Impact of Four-Dimensional Variational Data Assimilation of Atmospheric Motion Vectors on Tropical Cyclone Track Forecasts

DONGLIANG WANG, XUDONG LIANG, AND YIHONG DUAN

Shanghai Typhoon Institute, and Laboratory of Typhoon Forecast Technique/CMA, Shanghai, China

JOHNNY C. L. CHAN

Shanghai Typhoon Institute, Shanghai, and Department of Physics and Materials Science, City University of Hong Kong, Hong Kong, China

(Manuscript received 8 September 2004, in final form 27 December 2005)

ABSTRACT

The fifth-generation Pennsylvania State University–National Center for Atmospheric Research nonhydrostatic Mesoscale Model is employed to evaluate the impact of the *Geostationary Meteorological Satellite-5* water vapor and infrared atmospheric motion vectors (AMVs), incorporated with the four-dimensional variational (4DVAR) data assimilation technique, on tropical cyclone (TC) track predictions. Twenty-two cases from eight different TCs over the western North Pacific in 2002 have been examined. The 4DVAR assimilation of these satellite-derived wind observations leads to appreciable improvements in the track forecasts, with average reductions in track error of ~5% at 12 h, 12% at 24 h, 10% at 36 h, and 7% at 48 h. Preliminary results suggest that the improvement depends on the quantity of the AMV data available for assimilation.

1. Introduction

The scarcity of observations, both near the storm center and in the surrounding environment, is a key factor in limiting the accuracy of tropical cyclone (TC) forecasts. With geostationary satellite imageries, some winds over the data-void regions could be derived by tracking cloud and water vapor features, which are generally referred to as atmospheric motion vectors (AMVs). This ability makes AMVs particularly useful for studying TCs (Velden et al. 1998). Although AMVs have been produced operationally for more than a decade, and recent enhancements in spatial resolution and radiometric sensitivity have significantly improved both the accuracy and density of the wind products (Velden et al. 2005), relatively few attempts have been made to determine the impact of AMVs on numerical TC track forecasts, particularly for those in the western North Pacific (WNP). Velden et al. (1992), Goerss et al. (1998), and more recently Soden et al. (2001) have used some simple schemes, for example, the three-dimensional optimum interpolation method, to assimilate the AMVs directly into various forecast models. They found that such assimilations resulted in reductions in the mean track error to some extent.

More recently, Xiao at el. (2002) examined the impact of satellite-derived wind observations on the prediction of a mid–Pacific Ocean cyclone with the four-dimensional variational (4DVAR) data assimilation method. Only a slight positive impact on the prediction was found, but their study was limited to only one mid–Pacific Ocean cyclone and five time-lag (interval = 12 h) experiments. Zhang and Wang (1999) used AMVs to correct the objectively analyzed wind fields and to construct an asymmetric bogus vortex. The method improved the forecasts of TC tracks.

As the 4DVAR data assimilation scheme has been demonstrated to be an effective tool to assimilate nonconventional data (e.g., Xiao et al. 2000), the objective of this study is therefore to investigate the impact of

Corresponding author address: Dongliang Wang, Shanghai Typhoon Institute, 166 Puxi Rd., Shanghai 200030, China.
E-mail: wangdl@mail.typhoon.gov.cn

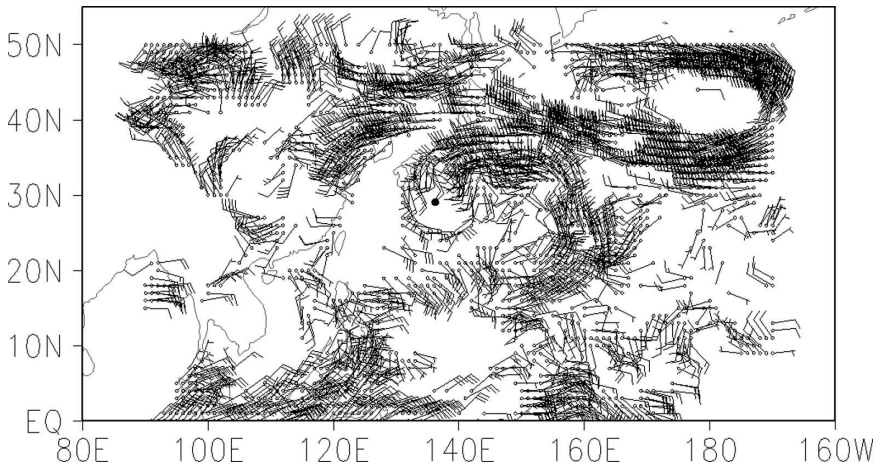

FIG. 1. *GMS-5* infrared and water vapor AMVs at 1200 UTC 17 Aug 2002 between 250 and 300 hPa. The dot indicates the TC center.

