A Hybrid GSI/ETKF Data Assimilation Scheme for WRF/ARW

Arthur P. Mizzi[1]
Zhiquan Liu, and Xian-Yu Huang

NCAR Earth System Laboratory
Mesoscale and Microscale Meteorology Division
Boulder, CO 80307

1. Introduction

Recently, there has been much interest in hybrid variational data assimilation research. *See e.g.*, Wang 2010, Bowler et al. 2009, and Wang et al. 2008a, b. This paper continues that research by extending the work of Wang et al. 2008a, b to a regional version of the NOAA Gridpoint Statistical Interpolation (GSI) global hybrid assimilation scheme described by Wang 2010. For this research, the NOAA Environmental Modeling Center (EMC) developed the GSI global hybrid code in collaboration with the University of Oklahoma. We modified that code to incorporate an ETKF interface and apply the GSI/ETKF hybrid in a regional setting with WRF-ARW.

2. The GSI/ETKF hybrid scheme

The GSI/ETKF regional hybrid cycling algorithm follows that of Wang et al. 2007. We start with an ensemble of background forecasts. The assimilation cycle begins by calculating the ensemble mean and perturbations. It proceeds by updating the mean with the GSI/ETKF regional hybrid and updating the perturbations with the ETKF. The mathematical formalism is identical to that of Wang et al. 2008a. The ETKF uses the method described by Bishop et al. 2001 with the inflation factor strategy described by Wang et al. 2007. For the ETKF, we use an averaging period of 10 days.

After updating the ensemble mean and perturbations, we reconstitute the total fields for each ensemble member, update the boundary conditions, and use WRF/ARW to make 12-hr forecasts for each ensemble member. Those forecasts become the background forecasts for the next cycle, and the assimilation process is repeated.

In the GSI/ETKF hybrid, the weighting between the variational increment and the ensemble increment is controlled by ß. When ß=1.0, we get the variational increment. When ß=0.0, we get the ensemble increment. When ß is between 0.0 and 1.0, we get the hybrid increment. Additional GSI/ETKF hybrid parameters are: (i) H – the horizontal localization length scale in km , (ii) V – the vertical localization length scale in number of vertical gridpoints, and (iii) N – the number of ensemble members. As the localization scales decrease, the radius of influence for a particular observation decreases, reducing the amount of noise in the increment. For the experiments discussed in this paper, we generally set ß=0.5, H=1,500 km, V=20 grid points, and N=20 ensemble members. The

---

[1] Corresponding author: Arthur P. Mizzi, NCAR/MMM P.O. Box 3000, Boulder, CO., mizzi@ucar.edu, 303-903-5544.

control experiment uses those parameter settings with the Wang et al. 2007 inflation algorithm and is denoted WG07.

3. Experimental design

The WRF model is described by Skamrock et al. 2005. We ran WRF on the CONUS domain, which covers North America and the surrounding oceans. We used a coarse resolution of 200 km on a 45 x 45 horizontal grid with 27 vertical levels. The model top was located at 50 hPa. This is the same configuration as was used by Wang et al. 2008a, b. The coarse CONUS domain was chosen for similar reasons as identified by Wang et al. 2008a, b i.e., to facilitate running a large number of experiments. The same cautions about applying our results to higher resolution experiments apply. We plan to extend the studies reported in this paper to higher resolution.

For the experiments reported here, WRF was initialized with GFS analyses and the GFS forecasts were used for the boundary conditions. We starting the assimilation experiments at 12Z on August 15, 2007 and cycled for 10 days until 12Z on August 25, 2007. During that period, Hurricane Dean formed within the WRF domain at 52° W, 12.5° N on August 15, 2007 and moved across the domain in a WNW direction. On August 19, 2007, Hurricane Dean made landfall at 97° W, 21° N in Mexico.

In our experiments, we used 12h-cycling. For any particular cycle, the GSI/ETKF hybrid used all conventional observations obtained from the NCEP operational datasets. The ETKF used only radiosonde wind, temperature, and moisture observations. The ETKF also excluded observations when the relative observation error (defined as observation error divided by the observation) exceeded 20 because the Wang et al. 2007 inflation algorithm behaved poorly when the innovations were relatively small and the observation error were relatively large.[2] Our experiments showed that the inflation algorithm behaved properly when the maximum relative observation error cutoff was set to 20. Similarly, we used a minimum relative observation cutoff of .0001 to exclude observations with near-zero observations error. We did not investigate the sensitivity of our results to those upper and lower relative error bounds.

For cycling, the initial ensemble was generated by adding random perturbations drawn from a normal distribution having the same distribution as the WRFDA background error covariance to the August 15, 2007 00Z analysis and boundary conditions. That initial ensemble was then used to obtain an ensemble of WRF 12-h forecasts. That forecast ensemble became the background forecasts for starting the assimilation cycling on August 15, 2007 at 12Z.

For verification, we used all radiosonde wind, temperature, and moisture observations and surface synoptic observations that pass the WRFDA quality control procedure based on the non-ensemble GSI background forecasts. Our verification began on August 17,

[2] When that happens, the sum of the ratio of those terms can make the inflation factor negative.

2007 at 12Z after two days of cycling. We verified our results on the entire horizontal and vertical domain.

