Impact of WRF-Var (3DVar) Background Error Statistics on Typhoon analysis and Forecast

Y.-R. Guo[1], H.-C. Lin[1], X. X. Ma[1], X.-Y. Huang[1], C.T. Terng[2], and Y.-H. Kuo[1]

[1] National Center for Atmospheric Research
P. O. Box 3000, Boulder, CO 80307

[2] Central Weather Bureau, Taipei, Taiwan

Abstract

The Background Error Statistics (BES) is a key component of the WRF-Var (3DVar) system. The climatological BES generated using one month or more forecast dataset with NMC-method may not be suitable for the analysis of a Typhoon case, such as Typhoon Haitang that occurred over the west Pacific from 14 to 20 July 2005, especially for typhoon bogus data assimilation. A series of cycling data assimilation experiments are conducted for Typhoon Haitang. We found that assimilation of typhoon bogus data with the tuned BES could significantly improve the typhoon analysis in terms of the storm position and intensity. The technique using multiple external loops with different tuning factors for BES improved analysis of the large-scale synoptic circulation, and the typhoon position and intensity, and as a result, the skill of the Typhoon track forecast is improved.

1. Introduction

For the analysis of tropical storms, hurricanes or typhoons, it is usually necessary to perform bogus data assimilation (BDA) with the variational techniques, 3DVar or 4Dvar (Zou and Xiao, 2000). However, the bogus data cannot *always* be assimilated properly with WRF-Var (a version of the 3DVar code developed by NCAR special for WRF-ARW model, Skamarock *et al.* 2005). With the MM5-3Dvar system (Barker *et al.* 2004), Guo et al. (2005) tuned the background error statistics and improved the forecast of a convective case. The success of the assimilation of typhoon bogus data may also depend on the BES and the storm location in the first guess (FG),…etc. Typhoon Haitang occurred over the west Pacific from 14 to 20 July 2005 is a super typhoon with the deepest central pressure of 912 hPa. To make the correct analysis of the initial conditions for forecast model and then improving the forecast are challenging problems for Typhoon Haitang.

In order to find a suitable technique with WRF-Var for Typhoon Haitang data assimilation, a series of analysis and forecast experiments were conducted. Using the NCEP GFS analysis as the first guess (Cold-start experiments), WRF-Var gave a better analysis of the Typhoon location (the storm re-location technique might already be applied in NCEP GFS analysis), but the intensity was too weak. With the cycling mode (Warm-start) experiments, the analysis and forecast of the Typhoon intensities were much better than those from the Cold-start experiments, but the poor analysis and forecast of Typhoon locations were obtained if the default, climatological BES were used in WRF-Var. In the cycling mode experiments, the technique using multiple external loops with different tuning factors to CV5 BES in WRF-Var improved the analysis of both the large-scale synoptic flow and the Typhoon position and intensity in the initial condition, and as a result, the skill of the Typhoon track forecast was improved.

2. Typhoon Haitang case

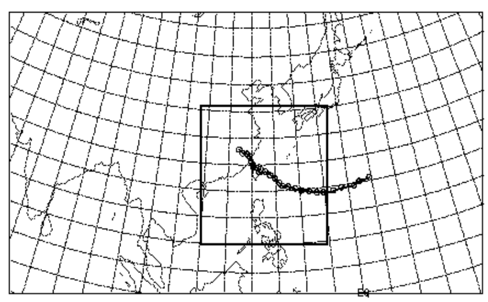

Fig. 1. Experiment (45/15km) domains and the Typhoon Haitang track from 0000 UTC 14 to 0600 UTC 20 July 2005.

On 11 July 2005, a tropical cyclone formed over Pacific Ocean around 22°N, 157°E. At 1800 UTC 13 July, it intensified to be a typhoon with the central pressure of 970 hPa and maximum wind of 33 m/s. It was named as Haitang and moved westward along the 20° latitude. At 1800 UTC 16 July, the central pressure dropped to 912 hPa and the maximum wind reached 55 m/s. The typhoon landed at the east coast of middle Taiwan at about 0000 UTC 18 July. Then, it moved across the Taiwan island, and made another landfall at mainland China at noon 19 July. It decayed and disappeared at 0600 UTC 20 July. Figure 1 showed the experiment domains with the Typhoon Haitang track from 0000 UTC 14 to 0600 UTC 20 July.

