Testing and Evaluation of GSI Hybrid Data Assimilation for Basin-scale HWRF: Lessons We Learned

Hui Shao¹, Chunhua Zhou¹, Mrinal K. Biswas¹, Ligia Bernardet²,⁴, Brian Etherton²
In collaboration with
Mingjing Tong³, John Derber³, Jeff Whitaker², and Henry Winterbottom²

¹National Centers for Atmospheric Research (NCEP), Boulder CO
²NOAA Earth System Research Laboratory (ESRL), Boulder, CO
³NCEP Environmental Modeling Center (EMC), College Park, MD
⁴Cooperative Institute for Research in the Environmental Sciences (CIRES), CU, Boulder CO

Special acknowledgement to HFIP
Mechanism for DTC Data Assimilation (DA) T&E

Operational GSI implementation and parallel test runs. **Focus on evaluating the overall performance of GSI.**

DTC real-time & retrospective GSI runs using functionally-similar operational environment: **Focus on testing incremental changes.**

- **Real-time**: “sync” testbed, uncover the issues
- **Short-term retrospective**: test individual changes, tackle the issues
- **Extensive retrospective**: impact study w/ SS, test research/developmental components

- Benchmark
- Parallel run config
- Archived data/background for retro runs

Pathway to “O”

- Benchmark
- Developmental config (suggested from the DTC)
GSI 3D-Var/Hybrid Ensemble-3DVar Cost Functions

\[
J_{3\text{DVAR}}(x') = \frac{1}{2}(x')^T B_f^{-1}(x') + \frac{1}{2}(Hx' - y')^T R^{-1}(Hx' - y')
\]

- **Fit to background**
- **Fit to observations**

\[
J_{\text{hybrid}} = \frac{\beta}{2} (x'_{\text{ens}})^T B_{\text{ens}}^{-1} (x'_{\text{ens}}) + \frac{1 - \beta}{2} (x'_{\text{background}})^T B_{\text{background}}^{-1} (x'_{\text{background}})
\]

- **\(B_f\):** (Fixed) Background error covariance (estimated offline)
- **\(H\):** Observation (forward) operator
- **\(R\):** Observation error covariance (Instrument + representativeness), where \(y^o\) are the observations
- **\(\beta\):** Weighting factor (0.25 means total \(B\) is \(\frac{3}{4}\) ensemble).

Cost function \((J)\) is minimized to find solution, \(x'\) \((x' = x^b + x')\)

(Courtesy from Jeff Whitaker, GSI Tutorial, 2012)
NCEP Dual-Res Global Coupled Hybrid

- Gets flow dependent background error covariance in 3D-Var by using an ensemble estimate.
- Ensemble perturbations are incorporated directly into cost function using extended control variable approach.

(From Daryl Kleist, AMS, 2011)
“Minimal” GSI-Hybrid System for Regional Applications

The minimal system was set up by NCEP/EMC in 2012:

- BE contributions: 25% (β) static (fixed) and 75% ensemble
- The ensemble input comes from NCEP’s Global Forecast System (GFS) ensemble.
- No feedback from deterministic analysis to ensemble analyses
- No extra computational cost due to ensemble generation

✔ Similar regional GSI-hybrid DA system has shown positive impacts in NCEP’s NAM applications.
✔ Would it be beneficial to TC forecasts? Issues? Limitations?
✔ What developmental direction should be taken for this specific scenario?
Objectives

In coordination with other regional GSI-hybrid DA development teams, the DTC

- Tests and evaluates the developmental GSI-hybrid DA model, currently the NCEP minimal regional GSI-hybrid, for Hurricane WRF (HWRF) applications.
  - Cross covariance of variables through hybrid/ensemble components
  - Sensitivity study of the weights for the static and ensemble BE statistics
  - Cycling schemes of the DA-forecast system.
- Develops user related interface.
  - Binary capability of the HWRF components
    - Reading big-endian files on little-endian platforms
- Leads the effort to make a code management plan for the GSI-hybrid code, including both variational and ensemble components.
Hurricane WRF components

HWRF Flow Diagram

- GFS Forecast
- Storm Message
- Atmos Preprocessor
- Geo-static data
- GFS Ensemble
- Vortex Improvement
- Previous 6-h forecast
- Observations
- Coupler
- Vortex Tracker
- Ocean initialization
- HWRF Ocean
- Clamotology
- HWRF Atmosphere
- Postprocessor

HWRF Components
- WRF model
- Pre-Processor (WPS)
- Vortex initialization
- Data assimilation (GSI)
- Coupler (NCEP)
- Ocean (POM-TC)
- Post-Processor (UPP)
- Vortex Tracker (GFDL)

(Courtesy from Ligia Bernadet)
Operational HWRF atmospheric config:
- Horizontal grid spacing: 27, 9, 3 km
- Inner nests move to follow storm
- Domain location vary from run to run depending on storm location
- 42 vertical levels
- Model top 50 hPa

Exp. HWRF atmospheric config:
- Horizontal grid spacing: 27 km
- No inner nests yet
- Domain is fixed
- 61 vertical levels
- Model top 2 hPa
## Experimental Design