4. Single observation experiments

We performed many single observation experiments to confirm the conceptual accuracy of the GSI/ETKF regional hybrid. For those experiments, we placed a single temperature observation of 1°K at ~500 hPa at the horizontal location indicated by the black dot in Figure 1. Due to space limitations, we present only a subset of those experiments.[3]

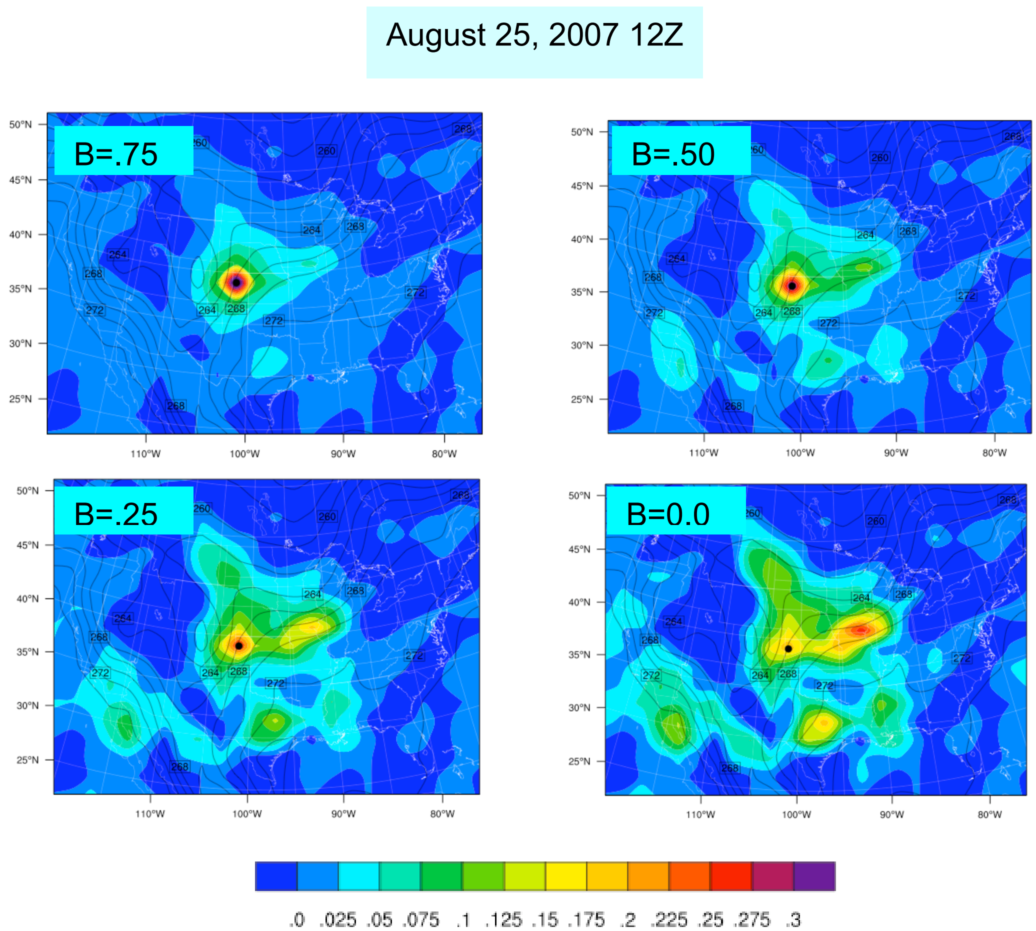

Figure 1. Single observation experiment weighting factor sensitivity. The solid black lines are contours of geopotential height in 10m. The shaded contours are the temperature increment in °C. ß represents the weighting between the variational increment (ß=1.0) and the ensemble increment (ß=0.0).

---

[3] In a single observation experiment for the initial ensemble, as the ensemble size increases, the ensemble analysis increment should approach the variational increment. However, here we generated the initial ensemble with the WRFDA background error covariance as opposed to the GSI error covariance. That means that as the ensemble size increases, the ensemble increment will approach WRFDA variational increment and be slightly different from the GSI variational increment. Although not shown, we have verified that as the ensemble size increases, the ensemble increment approaches the variational increment.

Figure 1 shows the single observation increment sensitivity to varying ß between 0.0 and 1.0. The background forecasts were taken from August 25, 2007 at 12Z, the last day of the cycling experiment, to demonstrate the increment's flow dependence. Figure 1 shows that for ß=0.0 the ensemble increment had horizontal structure and was stretched along the geopotential contours. As ß approaches 1.0, the horizontal structure lessened, and the increment more closely resembled the stretched variational increment. Although not shown, we have confirmed that similar stretching occurred in the vertical.

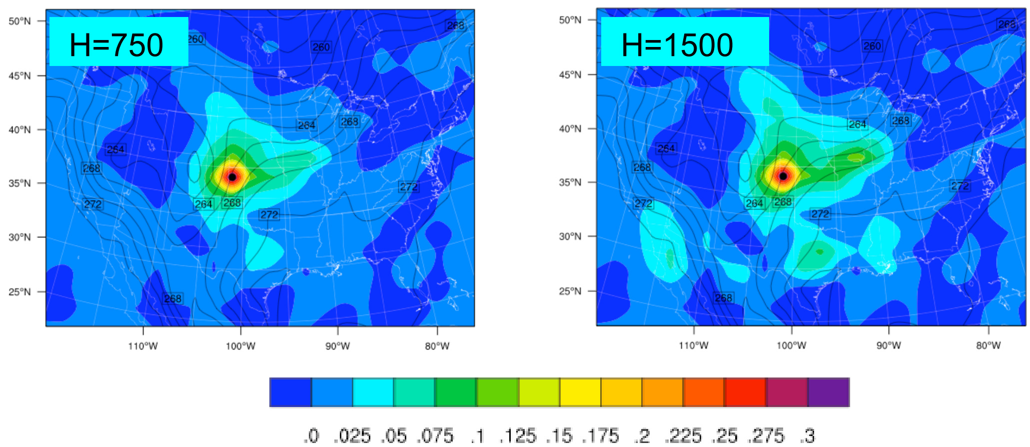

Figure 2. Single observation experiment horizontal localization sensitivity. The solid black lines are contours of geopotential height in 10m. The shaded contours are the temperature increment in °C. H represents the horizontal localization length scale in km.

