Improving Convective Forecast Skill via Ensemble Data Assimilation with WRF-DART

Glen Romine
NCAR Boulder, CO

Data Assimilation Research Testbed (DART) – A community facility for ensemble data assimilation

DART provides a tool for generating ensemble initial conditions consistent with the forecast model. Ensemble forecast can be leveraged in targeted observation studies.

Goal: Reliable mesoscale forecasts of intense convection - e.g. 6-18 Fhr; severe weather ‘watch’ guidance

WRF/DART forecast system realtime demonstration: Mesoscale Predictability Experiment (MPEX) – Spring 2013

DA driven field campaign
Sought ‘mesoscale’ features associated with mid-tropospheric disturbances that might modify the near storm environment later in the day.
MPEX – Mesoscale Predictability Experiment

Goals

1) Improve convection permitting forecasts by reducing initial condition uncertainty through targeted sub-synoptic observations upstream of anticipated convective events

2) Sample the near storm environment to better understand how developing convection impacts subsequent predictability

Ops from 15 May – 15 June 2013, 15 flights, 18 upsonde missions
WRF and DART configuration options

WRF V3.3.1
- 415x325x40 [1045x870] (E-W)x(N-S)x(B-T), model top 50 mb
- 15 [3] km grid spacing
- Key physics options: Tiedtke, RRTMG, Thompson, MYJ, NOAH
- Ensemble forecasts – 30 members + GFS control 12/00 UTC daily

DART development branch (approx. Kodiak release)
- 50 member ensemble
- 6 hourly continuous cycling assimilation
- adaptive prior inflation, sampling error correction, adaptive localization
- conventional obs (ACARS, METAR, Radiosondes, Marine, Profiler, CIMMS motion vectors), ~180k obs/day

Continuous cycling – for 46 days

See Romine et al. 2013 for more details
MPEX – enhanced observation density

Number of computer model grid points >> observation points, fewer fields measured

Each marker is location of an observation used in an analysis (at any height)

Red ‘X’ are balloon soundings which give most detailed information in vertical column
MPEX – enhanced observation density

Each marker is location of an observation used in an analysis (at any height)

Red ‘X’ are balloon soundings which give most detailed information in vertical column

Able to sample up to $\frac{1}{3}$ of drop points during MPEX, equivalent to more rawinsondes, as well as MTP and flight level

Number of computer model grid points $\gg$ observation points, fewer fields measured
Case 1: 2013-05-15  GOES water vapor channel

GOES-15 gvar_ch3 brightness  Date: 2013-05-15_1200

Brightness (K)

MAX
-17.85
MIN
-55.35
Case 1: 2013-05-15  water vapor + ENS vorticity

Ensemble mean analysis pos.
absolute vorticity

Exploring synthetic radiance products to compare against GOES for verification

Spatial similarities between analysis kinematics and upper tropospheric moisture

Abs. vorticity, 10-45 by 5 ( $10^{-5}$ s$^{-1}$ )
Case 1: 2013-05-15 water vapor + ENS vorticity + Sondes

Ensemble mean analysis pos.
absolute vorticity

Drop locations & NWS sondes

Sampled upshear side of upper level disturbance in TX – How were these points selected?

GOES-15 gvar_ch3 brightness  Date: 2013-05-15_1200

NWS - circles, MPEX drops - stars
Abs. vorticity, 10-45 by 5 (10^-5 s^-1)

Brightness (K)

MAX -17.85
MIN -55.35

Ensemble Sensitivity Analysis (ESA)

\[
\frac{\partial J_e}{\partial x_j} \equiv \text{cov}(\delta J, \delta x_{o,j}) D_j^{-1} = \frac{\text{cov}(J, X_j)}{\text{var}(X_j)}
\]

Ancell and Hakim 2007, Hakim and Torn 2008

- Ensemble-based method of computing forecast sensitivity to the initial conditions (or prior forecast states)
- From linear regression based on ensemble:
  - Dependent variable is ensemble estimate forecast metric (e.g. average accumulated precipitation over an area)
  - Independent variable is ensemble estimate of state variable (e.g. mid-tropospheric humidity)
- Works best when the forecast metric is more continuous
- Can also compare subset of members that have particular metric properties (e.g. max – min metric groups)
Sample ensemble sensitivity

Warm (cool) colors – increase (decrease) in field at 12 UTC associated with more precipitation in area at right from Fhr 33-36

500 hPa vorticity valid 2013051512 (F024)

Shifting shortwave in SW Texas further ESE is associated with more precipitation in box
Hypothetical observation impact

- Ensemble-based method allows for estimate of observation impact
  - Can get change in metric value if you know observation properties, ensemble metric values and observation value itself
  - Can still get reduction in variance knowing only first two above (no need for observation)

\[
\delta J = J(HX^b)^T(HP^bH^T + R)^{-1}[y - H(x^b)]
\]

Change in forecast metric
Forecast metric observation covariance \times inverse innovation covariance

\[
\delta \sigma = -J(HX^b)^T(HP^bH^T + R)^{-1}HX^bJ^T
\]

Change in forecast metric variance
Forecast metric observation covariance \times inverse innovation covariance

See Torn and Hakim 2008, MWR
Hypothetical dropsonde impact

Change in forecast metric variance for hypothetical dropsonde locations. Bkgd analysis would be ‘sensitive’ to new information.