3. Analysis experiments

The current version of WRF-Var (3DVar) has the capability of assimilating the TC (tropical cyclone) bogus sounding data. To assess the performance of bogus data assimilation with WRF-Var, a series of analysis experiments were carried out. In these experiments, the first guess (FG) is a 6-h forecast initiated at 1800 UTC 16 July from NCEP GFS analysis (Fig.2a) and only the bogus sounding data: SLP (Sea Level Pressure) and winds at mandatory levels from 1000 to 400 hPa from the Central Weather Bureau (CWB), Taiwan, are assimilated (Fig. 2b). In this FG, the Typhoon was located about 5 grid-points (80-km) west of the observed position in the 15km domain2 (Table 1). There are total of 64 bogus soundings (40 TC bogus and 24 global bogus) available.

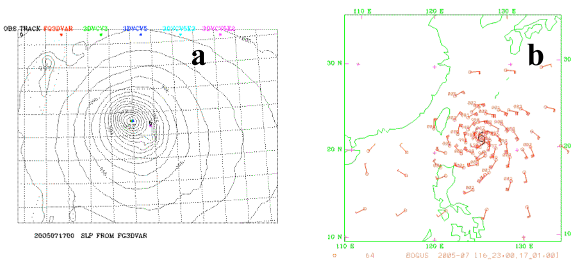

Fig. 2. a) First guess SLP over a sub-domain, and b) the bogus SLP and 1000 hPa wind at 0000 UTC 17 July.

In WRF-Var system, several background error statistics (BES) options for control variable are available. For control variables in physic space, the NCEP (CV3) BES could be used, and for control variables in eigenvector space, the NCAR (CV5) BES was derived based on the one month forecast dataset. The first two experiments used the default settings of the CV3 and CV5 BES called: 3DVCV3 and 3DVCV5.

Table 1. Typhoon Haitang central pressure and position at 2005071700Z for observation, FG, and experiments

Exp.	pressure	latitude	Longitude
Observation	912.00	21.50	125.80
FG3DVAR	957.29	21.63	124.93
3DVCV3	950.71	21.63	124.96
3DVCV5	951.42	21.63	125.01
3DVCV5E3	949.03	21.40	125.81
3DVCV5E2	947.69	21.41	125.81

From Table 1 and Fig. 3a to 3d, the central pressure, compared with the FG, was deepened by about 6~7 hPa with 3DVCV3 and 3DVCV5, but the location was almost same as that in FG. We also tried to assimilate the 40 TC bogus soundings only, and only single point of SLP data at the typhoon center. The analysis of the typhoon position was not improved at all, just the central pressure was deepened to 904 hPa when only single point of SLP was assimilated (not shown).

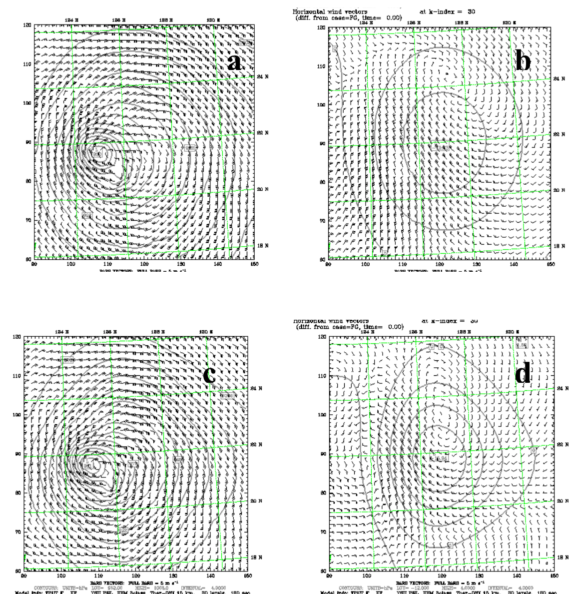

Fig. 3. Analysis of the Typhoon Haitang a) SLP field and surface wind b) increment of SLP and surface wind from 3DVCV3, c) SLP field and surface wind d) increment of SLP and surface wind from 3DVCV5.