Figure 2 displays the single observation increment sensitivity to changes in H, the horizontal localization length scale. It shows a reduction in the lateral fine structure as H decreased from 1,500 km to 750 km. When H was increased to from 1,500 km to 4,500 km with an increment of 1,500 km (not shown), there was no noticeable change in the lateral fine structure. When V, the vertical localization length scale was varied from 10 grid points to 40 grid points with an increment of 10 (not shown), we found similar variations.

5. Sensitivity experiments to determine the optimal ß and H

In this section, we study the optimal ß and H configuration for the GSI/ETKF regional hybrid. We ran 10-day cycling experiments for each configuration with ß set to 0.0, 0.25, 0.50, 0.75, and 1.0, and H set to 750 km, 1,500km, 3,000km, and 4,500 km for a total of 20 experiments (individual experimental results not shown). We found that: (i) for non-surface u, v, T, and q, the optimal configuration was ß=0.75 and H=750 km, (ii) for near-surface u, the optimal configuration was ß=1 and H=NA, (iii) for near-surface v, the optimal configuration was ß=.75 and H=3,000 km, (iv) for near-surface T, the optimal configuration was ß=.5 and H=750 km, and (v) for near-surface q, the optimal configuration was ß=.5 and H=1,500.

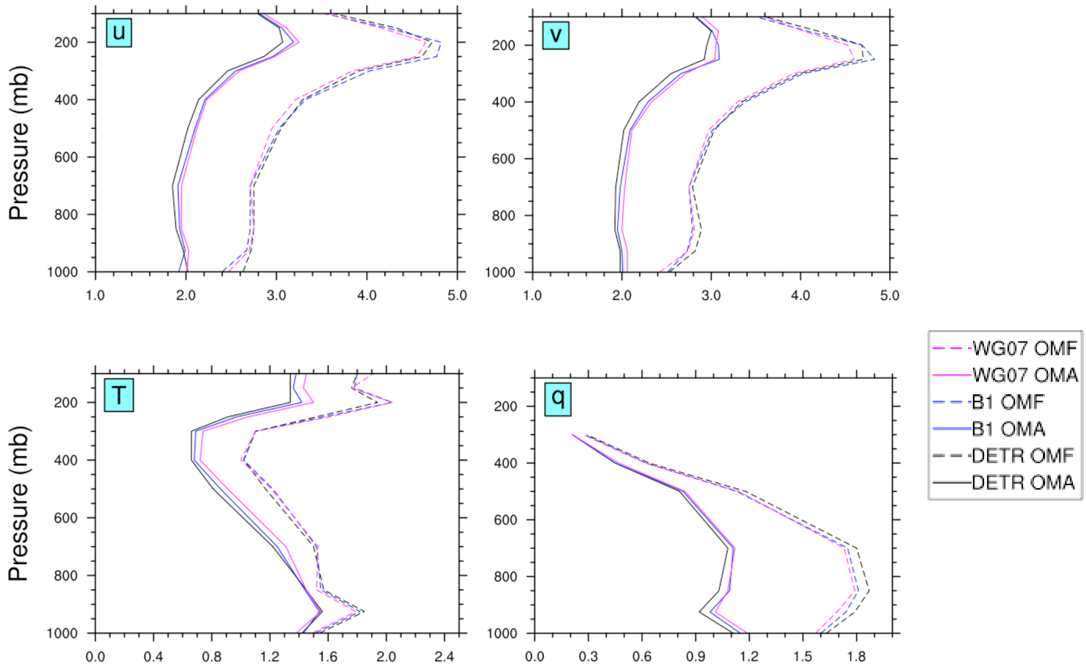

Figure 3. Vertical profiles of analysis and forecast RMSE for the control (WG07), ensemble/non-hybrid GSI (B1), and the non-ensemble GSI (DETR). The solid lines display the analysis RMSE (denoted OMA), and the dashed lines display the forecast RMSE (denoted OMF).

6. Analysis and forecast verification

In this section, we compare the analysis and forecast root mean square error (RMSE) from WG07, the ensemble/non-hybrid GSI (B1), and the non-ensemble GSI (DETR) experiments.

Figure 3 shows vertical profiles of analysis and forecast RMSE for WG07, B1, and DETR. The solid lines represent the analysis RMSE, and the dashed lines represent the forecast RMSE. Figure 3 shows that the DETR analyses fit the observations better than the B1 and WG07 analyses. That result is consistent with the findings previous investigators. Figure 3 also shows that the WG07 forecasts fit the observations better than the B1 and DETR forecasts. The fit was better for: (i) u and v in the middle and upper troposphere, (ii) T in the lower troposphere, and (iii) q in the middle and lower troposphere.