assimilating AMVs with the 4DVAR method on the prediction of WNP TC tracks. Section 2 gives a brief description of the AMV data used and the pretreatment applied, as well as the TC studied. The assimilation methodology and experimental design are introduced in section 3. The results and their discussions are presented in section 4 while section 5 gives a summary.

2. Datasets

a. Satellite winds

High-resolution visible (0.6 μm) imageries are used to track shallow cumulus clouds in the lower troposphere (600–900 hPa), while the upper winds are derived from the infrared (11-μm window) imageries. In clear regions in the upper (150–350 hPa) and middle (350–550 hPa) troposphere, winds are derived by tracking water vapor structures observed in the water vapor channels [6.7, 7.0, and 7.3 μm; Nieman et al. (1997)]. All the data assimilated in this paper are derived from the *Geostationary Meteorological Satellite-5* (*GMS-5*) infrared and water vapor imageries provided by the China National Satellite Meteorological Center.

On average, 1344 AMV observations are available within the model domain (about 3825 km $\times$ 4095 km), 95% of which are observed above 400 hPa and 55% are concentrated between 200 and 300 hPa. As an example, consider the AMV data available at 1200 UTC on 17 August 2002. About 52% of the multispectral winds are observed between 200 and 300 hPa. The AMVs between 250 and 300 hPa (Fig. 1) depict the important synoptic and subsynoptic-scale circulation features over the WNP. The upper-tropospheric flow is characterized by an anticyclonic circulation around (30°N, 138°E) and a cyclonic circulation in the northeast. With such comprehensive coverage over the oceanic regions, it is anticipated that these data should improve the forecast model initial conditions.

Because the AMV data are derived with reference to direct surface or upper-wind observations, their errors have a larger randomness. A simple quality control similar to the first-guess check applied in the European Centre for Medium-Range Weather Forecasts (ECMWF) prediction system (Bormann et al. 2003) is used to reduce the impact of this kind of error. According to the statistics from Wang et al. (1997), the error in the upper-tropospheric AMVs is 6 m s^{-1}. Any AMV with a difference from the original model analysis >3 times the error, that is, 18 m s^{-1}, is therefore rejected.

b. TC data

Twenty-two cases are chosen from eight different WNP TCs in 2002 (Table 1). Because the experiments are not performed in real time, the selection of the individual cases is based upon the availability of both the AMVs and National Centers for Environmental Prediction (NCEP) analyses for initializing the model. The model forecast tracks are verified against the best-track analyses of the Annual Typhoon Report of the China Meteorological Administration.

3. Assimilation methodology and experimental design

Experiments are carried out using the fifth-generation Pennsylvania State University–National Center for Atmospheric Research nonhydrostatic Mesoscale Model version 5 (MM5) and its 4DVAR sys-

TABLE 1. The name, number of cases, forecast dates, and current intensities used for the CTRL and WIND experiments.

Name	No. of forecasts	Time and date (during 2002) and intensity (hPa)
Rammasun	4	1200 UTC 28 Jun, 1000; 1200 UTC 1 Jul, 980; 0000 UTC 2 Jul, 970; 1200 UTC 3 Jul, 950
Chataan	2	1200 UTC 1 Jul, 990; 1200 UTC 4 Jul, 980
Fengshen	5	0000 UTC 19 Jul, 925; 0000 UTC 20 Jul, 930; 0000 UTC 24 Jul, 965; 0000 UTC 25 Jul, 975; 0000 26 Jul, 975
Fungwong	2	0000 UTC 24 Jul, 970; 0000 25 Jul, 975
Phanfone	2	1200 UTC 16 Aug, 950; 1200 UTC 17 Aug, 950
Vongfong	1	1200 UTC 17 Aug, 990
Rusa	2	1200 26 Aug, 950; 1200 UTC 30 Aug, 960
Sinlaku	4	1200 UTC 2 Sep, 960; 1200 UTC 3 Sep, 960; 1200 UTC 4 Sep, 960; 1200 UTC 6 Sep, 960