3DVCV5 gave larger increment of SLP and similar influence range to 3DVCV3, but the final SLP analyses were very similar. This may be because the CV3 BES and CV5 BES, though derived from different forecast datasets, they were all computed with NMC-method (parrish and Derber, 1992). The former used the differences of NCEP GFS 24- and 48-h forecasts over 49 cases over a year, and the latter used the differences of the WRF-ARW model 12- and 24-h forecasts over the month of July 2005, which included the period of Typhoon Haitang. It is possible that the background error variances in NCEP CV3 BES might be underestimated for this typhoon case, which could lead to smaller increments of the SLP than those with CV5 BES.

Two questions are raised i) why did it fail to assimilate the TC bogus data with WRF-Var here?", and ii) "Is it possible for WRF-Var system to successfully assimilate the TC bogus data for Typhoon Haitang?"

Figure 4 may explain the reasons for failure of the TC bogus data assimilation in this case: a) the position of storm in FG is too far off the observed one. This may always happen in cycling mode because the FG is a forecast from the previous cycle. With cold-start run, i.e. the FG obtained from a large domain analysis (such as NCEP GFS), the storm position may be close to the observed one because certain re-location technique has already been implemented in these analyses. b) the scale-length in BES used in the recursive filter is too long (thin short dash line in Fig. 4). Thus, in the final analysis, the position of the storm is very close to the original one in FG (thick dot-dash line in Fig. 4). When the scale-length is shortened (thin long dash line in Fig. 4), the storm position in the final analysis (thick double-dot dash line in Fig.4) is close to the observed one.

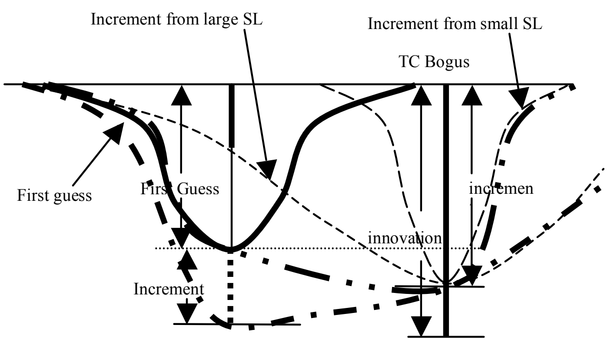

Fig. 4. Schematic of the impact of the scale-lengths on the final analysis.

Typhoon Haitang is a small-scale system. In GSI (Grid Statistics Interpolation) system (Wu et al 2002), two different scales were used in its recursive filter to achieve a fat-tailed spectrum for background error, in order to improve the analysis on the smaller scales. In Cressman-type objective analysis approach (Cressman, 1959), the different influence radii were used in the successive correction to account for the different scale analysis. However, WRF-Var did not consider the multiple scales in application of the recursive filter in an internal minimization loop, as in GSI. Rather WRF-Var has the multiple external loops function. Therefore implementation of the different tuning factors to BES for the different external loops in WRF-Var system will mimic the multiple scales in recursive filter in GSI and multiple influence radii in successive correction approach. Here, we conduct two additional analysis experiments,

3DVCV5E3: CV5 BES with 3 sets of tuning factors for 3 external loops as below,

VAR_SCALING1 = 1.50, 1.00, 0.50, (ψ)
VAR_SCALING2 = 1.50, 1.00, 0.50, (χ_u)
VAR_SCALING3 = 1.50, 1.00, 0.50, (T_u)
VAR_SCALING4 = 1.00, 1.00, 0.50, (rh)
VAR_SCALING5 = 1.50, 1.00, 0.50, (Psfc_u)

LEN_SCALING1 = 1.00, 0.50, 0.25, (ψ)
LEN_SCALING2 = 1.00, 0.50, 0.25, (χ_u)
LEN_SCALING3 = 1.00, 0.50, 0.25, (T_u)
LEN_SCALING4 = 1.00, 0.50, 0.50, (rh)
LEN_SCALING5 = 1.00, 0.50, 0.20, (Psfc_u)

3DVCV5E2: CV5 BES with 2 sets of tuning factors for 2 external loops. The second set of the tuning factors was removed.