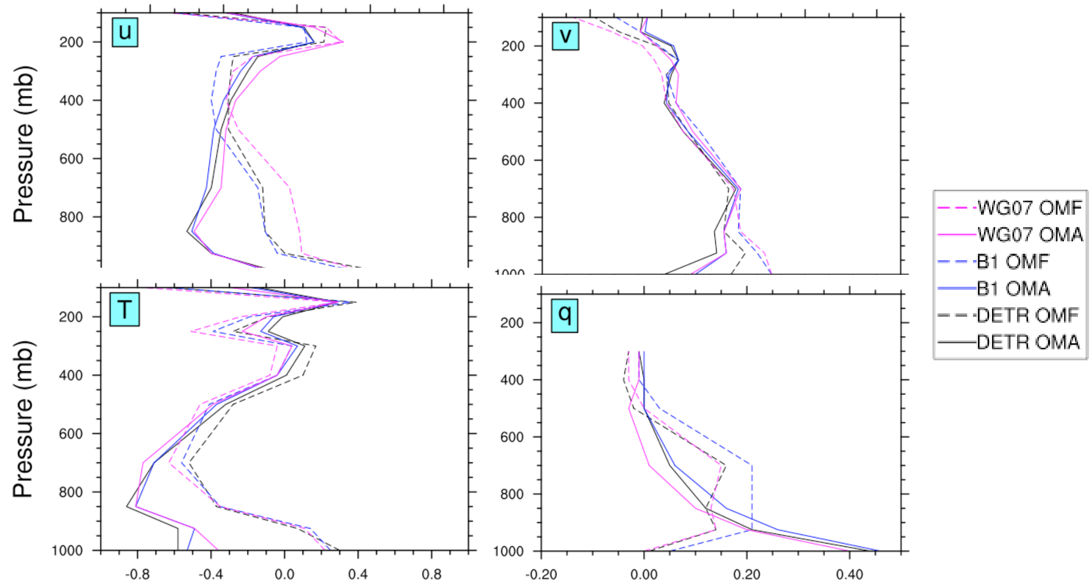

Figure 4. Vertical profiles of analysis and forecast bias for WG07, B1, and DETR. The solid lines display the analysis bias (denoted OMA), and the dashed lines display the forecast bias (denoted OMF).

Figure 4 is the same as Figure 3 except it displays vertical profiles of the bias. Generally, the WG07 analysis bias was less than the B1 and DETR analysis biases. *See*: (i) u in the middle and upper troposphere, (ii) T in the lower and upper troposphere, and (iii) q in the middle and lower troposphere. In Figure 4, it is difficult to characterize the forecast biases because they varied with the meteorological variables.

7. Sensitivity to inflation factor scheme

In this section, we investigate the sensitivity of RMSE to the ETKF inflation factor scheme. We consider four algorithms: (i) WG03 – the Wang and Bishop 2003 scheme which used innovation averaging, (ii) WG07 – the control experiment which used the Wang et al. 2007 scheme with innovation and projection factor averaging, (iii) BW08 – the Bowler et al. 2008 scheme which used the ratio of the spread at the previous and current cycle, and (iv) TRNK – an experimental scheme that used the Wang and Bishop 2003 scheme but averaged the final inflation factor as opposed to the innovations. We ran 10-day cycling experiments with those inflation schemes for the Hurricane Dean case. All parameters were the same as in WG07 except for the inflation factor scheme. The results from the inflation factor scheme sensitivity experiments are displayed in the left panel of Figure 5.

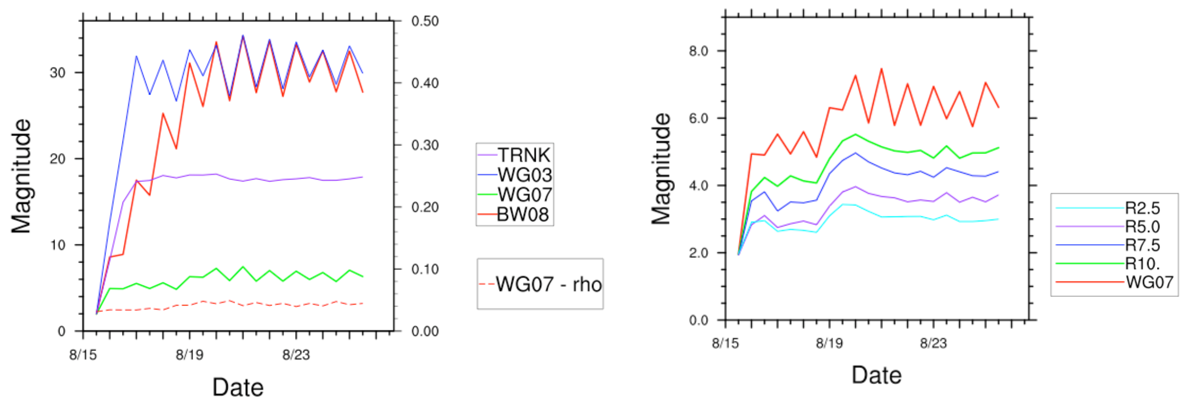

Figure 5. Left Panel – inflation factor time series for the WG03, WG07, BW08, and TRNK schemes with magnitude displayed on the left ordinate axis. WG07-rho denotes the Wang et al., 2007 project factor time series with magnitude displayed on the right ordinate axis. Right Panel – inflation factor time series for the WG07 and R( ) experiments.

The left panel of Figure 5 shows that WG03, WG07 and TRNK became stable after four cycles. BW08 became stable after eight cycles. WG03 and BW08 had the largest magnitudes, TRNK had moderate magnitudes, and WG07 had the smallest magnitudes. WG07 had the smallest inflation factors because it used an adaptive projection factor (WG07-rho in the left panel of Figure 5) that corrected for errors in the eigenvector projections. The results for all schemes, except TRNK, had bi-cycle oscillations that have been observed by other researchers and attributed to over/under inflation. *See* Bowler et al. 2008. Our results show that those oscillations were due to the number of ETKF observations.

The right panel of Figure 5 displays inflation factor time series for experiments that held to number of observations entering the ETKF (ETKF observations) constant from one cycle to the next. For those experiments: R2.5 denotes 2,500 randomly selected ETKF observations per cycle, R5.0 denotes 5,000, R7.5 denotes 7,500, and R10 denotes 10,000. For the R( ) experiments, the ETKF observations were selected randomly from the set of between 15,000 and 18,000 observations used by the WG07 ETKF. The right panel of Figure 5 shows that when the number of ETKF observations was held constant, the bi-cycle oscillation disappears.