If the 24/36h ensemble forecasts were accurate, including new information at the points with the warmer colors would lead to the largest impact on the 12 UTC analysis.

Point values shown include vertical and horizontal averaging.
Does ESA provide useful guidance?

Plan to investigate ensemble sensitivity for targeted obs – did our strategy work reasonably well for targeting mesoscale observed features?

Preliminary assessment of all case events:
- All cases had convective development
- Reliance on accurate 24 h ensemble forecasts of small scale (often) weak disturbances
- Realtime metric regions were automatically generated, rarely overlapped exactly in long to shorter range forecasts
- Small variance in analysis state, rarely with statistically significant pattern sensitivity
Control forecast performance:
Case 1: 2013-05-15 – 12 UTC forecast

On ‘watch guidance’ time windows, NCAR ensemble consistently had high POD for severe storms (e.g. in North-Central TX here), but also high FAR (e.g. in KS)

Quality of timing in forecast threats at a specific point varied (examples follow)
Probability of organized convection from 12 UTC fcsts

30 member, 50 km neighborhood, 6-15 Fhr
High POD but also high FAR

Probability of updraft helicity > 75 m$^2$s$^{-2}$
Shortcomings in timing, location, dominant mode, but mesoscale forecast is still useful guidance.

Max column reflectivity Spaghetti  Fhr 6

May 15, 2013
North-Central TX tornadoes
May 31, 2013
El Reno, OK tornado,
Oklahoma City flood

Here, smaller forecast variance and some details such as the threat of flash flooding in Central OK are well forecast.

Does smaller forecast variability consistently indicate reliable forecast, or is ensemble under dispersive?

Need several seasons of forecasts

Threshold > 45 dBZ
Convective scale ensemble design

30 member ensemble, 32 days, 36 hr forecasts from 00 UTC:

**Control** – WRF-DART ensemble DA providing initial state

**Perturbed lateral boundaries** – aids dispersion/reliability later in the forecast period as LBCs (beyond 18 hrs)

**Stochastic Kinetic Energy Backscatter (SKEBS)** – represents uncertainty in the turbulent energy cascade, projects up to larger scales

**Stochastically perturbed physics tendencies (SPPT)** – represents uncertainty in the output from existing physical parameterizations
Model error representation (MER) in ensemble forecasts - preliminary

700 mb Temperature (°C)
Ensemble mean difference from control

Ensemble spread

MER improves spread, forecast reliability and even skill for light rain intensity
Mean difference (5/14-6/15) between WRF-DART analysis and downscaled GFS analysis temperature on nest domain for 12 UTC initial conditions – WRF physics related drift?

Will need to evaluate against obs

Will explore options to control drift, does MER help?
Plan of attack: MPEX dropsonde impact on forecast skill

Control: Hourly cycling from 00 UTC based on the realtime continuously cycled 6 hourly analysis. Ensemble forecasts from 16 UTC (after all drops complete for all cases).
   Includes additional conventional observations: hourly windows + GPS, mesonet, OK mesonet

Dropsondes: Same as control – but assimilate available dropsonde observations nearest in time to each hour (based on mid-time from release to reaching surface)

Verification: Forecasts against Stage IV accumulated precip, POD for severe storms (obs are difficult here), possibly GOES radiance (exploring)

Reliable skill: Are there identifiable characteristics of more skillful forecasts? Under-dispersive, so forecast variance is insufficient by itself.
Status and future work

• WRF-DART initialized ensemble forecasts with convection-permitting grid spacing provided useful guidance during MPEX of significant severe weather hazards, particularly during day 1 of the forecast - many strongly forced events

• Ensemble sensitivity analysis applied to targeted observing strategies will be further explored, reliance on accurate 24 h ensemble forecasts of small disturbances is a weakness

• We will be assimilating MPEX sondes in retrospective studies with WRF-DART with subsequent CP ensemble forecasts (data denial obs impact experiments)

• Evidence of ‘drift’ in continuously cycled WRF model analysis/forecasts, will be exploring impact of model error representation schemes to improve forecasts, perhaps also analysis system