To ingest more information of the observations for larger scales (LEN_SCALING1=1.0, etc.) in the first external loop, the larger variance tuning factors (VAR_SCALING1=1.50, etc.) are used. With the fidelity of the FG from the previous external loops become higher, the variance tuning factors are decreased. Meanwhile, the scale-lengths tuning factors are reduced to account for the smaller scale analysis. The more external loops in WRF-Var are executed, the more computing cost is required. To speed up the WRF-Var, in addition to 3DVCV5E3, 3DVCV5E2 with two external loops is carried out by ignoring the middle tuning factors.

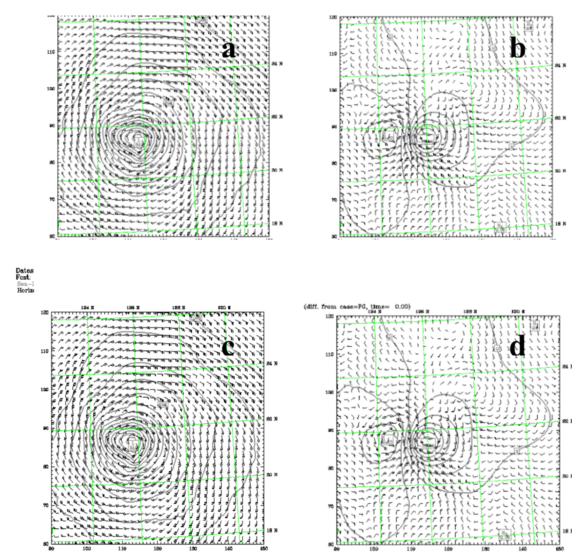

Fig. 5. Analysis of Typhoon Haitang a) SLP and surface wind fields and b) increment of SLP and wind from 3DVCV5E3, c) SLP and surface wind fields and d) increment of SLPand wind from 3DVCV5E2.

From Fig. 5 and Table 1, 3DVCV5E3 gave very good analysis of the typhoon position after the bogus data assimilation with WRF-Var. The scales of the SLP increment field are much smaller than those from 3DVCV3 and 3DVCV5, and there is a positive increment center west of the negative increment center (Fig. 5b). This made the pressure increasing at the FG storm center and pressure deepening at the observed center. So finally the correct typhoon position analysis is obtained. The central pressure from 3DVCV5E3 is also a few mb deeper than 3DVCV3 and 3DVCV5 (Tale 1). 3DVCV5E2 gave very similar analysis to 3DVCV5E2 with even deeper central pressure (Fig. 5c and 5d, Table. 1).

It is clear that the technique using multiple external loops with the different tuning factors to BES in WRF-Var can successfully assimilate the TC bogus data, and obtain improved analysis of yphoon position and intensity. Next question is "Does improved initialization with this technique in WRF-Var lead to improved forecast?"

4. Forecast experiment design

To answer this question, six experiments with WRF-Var and WRF-ARW forecast model were conducted: two in cold-start mode and four in warm-start (cycling) mode.

Cold_SI : Cold-start initiated by WRF_SI based on the NCEP GFS analysis
C3DVCV3 : Warm-start (cycling) runs initiated by WRF-Var with CV3 BES
C3DVCV5 : Warm-start (cycling) runs initiated by WRF-Var with CV5 BES
C3DVCV5E3: Warm-start (cycling) runs initiated by 3 External loops WRF-Var with 3 different tuning factors to CV5 BES for both domains
C3DVCV5E2: Warm-start (cycling) run initiated by 1 External loop WRF-Var for domain1 (45-km) and 2 External loops with different tuning factors to CV5 BES for domain2 (15-km)
Cold_CV5 : Cold-start initiated by WRF-Var with CV5 BES

All experiments with 4 letters "C3DV" means the 45/15km two domains (Fig.1) 6-h cycling mode experiments with WRF-Var/WRF starting from 0000 UTC 14 July 2005. The beginning initial condition at 0000 UTC 14 July 2005 and all the boundary conditions are created by WRF_SI from NCEP GFS analysis. The 72-h forecasts are made for four initial times, 0000 UTC and 1200 UTC 16, 0000 UTC and 1200 UTC 17 July (Fig. 6).