The bi-cycle oscillation from Figure 5 is also present in ensemble spread time series. *See* Figure 6. The upper panels of Figure 6 show the analysis ensemble spread, i.e., the ETKF ensemble spread, for the different inflation schemes when the number of ETKF observations was not held constant. The lower panels of Figure 6 show the analysis ensemble spread for WG07 when the number of ETKF observations was held constant. For the upper panels, at 00Z the number of ETKF observations was relatively small O(15,000) and resulting analysis ensemble spread was relatively large (number of observations not shown). At 12Z, the number of ETKF observations was relatively large O(17,000) and the analysis ensemble spread is relatively small (number of observations not shown). The lower panels show that the oscillation disappears when the number of ETKF observations was held constant.

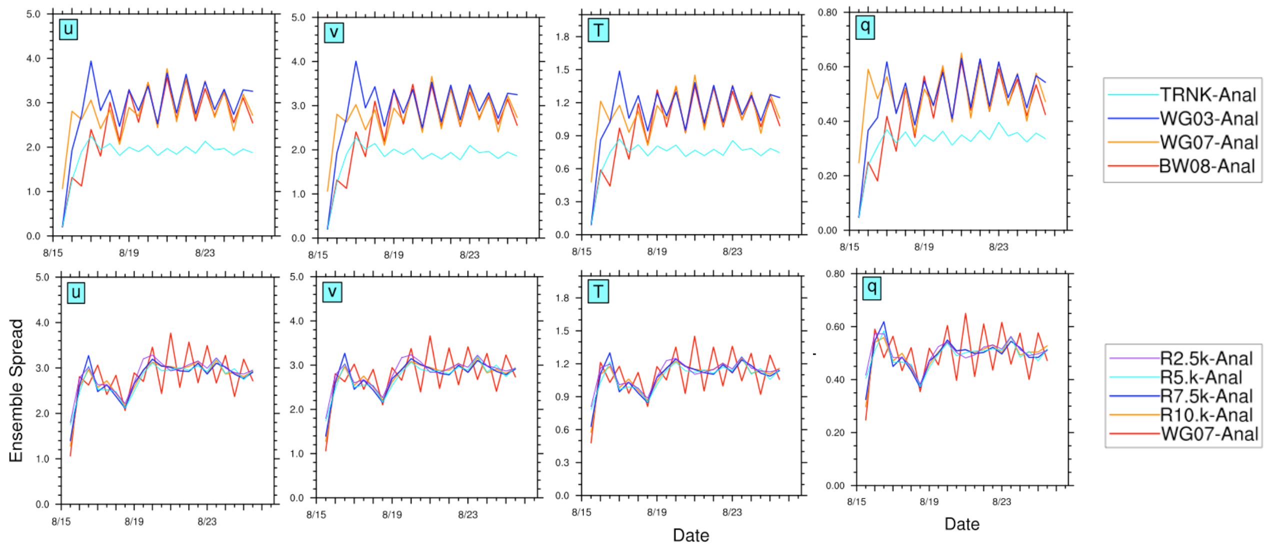

Figure 6. Upper Panels – ensemble spread time series for u, v, T, and q from the WG03, WG07, BW08, and TRNK schemes. Lower Panels – ensemble spread time series for the WG07 and R( ) experiments.

We also ran verifications on the R( ) experiments to determine the sensitivity of analysis and forecast RMSE to varying the number of ETKF observations. Those results are shown in Table 1. The yellow shading identifies the configuration with the lowest forecast RMSE for each meteorological variable. Table 1 shows that reducing the

UPR	R 2.5	R 5.0	R 7.5	R 10.	WG07
u (m/s)	3.234 (2.323)	3.232 (2.326)	3.230 (2.325)	3.240 (2.324)	3.232 (2.327)
v (m/s)	3.250 (2.363)	3.249 (2.364)	3.251 (2.365)	3.254 (2.362)	3.257 (2.369)
T (K)	1.487 (1.210)	1.484 (1.208)	1.485 (1.207)	1.485 (1.214)	1.491 (1.211)
q (g/kg)	1.297 (0.867)	1.298 (0.872)	1.299 (0.866)	1.302 (0.869)	1.305 (0.870)

Table 1. RMSE for the WG07 and R( ) experiments for non-surface meteorological variables. The forecast RMSE is outside the parentheses, and the analysis RMSE is inside the parentheses. The yellow shading identifies the lowest forecast RMSE for each meteorological variable.

number of ETKF observations improved the analysis and forecast RMSE. Table 1 also shows that the optimal number of ETKF observations ranged between 2,500 and 7,500 and was variable dependent.

The RMSE changes between the different experiments in Table 1 are small and may not be significant. We compared those changes with RMSE changes obtained by increasing the ensemble size from 20 to 80 in increments of 20 (not shown) and found that the changes in Table 1 were comparable. Similarly, we compared the RMSE magnitudes and changes in Table 1 to those in Table 2 from Wang et al. 2008b and found that they were comparable. Thus, we conclude that the results in Table 1 are meaningful.

Next, we review verification of the inflation factor scheme sensitivity experiments. Those results are displayed in Table 2 for the non-surface meteorological variables. We performed a similar verification for the near-surface variables, but those results are not presented. Table 2 shows that: (i) WG07 was the optimal inflation factor scheme based on the u, v, and q forecast RMSE and (ii) TRNK was the optimal scheme based on the T forecast RMSE. For the near-surface variables, we found similar results: (i) WG07 was optimal based on the u, T, and q forecasts RMSE, (ii) BW08 was optimal based on the v forecast RMSE, and (iii) DETR was optimal based on the surface pressure, Ps forecast RMSE.