tem. The model with a single domain has a grid spacing of 45 km, and 85 × 91 grid points. The model center is located at the best-track center of the TC. There are 23 layer half-σ levels in the vertical and the pressure at model top is 10 hPa. Considering the AMVs are largely concentrated in the upper troposphere, the upper and lower vertical layers are denser. Physics options used for the variational data assimilation system are the same as those for the forecast model, including the Kuo (1974) cumulus parameterization, the Blackadar (1979) high-resolution planetary boundary layer parameterization scheme, and an explicit moisture scheme (Dudhia 1989). The first-guess conditions are obtained from the NCEP 0-h analysis or NCEP 12-h forecast. To assess the impact of the AMVs, no bogus scheme is used.

A 6-h assimilation window is used for the MM5 4DVAR system to incorporate the *GMS-5* AMV data at the initial time and 6 h later. The cost function to be minimized is

$$\mathbf{J} = \sum_{m=1,2} \mathbf{J}_m + \mathbf{J}_b, \tag{1}$$

where

$$\mathbf{J}_b = \frac{1}{2}(\mathbf{X}_0 - \mathbf{X}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{X}_0 - \mathbf{X}_b) \quad \text{and} \tag{2}$$

$$\mathbf{J}_m = \sum_r [(\mathbf{H}_l\mathbf{u} - \mathbf{u}^{\mathrm{sat}})^T\mathbf{W}_u(\mathbf{H}_l\mathbf{u} - \mathbf{u}^{\mathrm{sat}}) + (\mathbf{H}_l\mathbf{v} - \mathbf{v}^{\mathrm{sat}})^{\mathrm{T}}\mathbf{W}_v(\mathbf{H}_l\mathbf{v} - \mathbf{v}^{\mathrm{sat}})], \tag{3}$$

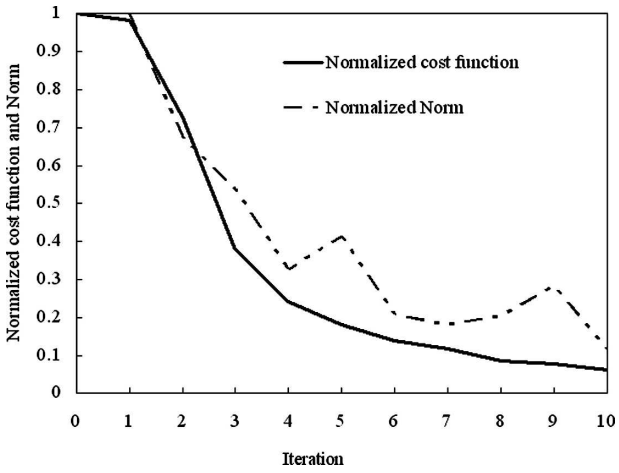

FIG. 2. Variations of the normalized cost function and gradient norm with the number of iterations for the case of Typhoon Phanfone at 1200 UTC 17 Aug 2002.

with $\mathbf{J}_b$ being the background partial cost function and $\mathbf{J}_m$ the partial cost function that measures the discrepancy between the model-predicted and AMVs. Here, $\mathbf{X}_0$ is the initial condition, which includes initial wind components u and v, temperature, specific humidity, and pressure perturbation; $\mathbf{X}_b$ is the background term, and has the same components as $\mathbf{X}_0$; $\mathbf{B}$ is an approximated background error variance matrix that relates to the MM5 analysis at the initial time and 12 h later; m is the number of times needed to assimilate the AMVs; r are the locations of the available wind data; $\mathbf{H}_l$ is a linear scheme used to interpolate the MM5-analyzed winds onto the target positions; and $\mathbf{W}_u$ and $\mathbf{W}_v$ are treated as constants, which are the inverse of the square of the AMV error, determined empirically to be 6 m s^{-1}.

The main steps are as follows. The cost function and forcing terms are calculated by integrating the forecast model forward for 6 h. In the course of integrating the adjoint-model backward, the forcing term is added at each observing time to obtain the gradient with respect to the initial condition. According to the gradient, the limited-memory quasi-Newtonian method is used to calculate the new initial condition, which can reduce the cost function. All these steps should be repeated until satisfactory convergence or the given number of iterations has been achieved. After obtaining the "optimal" initial conditions, forward runs are started from the optimal initial conditions using the MM5 for a 48-h simulation.