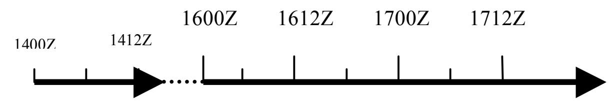

Fig. 6. Schematic of cycling mode WRF-Var/WRF experiments.

For the cold-start experiments, the initial conditions at the four initial times are generated from NCEP GFS analysis by WRF_SI (an interpolation procedure) or WRF_SI plus 3DVCV5.

In WRF-Var, in addition to the conventional GTS data (SYNOP, METAR, SHIP, BUOY, TEMP, AIREP, PILOT, SATEM, and SATOB), the GPS REF, QuikScat, and TC and Global Bogus data, are also assimilated. The results for 4 initial times: 1600Z, 1612Z, 1700Z, and 1712Z, are summarized below. Note that at 1600Z, the warm-start experiments have already gone through eight 6-h cycles starting from 1400 UTC July 14, 2005.

4.1 Analysis errors at initial times

Table 2a. Analysis errors (km) of Typhoon Haitang location at the initial time

Initial time	1600	1612	1700	1712	Mean
Cold_SI	32.9	42.7	42.1	0.2	29.5
C3DVCV3	170.9	108.5	77.0	107.7	116.0
C3DVCV5	78.8	48.6	81.3	61.0	67.4
C3DVCV5E3	22.1	6.6	52.7	12.2	23.4
C3DVCV5E2	3.5	14.9	92.7	21.2	33.1
Cold_CV5	9.8	5.5	11.7	4.1	7.8

Table 2b. Analysis errors (hPa) of Typhoon Haitang central pressure at the initial time

Initial time	1600	1612	1700	1712	Mean
Cold_SI	52.8	69.6	73.0	60.7	64.0
C3DVCV3	26.2	40.1	41.6	25.7	33.4
C3DVCV5	28.2	33.9	39.0	28.0	32.3
C3DVCV5E3	22.8	25.4	35.4	24.2	27.0
C3DVCV5E2	23.3	27.4	32.7	23.4	26.7
Cold_CV5	40.5	55.4	57.8	51.3	51.3

Table 2a and 2b gave the analysis errors of Typhoon Haitang's location and central pressure. Note that there are two differences from the analysis experiments at 0000 UTC 17 July presented in the last section: i) all 12 types of available observations are assimilated; ii) the FG is obtained from the 6-h forecast in cycling mode runs or from the WRF_SI or WRF_SI plus 3DVCV5. The experiments here are more realistic in terms of operational configuration.

From Table 2a, the Typhoon location from Cold_SI and Cold_CV5 are rather accurate, which means that the NCEP GFS analysis may already have certain re-location technique implemented. WRF_SI plus WRF-Var/3DVCV5 placed the typhoon very close to the observed location, only 7.8 km departure averaged over four times. But the central pressure errors from the cold-start mode experiments are extremely large, 64 hPa for Cold_SI and 51.3 hPa for Cold_CV5 averaged over four analysis times.

From Table 2b, it is clear that the reduction of the central pressure errors is achieved through cycling. Even with the climatological BES without any tuning, the central pressure errors are reduced significantly to 33.4 hPa for C3DVCV3 and 32.3 hPa for C3DVCV5.

C3DVCV5 with CV5 BES works better than C3DVCV3 (CV3 BES), the averaged analysis location error is 67.4 km, only about half of that from C3DVCV3, 116.7 km. Presumably, the CV5 BES derived from the same month (July 2005) and model (WRF-ARW) forecast dataset is better than NCEP CV3 BES.

C3DVCV5E3 gave the second accurate location analysis (23.4 km), after Cold_CV5, but the averaged central pressure error is 27 hPa, only half of that from Cold_CV5. C3DVCV5E2 gave the worse location analysis (33.1 km) and a little better pressure analysis (26.7 hPa) than C3DVCV5E3.

Overall, in a realistic environment, C3DVCV5E3, the technique using 3 external loops with the different tuning factors to CV5 BES, obtained the best analysis results. Next question is "how about the forecast?"