UPR	**DETR**	**TRNK**	**WG03**	**WG07**	**BW08**
U (m/s)	3.294 (2.293)	3.245 (2.341)	3.259 (2.330)	3.232 (2.327)	3.256 (2.327)
V (m/s)	3.320 (2.267)	3.259 (2.376)	3.286 (2.370)	3.257 (2.369)	3.272 (2.368)
T (K)	1.487 (1.142)	1.473 (1.201)	1.492 (1.200)	1.491 (1.211)	1.488 (1.200)
q (g/kg)	1.361 (0.830)	1.308 (0.870)	1.311 (0.871)	1.305 (0.870)	1.308 (0.871)

Table 2. RMSE for the inflation factor sensitivity experiments for non-surface meteorological variables. The forecast RMSE is outside the parentheses, and the analysis RMSE is inside the parentheses. The yellow shading identifies the lowest forecast RMSE for each meteorological variable.

8. Spread verification of the inflation factor schemes

One measure of whether a perturbation generation strategy works properly in an ensemble prediction system is the degree to which the resulting ensemble spread predicts the background forecast errors for the next cycle. *See* Wang et al. 2008a, b. Wang et al. 2008a. b and Wang et al. 2007 calculate the background forecast error as the RMS innovation minus the observation error. *See* Eq. 1, Wang et al. 2008b. Figure 7 displays vertical profiles of the background forecast error and ensemble spread. Ideally, those

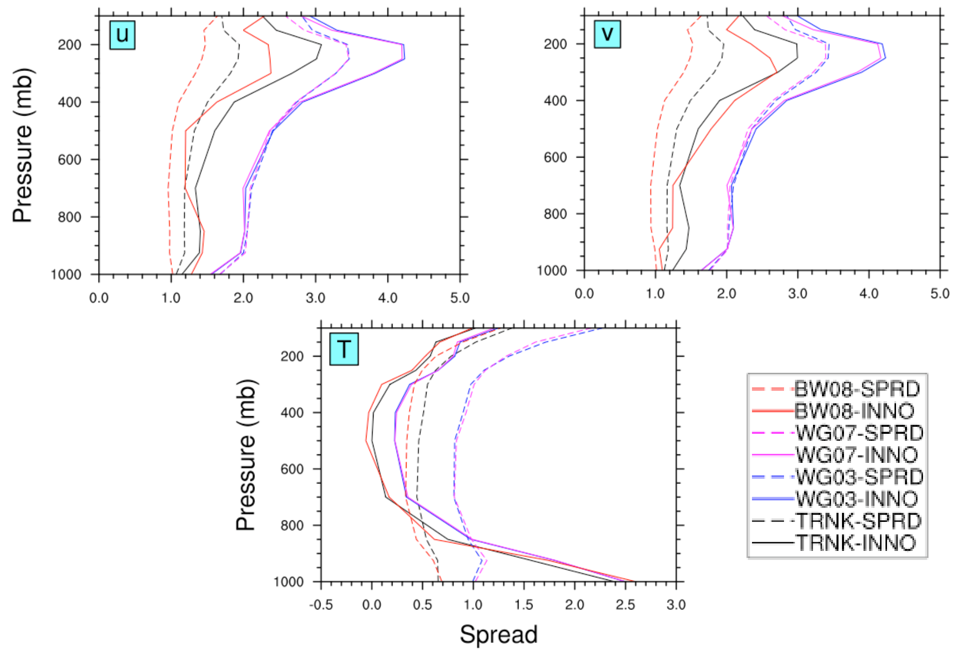

Figure 7. Vertical profiles of background forecast error (denoted INNO) and the ETKF ensemble spread (denoted SPRD) for the inflation factor scheme sensitivity experiments. The background forecast error is defined as the RMS innovations minus the observation errors. Experiment identifiers are as defined earlier.

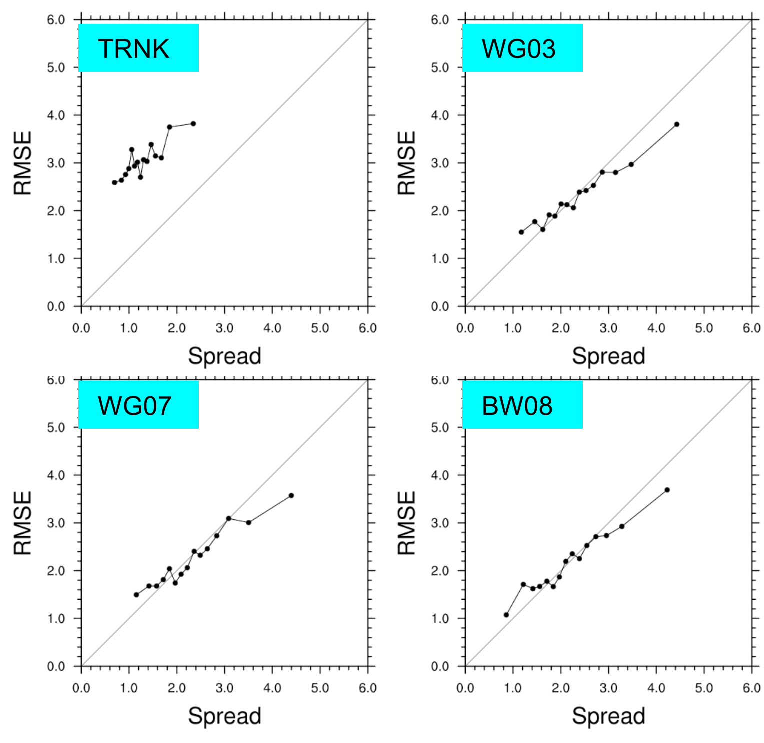

Figure 8. Plots of background forecast error (denoted INNO) and the ETKF ensemble spread (denoted SPRD) for u at 500 hPa from the inflation factor scheme sensitivity experiments. Background forecast error is as defined in Figure 7. Experiment identifiers are as defined earlier.

profiles should be coincident. Figure 7 shows results for u, v, and T. It does not include results for q because the associated relative observation errors fell below the minimum relative observation error cutoff. That was an inadvertent exclusion that is not thought to have had a material impact of our results.