<table>
<thead>
<tr>
<th>Cycling experiment</th>
<th>Config. ID</th>
<th>% Static BE</th>
<th>% Ens. BE</th>
<th>ICs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTL</td>
<td></td>
<td></td>
<td></td>
<td>GFS analysis</td>
</tr>
<tr>
<td>COLD</td>
<td>25%</td>
<td>75%</td>
<td></td>
<td>Cold start with GFS forecasts</td>
</tr>
<tr>
<td>CYC</td>
<td>25%</td>
<td>75%</td>
<td></td>
<td>1-day GSI cycling prior to analysis time</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BE weighting experiments</th>
<th>Var00</th>
<th>Var10</th>
<th>Var25</th>
<th>Var50</th>
<th>Var75</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Static BE</td>
<td>0</td>
<td>10</td>
<td>25</td>
<td>50</td>
<td>75</td>
</tr>
<tr>
<td>% Ens. BE</td>
<td>100</td>
<td>90</td>
<td>75</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>ICs</td>
<td>Cold start with GFS forecasts</td>
<td>Cold start with GFS forecasts</td>
<td>Cold start with GFS forecasts</td>
<td>Cold start with GFS forecasts</td>
<td>Cold start with GFS forecasts</td>
</tr>
</tbody>
</table>

Only conventional observations and TCvital were assimilated.
Testing Period

Dates: Aug 1\textsuperscript{st} 2012 – Aug 13; Aug 22\textsuperscript{nd} – Aug 28\textsuperscript{th}, 2012

Storms covered: Ernesto, Florence, Helene, Isaac, Joyce

<table>
<thead>
<tr>
<th>Type/Cat</th>
<th>Name</th>
<th>Dates</th>
<th>Max Wind (mph)</th>
<th>Min Press (mb)</th>
<th>Deaths U.S.</th>
<th>Damage</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2</td>
<td>Ernesto</td>
<td>1 – 10 Aug</td>
<td>100</td>
<td>973</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>Florence</td>
<td>3 – 6 Aug</td>
<td>60</td>
<td>1002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>Helene</td>
<td>15 – 20 Aug</td>
<td>110</td>
<td>965</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>Isaac</td>
<td>21 Aug – 1 Sep</td>
<td>80</td>
<td>965</td>
<td>34</td>
<td>$2.35B</td>
</tr>
<tr>
<td>TS</td>
<td>Joyce</td>
<td>22-24 Aug</td>
<td>40</td>
<td>1006</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Info from NWS National Hurricane Center webpage
“Minimal” GSI-hybrid Versus GFS

Aggregated track errors

Aggregated intensity abs. errors

# of cases: 84 81 79 75 71 67 64 61 58 55 50 48 46 43 41 39 35 33 31 29 28
Hurricane Isaac Tracks

2012082212
Best Track
CTL
COLD
CYC

2012082400

2012082412

2012082500
Isaac Analysis Initiated at 2012082500

1000 hPa Specific Humidity

Obs Location

COLD-CTL

CYCL-CTL
Isaac 72hr Forecasts Initiated at 2012082500

1000 hPa Specific Humidity

CTL

COLD-CTL

CYCL-CTL
Isaac Analysis Initiated at 2012082500

Tropospheric deep-layer mean (DLM) wind vector and speed (kts)
Isaac 72hr Forecast Initiated at 2012082500

Tropospheric deep-layer mean (DLM) wind vector and speed (kts)
Varying Weights of Static and Ensemble BE

*Var 25: 25% Static BE and 75% ensemble BE*
Isaac: Vertical profiles of RMSEs for q and T Analyses

Ensemble contributions degrade analysis of T at most levels and q at low levels. Similar results were found in biases and for Ernesto.

*Var 25: 25% Static BE and 75% ensemble BE
Ensemble spread 2012082900 (Isaac)
Summary and Conclusions

- DTC built and configured a testing environment similar to NCEP/EMC (Same system, script, input data, different linux machines).
- The current GSI-hybrid system for basin-scale HWRF shows minimal or no improvement on TC intensity and track compared with GFS analysis initiated forecasts.
- Cycling the minimal GSI-hybrid system shows negative impacts on TC tracks and intensity by average. These impacts might be related to limitation of the DA in ocean areas? Further study is needed.
- No significant impacts on TC track and intensity from changing the relative weights of the static and ensemble BE statistics.
- For the case study, increased weighting of ensemble BEs gives more degradation to the biases and RMSEs of T at most levels and q at lower levels. However, ensemble contributions help the q bias at higher levels. The GFS ensemble used here should be further examined for representing regional errors?
Future Work

- Regional ensemble versus global ensemble
  - Ongoing effort over NOAA/PSD on developing regional EnKF for regional ensemble update
- Radiance/Cloudy radiance DA
  - NCEP EMC added new bias correction scheme for regional radiance DA
  - Ongoing effort on cloudy radiance over NCEP/EMC and other development teams
- Moving nests/high resolution DA
- Observation impact study for TC forecasts
Pseudo-single Obs Test

$q = 1\text{g/kg at 700mb at 28.9N, 270.5E (Isaac center)}$

Analysis increments of specific humidity

Analysis increments of temperature

3DVAR  Hybrid (0.25)  Ensemble
Real Obs Test  AMSU-A radiance channels at 272E, 25.12N

background

ENS-Background

3DVAR-Background

Hybrid-Background