4.2 Track forecast errors

Figure 7 shows the track forecast for Typhoon Haitang at four different initial times: 0000 and 1200 UTC 16 and 0000 and 1200 UTC 17 July 2005. At the second and third initial times: 1200 UTC 16 and 0000 UTC 17 July, the track forecasts from all experiments are close to each other while at the first and last initial time: 0000 UTC 16 and 1200 UTC 17 July, the forecast tracks are diverged. In general, the track forecasts with CV5 BES are better than those with C3DVCV3 and Cold_SI. Figure 8 showed the track forecast errors at the different initial times, and the mean errors for all four initial times. From the mean track forecast errors, C3DVCVE3 is ranked the first, then Cold_CV5, C3DVCV5E2, C3DVCV5, Cold_SI, and final C3DVCV3. Although C3DVCV3 is the worst one here, we want to emphasize that the CV3 BES used here is not derived specially for typhoon cases and no tuning factors is applied to it. This result does not imply that the approach of using BES in physical space is not working. Rather, people should use it with care. Occasionally, at initial time 0000 UTC 17 July, C3DVCV3 gave the smallest track forecast error.

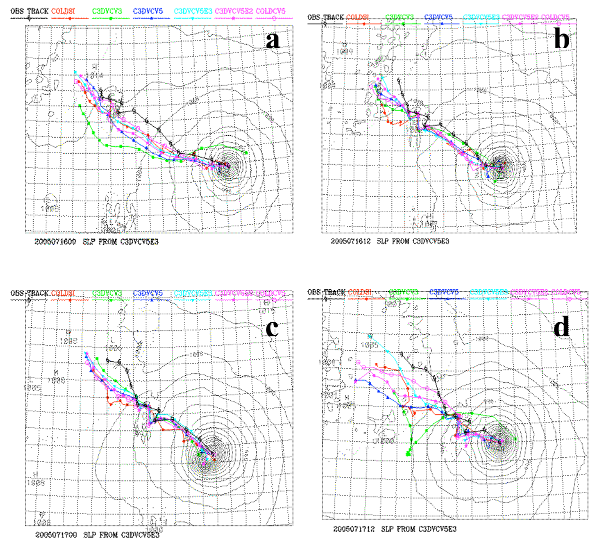

Fig. 7. The 72-h track forecast for different experiments at the different initial times: a) 1600Z, b) 1612Z, c) 1700Z, and d) 1712Z July 2005. The black lines are the best track. The contour lines are the SLP analysis from C3DVCV5E3.

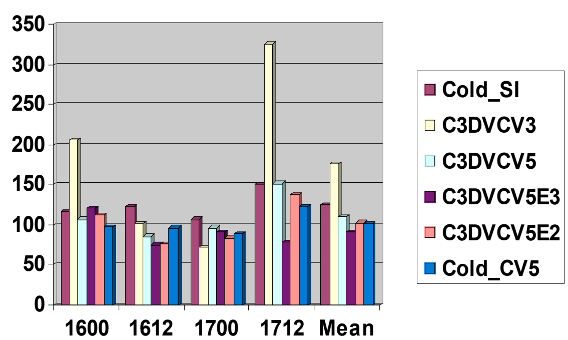

Fig. 8. The track forecast errors (km) over 72-h period for different experiments at the different initial times, and the mean errors averaged for four initial times.

Based on the best track data from CWB, we have the locations and central pressure data every 6-h for Typhoon Haitang. The forecast errors presented above are *pure* forecast errors, excluding the errors at the initial times, are calculated from 6-h to 72-h forecast locations against the 12 corresponding locations of the best track during each of 72-h periods. For details, we also calculated the track forecast errors for each of 24-h periods, i.e. 6 to 24-h, 30 to 48-h, and 54 to 72-h. In each of the 24-h periods, there are 4 observed locations. The results are shown in Fig. 9.