Figure 7 shows that for u and v the WG03 and WG07 background forecast error profiles were nearly coincident with the ETKF ensemble spread profiles except in the upper troposphere. The BW08 and TRNK profiles had a gap between the profiles throughout the troposphere. Figure 7 also shows that the T profiles had large gaps in the middle and lower troposphere. Wang et al. 2008b found similar result for T and suggested it was associated with errors in WRF's boundary and surface parameterizations.

Next we consider whether the ETKF spread can distinguish between large and small background forecast errors. We use the method of Majumdar et al. 2001 as applied by Wang et al. 2008a and Wang et al. 2007. We collected u innovation and ensemble spread data from observation locations at 500 hPa and processed it according to the procedure of Wang et al. 2008a. We used 15 bins and had 130 data points per bin. The results are presented in Figure 8 with ETKF ensemble spread along the abscissa and background forecast error along the ordinate. Ideally, the plotted points should fall along the 45º line that is included for reference.

Figure 8 shows that for TRNK all points fell above the 45º line. We suspect that TRNK's final inflation factor averaging caused the ensemble spread to systematically underestimate the background errors i.e., TRNK systematically underestimated the spread. Data from WG03, WG07, and BW07 generally fell along the 45º line. For those schemes, the ensemble spread distinguished between large and small background forecast errors.

9. WG07 and DETR 12-h forecast differences valid at August 22, 2007 12Z

Finally, we look at differences between the WG07 and DETR 12-h forecasts for a time when the WG07 u and v forecasts fit the observations better than the DETR forecasts. The forecast differences are plotting in Figure 10 for the 700 hPa and 300 hPa levels. In Figure 10, geopotential height is plotted in m with the blue contours, wind speed is plotted in m/s with the shaded contours, and the wind vectors are in m/s.

The largest differences are in eastern Pacific equatorial region where DETR over-predicted the easterlies. DETR also over-predicted the 700 hPa anti-cyclonic circulation, 300 hPa cyclonic circulation over southern Baja and western Mexico. As discussed by Wang et al. 2008b, such anomalous circulations can have detrimental effects of the precipitation and moisture analyses and forecasts. We have not investigated the anomalous circulation impacts for this paper.

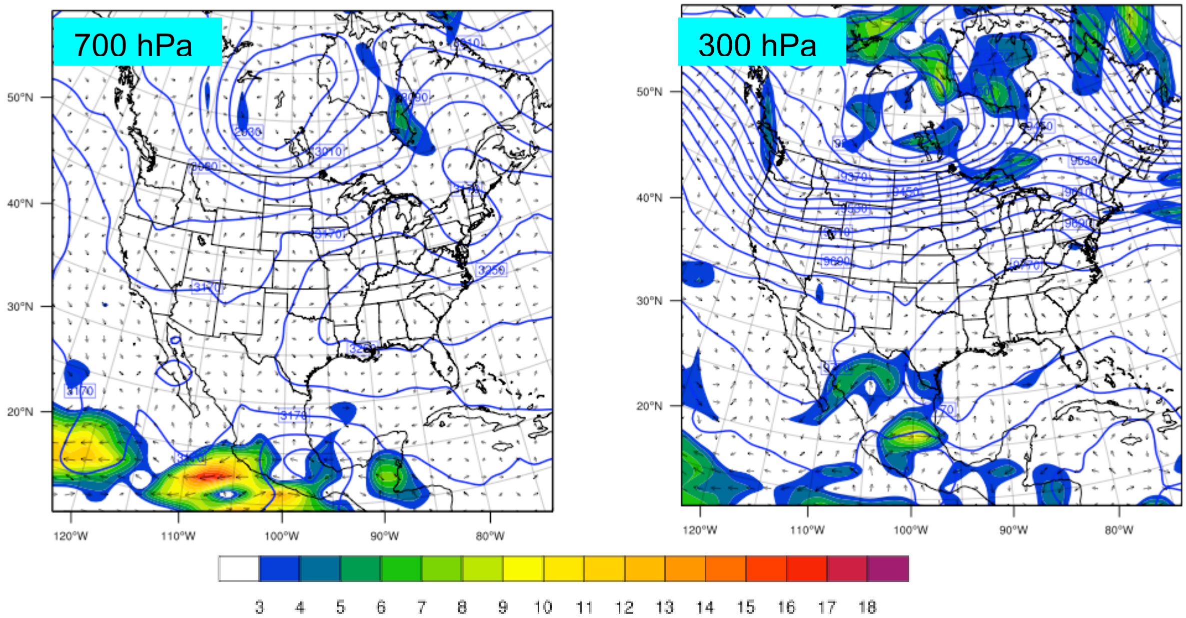

Figure 9. DETR minus WG07 12-h forecast differences valid at August 22, 2007 12Z. The blue contours represent geopotential height in m, shaded contours represent wind speed in m/s, and the wind vectors are in m/s.

10. Discussion and conclusions

In this paper, we presented results from testing the GSI/ETKF regional hybrid and comparing various ETKF inflation factor algorithms. As part that work, we conducted single observation and hybrid parameter sensitivity experiments.

Our results showed that for a single T observation of 1°C located at 500 hPa in a geoptential gradient, the GSI/ETKF regional hybrid behaved as expected. Specifically, the ensemble increment (ß=0) had horizontal structure and was stretched along the geopotential contours. The variational increment (ß=1) was not shown but was confirmed to be symmetric. The hybrid increment (0.0<ß<1) showed that as ß approached 1.0 from 0.0, the horizontal structure lessened and the increment tended toward a stretched version of the variational increment. Although not shown, similar behavior occurred in the vertical.