In the study of Zou and Xiao (2000), the largest adjustment in the model fields during the minimization procedure occurred at the second iteration, and the second largest adjustment at the fifth iteration. About

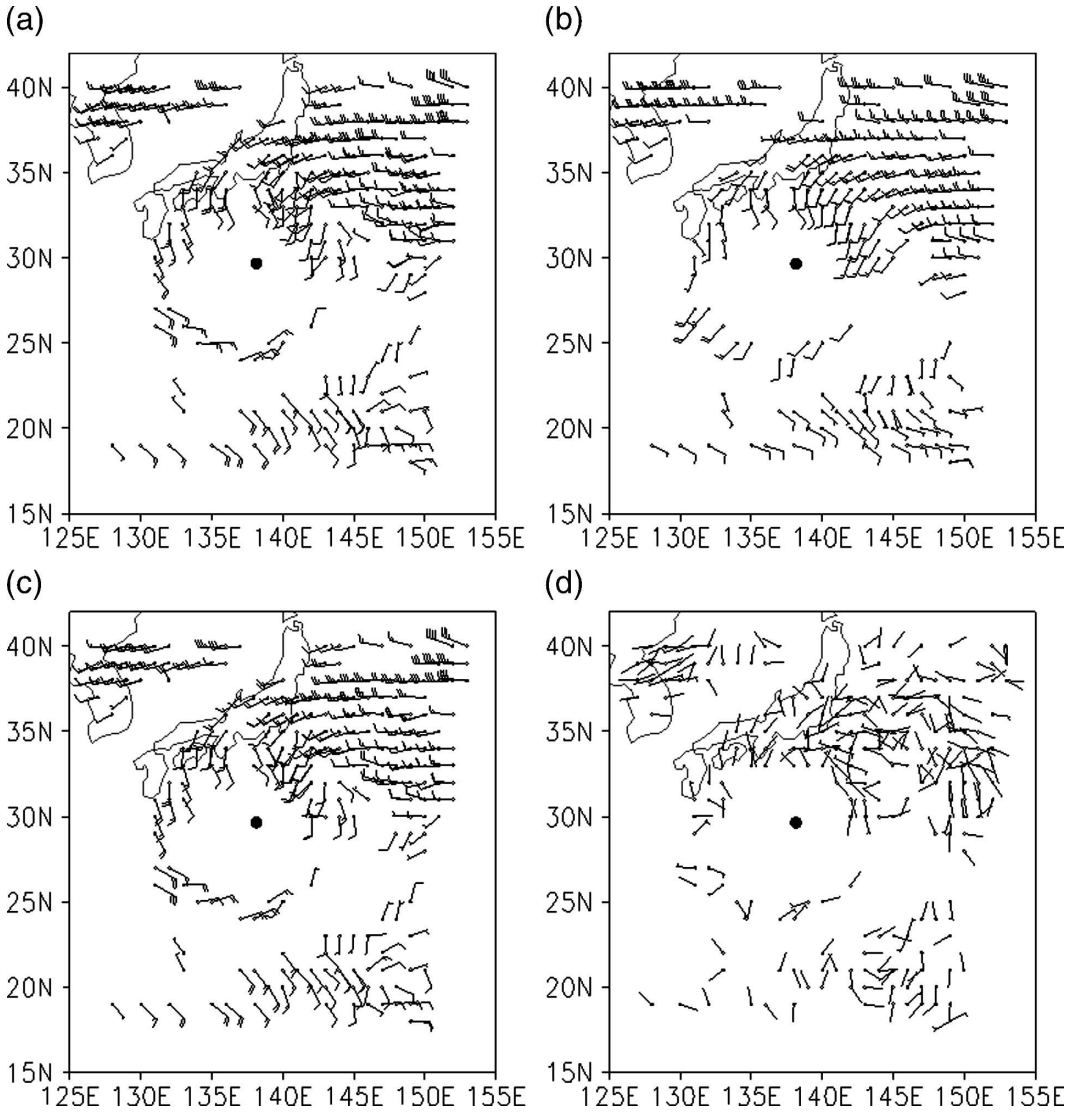

FIG. 3. (a) AMVs, (b) original initial winds from the MM5 analysis, (c) optimal initial winds, and (d) differences between (a) and (c) at 1200 UTC 17 Aug 2002 between 250 and 300 hPa. The dot indicates the TC center.