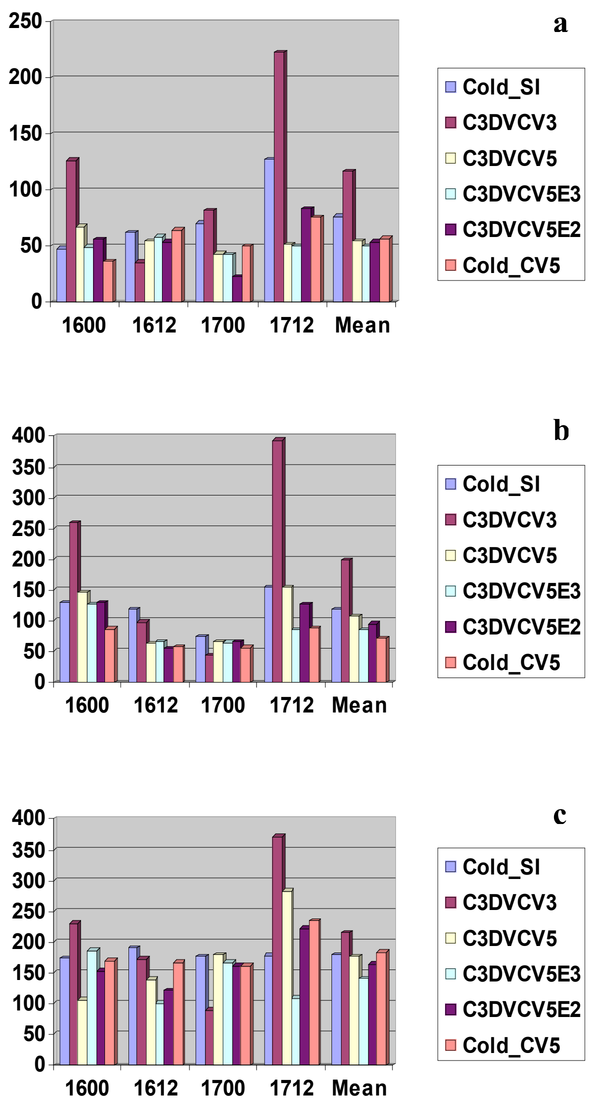

Fig. 9. Track forecast errors for different 24-h periods: a) 6 to 24-h, b) 30 to 48-h, and c) 54 to 72-h.

From Fig. 9, the track forecast errors from the Exp with CV5 BES usually are smaller than Exp Cold_SI and Exp with CV3 BES. Among all experiments with CV5 BES, the C3DVCV5E3 gave the smallest mean errors, even smaller than Cold_CV5 for the first and last 24-h periods, i.e. the warm-start (cycling) run is superior to the cold-start run.

4.3 Central pressure forecast

For the central pressure forecast, it can be clearly seen in Fig. 10 that the cycling mode experiments produce superior results to the cold-start mode runs.

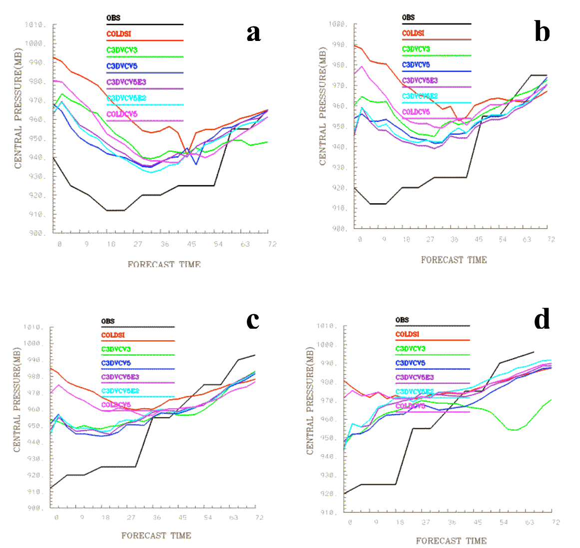

Fig. 10. The central pressure forecast for the different experiments at the different initial times: a) 1600Z, b) 1612Z, c) 1700Z, and d) 1712Z July 2005.

At the initial times, the central pressure from the cycling mode runs are 15 to 30 hPa lower than those of the cold-start runs, Cold_SI and Cold_CV5. These differences are gradually decreased as the forecast time advances, but still much lower than the cold-start runs, which, in turn, is still much higher than the observed central pressure.