We also examined the single observation increment sensitivity to changes in H (shown) and V (not shown) and found that as the localization scales were reduced far-field fine structure was reduced. In summary, the single observation experiment suggested that the GSI/ETKF regional hybrid is working properly.

Next, we conducted hybrid sensitivity experiments to determine the optimal values of ß and H. For non-surface meteorological variables (u, v, T, and q) the optimal configuration was ß=0.75 and H=750. For near-surface variables (u, v, T, q, and Ps) was variable dependent. ß ranged between 0.5 and 1.0, and H ranged between 750 km and 3,000 km. We suspect that those results are application dependent. We used coarse resolution and a subset of conventional observations (only the radiosonde u, v, T, and q observations) as ETKF observations. If the resolution were increased, the optimal values of ß and H would likely change, so caution should be used when applying these results to other resolutions. The optimal ß and H configurations would likely be unaffected by changes in the types of ETKF observations.

We compared the analysis and forecast RMSE for the GSI/ETKF regional hybrid (WG07) with the ensemble/non-hybrid GSI (B1) and ensemble GSI (DETR) RMSEs. We found that WG07 generally had higher RMSE for the analyses and lower RMSE for the forecasts. Those results were consistent with the results of Wang et al. 2008b for WRFDA with real data. We also compared the WG07 bias with the B1 and DETR biases. We found that the WG07 analysis bias was generally less, but the forecast bias was difficult to characterize because the comparison results varied with meteorological variable.

We also conducted an ETKF inflation factor sensitivity study to determine whether there is an optimal inflation algorithm. We compared the WG03, WG07, BW07, and TRNK schemes and found that WG07 was preferred. (Our results showed that: (i) WG07 was optimal for u, v, and q, and (ii) TRNK was optimal for T. However, TRNK did not perform well in the spread verification analysis. Thus, we conclude that WG07 is the optimal inflation algorithm.)

The inflation factor sensitivity experiments showed a bi-cycle oscillation in the inflation factor and in the ETKF ensemble spread. We showed that the oscillation was due to variations in the number of ETKF observations at the different cycle times, i.e, at 00Z there were O(15,000) observations with low inflation and high spread, and at 12Z there were O(17,000) observations with high inflation and low spread. To address that issue, we conducted as series of experiments that held the number of ETKF observations constant between cycles, identified as the R( ) experiments. Results from those experiments showed that holding the number of ETKF observations constant solved the oscillation issue without adversely impacting the ETKF ensemble spread. The spread was unaffected because the inflation factor compensated for the variations in the initial spread estimate and the final spread estimate properly attained the target.

The R( ) experiments also showed that reducing the number of ETKF observations improved the WG07 analysis and forecast RMSE. Based on the forecast RMSE, the optimal number of ETKF observations ranged between 7,500 and 2,500 depending on the meteorological variable. We suspect that reducing the number of ETKF observations improved the forecast RMSE for the following reason. The ETKF has a disparity between the ensemble size and the number of ETKF observations. Due to that disparity, the initial ETKF estimate of the ensemble spread is unrepresentative. The inflation factor attempts to correct for that unrepresentativeness. If the initial estimate of the spread is too low, then reducing the number of ETKF observations elevates that initial estimate putting it closer to the correct spread before inflation. Since the horizontal structures of the different initial estimates are different, the inflation factor cannot make this type of correction without localization. We also suspect that these results are application dependent. They would likely vary if the resolution were increased. Also, they would likely vary if the types of ETKF observations were changed. If separate inflation factors were used for each meteorological variable, we interpret these results as indicating that the optimal number of ETKF observations would be different for wind (u,v), T, and q. That suggests that the ETKF and inflation factor should be applied separately to each variable.

Finally, we preformed spread verification on the WG03, WG07, BW08, and TRNK inflation schemes and found: (i) for vertical profiles of background forecast error and ETKF ensemble spread, the results from all inflation algorithms resembled the results from Wang et al. 2008a, b, (ii) TRNK's final inflation factor averaging caused the ETKF ensemble spread to systematically under-predict the background forecast errors i.e., the inflation algorithm systematically underestimated the ensemble spread, and (iii) WG03, WG07, and BW08 were able to distinguish between large and small background forecast errors so their associated ensemble spread properly predicted the background forecast errors.

11. References

Bowler, N.E., A. Arribas, K.R. Mylne, K.B. Robertson, and S.E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Quart. J. Meteor. Soc.*, **134**, 703-722.

Majumdar, S.J., C.H. Bishop, B.J. Etherton, I. Szunyogh, and Z. Toth, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast error variance produced by targeted observations? *Quart. J. Meteor. Soc.*, **127**, 2803-2820.

Skamrock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, W. Wang, and J.G. Powers, 2005: A description of the Advanced Research WRF Version 2. NCAR Tech. Note 468+STR, National Center for Atmospheric Research, Boulder. CO 88 pp. [Available from UCAR Communications, P.O. Box 3000, Boulder, CO 80307].

Wang, X., 2010: Incorporating ensemble covariance in the gridpoint statistical interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990-2995.

Wang, X., and C.H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecasts schemes. *J. Atmos. Sci.*, **60**, 1140-1158.

Wang, X., T.M. Hamill, J.S. Whitaker, and C.H. Bishop, 2007: A comparison of hybrid ensemble transform Kalman filter-OI and ensemble square-root filter analysis schemes. *Mon. Wea. Rev.*, **135**, 1055-1076.

Wang, X., D.M. Barker, C. Snyder, and T.M. Hamill, 2008a: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model, Part I: Observing system simulation experiment. *Mon. Wea. Rev*., **136**, 5116-5131.

Wang, X., D.M. Barker, C. Snyder, and T.M. Hamill, 2008b: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model, Part II: Real observation experiment. *Mon. Wea. Rev*., **136**, 5132-5147.