98% of the total adjustment was completed in seven iterations. According to the experiments of Xiao et al. (2002), the major reductions of the cost function and gradient norm occurred in the first five iterations. Therefore, to save computation time, only 10 iterations are carried out in this study.

For each case, 48-h forecasts are performed. A control run (CTRL) without the AMVs is also made for comparison.

4. Results

a. Impact on the initial conditions

The convergence effectiveness of the minimization procedure can be seen from the variation of the normalized cost function and gradient norm with the number of iterations, as illustrated for the case of Typhoon Phanfone at 1200 UTC on 17 August 2002 (Fig. 2). The major reductions occur in the first five iterations and only small reductions are noticed after eight iterations. After 10 iterations, the cost function has decreased by nearly 1.5 orders of magnitude, and the gradient norm has also decreased by 1 order. This demonstrates that the minimization procedure of the MM5 adjoint system works well and 10 iterations should be enough to assimilate the AMV data.

The effectiveness of the assimilation procedure in adjusting the initial conditions can be evaluated by comparing the original (Fig. 3b) and the adjusted (Fig. 3c) initial winds between 250 and 300 hPa with the same

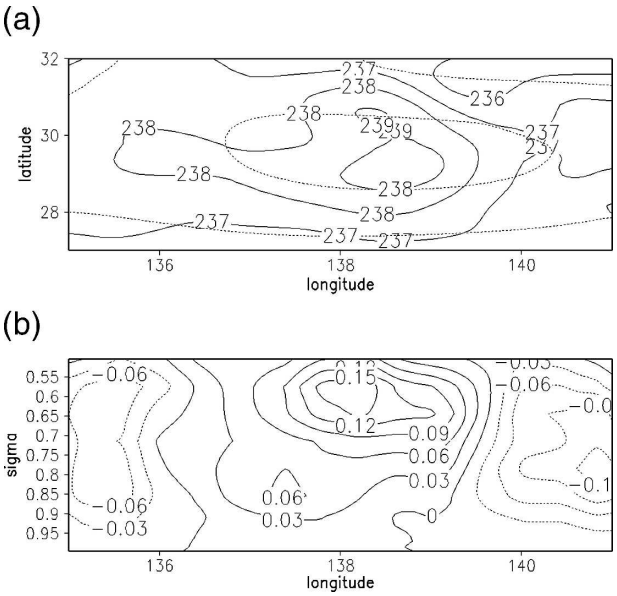

FIG. 4. (a) Distribution of initial temperatures (K) at 200 hPa (solid line, optimal temperature; dashed line, initial temperature; interval 1 K) and (b) east–west cross section of the difference in specific humidity (g kg^{-1}) through the TC center (138°E) below 500 hPa before and after assimilation (interval, 0.03 g kg^{-1}).

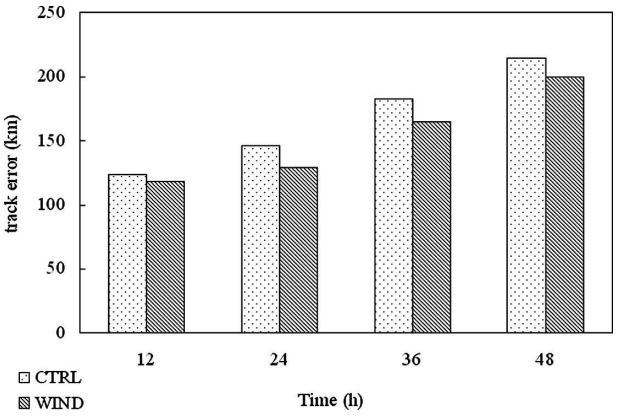

FIG. 5. Mean forecast track errors (km) for the 22 cases: CTRL and WIND experiments.

case of Typhoon Phanfone. For easy comparison, the MM5-analyzed (original) winds are interpolated onto the target positions. The largest adjustment through assimilation occurs south of the TC center (30°N, 138°E). The MM5 analysis from the CTRL run produced inaccurate south to southwesterly winds, which are changed into southeasterly to northeasterly winds in the assimilation (WIND experiment hereafter). Apparently, assimilation of the AMV data generates an anticyclonic circulation. Comparing Figs. 3a and 3c, it appears that the difference between the adjusted winds and the AMVs is quite small. To highlight that differences do exist, the difference vectors are plotted in Fig. 3d. In the lower troposphere, although the number of retrieved AMVs is less than that in the upper troposphere, and the differences are smaller, the adjusted winds still agree more closely with the AMV observations than the original winds (not shown).