At the beginning 6-h forecast, there are a few to ten hPa jump in the experiments with WRF-Var (3DVar). This is because the imbalance exists in the initial conditions generated by WRF-Var. Although WRF-Var is a multivariate analysis system with certain balance relationships in the background error statistics, these balance relationships may still be different from those in the WRF forecast model. A certain filtering or initialization procedure may be needed prior to integrating the forecast model, such as Digital Filter Initialization (DFI) (Huang and Lynch, 1993). How to best couple the WRF-Var with the initialization procedure to improve the forecast is an interesting research topic. Using an off-line DFI or implementation of DFI as a weak constraint in 4DVar (Wee and Kuo 2004) in WRF-Var, etc. to improve the forecast are worth to try.

5. **Summary and conclusions**

With an objective to obtain improved typhoon analysis, we perform bogus data assimilation with the current version of WRF-Var. Using WRF-Var as a black-box with the default, climatological BES cannot produce the correct analysis of the typhoon location and intensity for Haitang. We found that a technique combining the external loops with the tuning factors to CV5 BES is the best choice. We then performed a series of forecast experiments, including the cold-start mode and warm-start (cycling) mode experiments, at four initial times to assess the impact of the different background error statistics used in WRF-Var on the typhoon track and intensity forecast. The results led to the following conclusions:

1) The technique using multiple external loops with the different tuning factors to CV5 BES can give the correct analysis of typhoon location. With the cycling mode runs, this WRF-Var technique, in general, also gave the best forecast of the typhoon track and intensity.

2) CV5 BES usually produced more accurate analysis and forecast than CV3 BES for Typhoon Haitang. This is because the CV5 BES was derived from the differences of 12 and 24h forecasts based on the same WRF model and in the same month, July 2005, in which Typhoon Haitang occurred. So the specially derived BES for the case of interest is recommended.

3) In the cold-start mode runs, more accurate analysis of the typhoon location, but not intensity, were obtained with WRF-Var. It seemed that certain re-location technique may have already been implemented in the global analysis, such as NCEP GFS. So the re-location technique implemented in the FG field may help the WRF-Var to obtain the accurate typhoon location analysis.

4) In the cycling mode experiment, the analysis and forecast of the typhoon intensity are much better than those from the cold-start experiments. With a 15-km model resolution, it is impossible to forecast the true intensity of the typhoon, for example, 912 hPa at 0000 UTC 17 July 2005. But the cycling mode runs can provide much improved the central pressure forecast than the cold-start runs.

Overall, from this study, the technique using the multiple external loops with the different tuning factors to CV5 BES in WRF-Var and analysis cycling gave the best analysis and forecast of Typhoon Haitang.

When this technique applied to Typhoon Dujuan occurred in August 2003 over western Pacific Ocean, similar results are also obtained. WRF-Var followed by a proper initialization procedure and a re-location technique implemented in the first guess prior to WRF-Var may further improve the WRF-Var/WRF system for the typhoon and hurricane analysis and forecast.

References:

Barker, D. Y.-R. Guo, W. Huang, A. J. Bourgeois, and Q. Xiao, 2004: A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results. *Mon. Wea. Rev.,* **132**, 897-914.

Cressman, G. P., 1959: An Operational Analysis Scheme. *Mon. Wea. Rev.,* **87**, 367-374.

Guo, Y.-R., H. Kusaka, D. M. Barker, Y.-H. Kuo, and A. Crook, 2005: Impact of Ground-based GPS PW and MM5-3Dvar Background Error Statistics on Forecast of a Convective Case. *SOLA,* Vol. **1**, 073-076, doi: 10.2151/sola.

Huang, X.-Y., and P. Lynch, 1993: Diabatic Digital-Filtering Initialization: Application to the HIRLAM model. *Mon. Wea. Rev.,* **121**, 589-603.

Skamarock, 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/TN-468+STR,* 88 pp.

Wee, T-K., and Y.-H. Kuo, 2004: Impact of a digital filter as a weak constraint in MM5 4DVAR: An observing system simulation experiment. *Mon. Wea. Rev*., **132**, 543-559.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Wea. Rev.,* **130**, 2905-2916.

Zou, X., and Q. Xiao, 2000: Studies on the Initialization and Simulation of a Mature Hurricane Using a Variational Bogus Data Assimilation Scheme. *J Atnos. Sci.,* **57**, 836-860.