<tr>
<td>4</td>
<td>Noah</td>
</tr>
<tr>
<td>5</td>
<td>Noah</td>
</tr>
<tr>
<td>6</td>
<td>Noah</td>
</tr>
<tr>
<td>7</td>
<td>RUC</td>
</tr>
<tr>
<td>8</td>
<td>RUC</td>
</tr>
<tr>
<td>9</td>
<td>RUC</td>
</tr>
<tr>
<td>10</td>
<td>RUC</td>
</tr>
</tbody>
</table>
The Kalman Filter (KF)  

Assume  
- $x^t \sim N(\bar{x}^f, P^f)$; Gaussian forecast errors  
- $\epsilon \sim N(0, R)$; Gaussian observation errors  

KF analysis implements Bayes rule for Gaussians  
- analysis equations:  
  \[ \bar{x}^a = \bar{x}^f + K(y - H\bar{x}^f) \quad ; \quad P^a = (I - KH)P^f, \]  
- Kalman gain  
  \[ K = P^fH^T(HP^fH^T + R)^{-1} \]  

Computationally difficult unless problem is small  
- $P^f, P^a$ are $N_x \times N_x$, w/ $N_x = \text{dim } x$
Ensemble Kalman Filter (EnKF)

- **EnKF analysis step**
  - As in KF analysis step, but uses sample (ensemble) estimates for covariances ⇒ the huge matrix $P_f$ is never explicitly computed.
  
  $P_f H^T = \frac{1}{N-1} \sum_{i=1}^{N} (x^f - \overline{x^f})(Hx^f - \overline{Hx^f})^T$

  $HP_f H^T = \frac{1}{N-1} \sum_{i=1}^{N} (Hx^f - \overline{Hx^f})(Hx^f - \overline{Hx^f})^T$

  where $\overline{x^f} = \frac{1}{N} \sum_{i=1}^{N} x^f$ and $\overline{Hx^f} = \frac{1}{N} \sum_{i=1}^{N} Hx^f$

  $y^f = Hx^f$ is the forecast, or prior observation.

  - Output of EnKF analysis step is ensemble of analyses

- **EnKF forecast step**
  - Each member integrated forward with full nonlinear model to provide flow-dependent background error covariance
  - Monte-Carlo generalization of KF forecast step
Ensemble Kalman Filter (EnKF) in DART

- Data Assimilation Research Testbed (DART) is general software for ensemble filtering:
  - Assimilation scheme(s) are independent of model
  - Interfaces exist for numerous models: WRF (including global and single column), CAM (spectral and FV), others
  - See http://www.image.ucar.edu/DAReS/DART/
Experiment design

Grids
D1: 123 x 99 (45-km)
D2: 163 x 106 (15-km)
41 levels, two-way nesting

IC/LBCs
- 1°x1° GFS analyses were used for initialization in both domains
- 1°x1° GFS forecasts were used to generate lateral boundaries at 45-km grid four times a day

Ensemble
- 50-member ensemble
- WRF/DART to generate analyses and forecast

Cycling period: 1-10 June 2008 (3-hrly)
Observations for data assimilation

- MADIS (Meteorological Assimilation Data Ingest System)
  - RAOB - u, v, t, td, surface altimeter
  - METAR - u, v, t, td, surface altimeter
  - Marine - u, v, t, td, surface altimeter
  - ACARS - u, v, t, td
  - Surface observations: metar (for assimilation) and integrated mesonet (for verification)
Obs-space diagnostics (mesonet verification)

**U10**

**V10**
Obs-space diagnostics (mesonet verification)
Obs-space diagnostics (sounding)
An MCS case in summer’08

2008-06-08_21:00 UTC

2008-06-09_06:00 UTC
Ensemble spread (3-h forecast)

2008-06-08_18:00:00 UTC

V_d02 prior spread

T_d02 prior spread
Analysis increment in ensemble mean

A-B at 2008060818

T[K] at k=1
12-H accumulated rainfall at 15-km grid

2008-06-09_06:00:00 UTC

NCEP Stage IV

Total precipitation [mm]
Summary for model errors in WRF/DART

• The meso-scale ensemble system generally suffers from under-dispersiveness.

• Including model error representation improves the analysis and the following forecast compared to the control-physics ensemble that uses the same physics combination for all members.

• The stochastic Kinetic Energy Backscatter scheme was well tuned to improve the atmospheric state near the surface. The SKEBS outperforms the multi-physics ensemble in the short-term forecast.

• Multi-physics ensemble needs to be more investigated for the mean bias errors and the overdispersiveness near the surface depending on the physics combinations.
Ongoing work at NCAR

• SKEBS released with WRF3.3.
• Development ongoing: plans to introduce flow-dependent dissipation and vertical structure
• Impact of multi-physics and stochastic backscatter scheme in ensemble data assimilation
• Understand differences between multi-physics and stochastic representation physically
• A perturbed physics-tendency scheme (Buizza et al., 1999) is currently being tested (revisiting from earlier work)
• Extend ensemble forecasts with different model error techniques for probabilistic verification