During the minimization procedure, the assimilation variable (wind) is forced toward the observed information, while all other variables (e.g., temperature and moisture) are free to develop in a model-consistent manner. As an example, the distribution of the initial temperature at 200 hPa after assimilation shows an increase of about 1 K in the warm core (Fig. 4a). In the low and midtroposphere, the area around the eye becomes moister (Fig. 4b). Apparently, the additional information contained in the AMVs can propagate to other levels and other model variables through model integration by the 4DVAR assimilation procedure.

b. *Impact on track forecasts*

In general, assimilation of the AMV data reduces the mean track error at all forecast times, with larger relative reductions occurring after 24 h (Fig. 5). The mean 24-h position error is reduced from 146 to 129 km, and the 48-h error from 215 to 200 km. Among the 88 forecasts for the 22 cases (12-h interval), 48 forecasts have a reduction in track error after assimilation of the AMV data, 21 forecasts show no change in track error, and only 19 forecasts produced an increased track error (sum of each row in Table 2a, respectively). A further examination of the individual cases suggests that the effectiveness of assimilation of the AMV data varies with the quantity of the AMVs used, especially the AMV data at lower levels. Among the 44 forecasts with more than 1344 AMV observations (the average number of AMVs used in these experiments) assimilated, 27 (top number in left column) show improvement and only 5 (bottom number in left column) have larger errors. The number of AMV observations in the other 14 forecasts (bottom number in right column) with increased track error is less than the average. This result suggests that the improvement in the track forecasts depends on the quantity of the AMVs assimilated.

Using the adjoint technique (Errico and Vukicevic 1992; Zou et al. 1993; Vukicevic and Raeder 1995), Xiao et al. (2002) studied the sensitivity of the forecast model to model initial conditions. They found that the sensitivity region is mainly located over the eastern Asia coast and WNP, and the TC forecast is more sensitive to winds at lower levels than at higher levels. This may explain why their assimilation of the AMVs, which are over the northeast Pacific, resulted in very little

TABLE 2. Number of cases with different changes in the forecast errors compared with the CTRL run when different quantities of AMV observations are assimilated. (a) Data at all levels are included and (b) only data below 400 hPa are included. The average number of AMV observations including all levels is 1344 and that for observations below 400 hPa is 57. For each case, an entry in the "more" than the average category refers to the case having the number of AMV observations greater than this average number [1344 in (a) and 57 in (b)], and similarly for an entry in the "less" than the average category.

(a)

Change of forecast error compared with CTRL	No. of cases with No. of AMV observations	
	more than the average	less than the average
Reduced	27	21
No change	12	9
Increased	5	14

(b)

Change of forecast error compared with CTRL	No. of cases with No. of AMV observations	
	more than the average	less than the average
Reduced	33	15
No change	2	19
Increased	1	18

impact on TC prediction. The sensitivity of the TC forecasts to low-level winds is therefore examined further although the quantity of the AMV data available below 400 hPa is relatively small, with only an average of 57 observations. Among the 22 cases, 36 forecasts have more than this average number (sum of left column in Table 2b), and only 1 (bottom number in left column) has an error larger than that of CTRL, while an overwhelming majority show a reduction in track error. The other 18 (bottom number in right column) forecasts in the sample with increased error all have AMV observations less than the average. This result suggests that for a better TC forecast, more AMVs at lower levels are needed. In addition, an increase in the AMVs at lower levels apparently has a more significantly positive impact on the TC track forecast.

5. Summary and conclusions

This paper assesses the impact of the *GMS-5* AMVs on the track prediction of WNP TCs. Much of the AMV data are assimilated into a 45-km-resolution MM5 model featuring 4DVAR data assimilation routine with a 6-h assimilation window. Twenty-two cases from eight TCs occurring over the western North Pacific in 2002 are studied.

The MM5 4DVAR assimilation system is found to work well in assimilating the AMVs. The optimal initial condition can provide additional useful information compared with the original analysis. It can depict the important synoptic and subsynoptic-scale circulation features in the troposphere.

Quite promising overall 48-h forecast results are obtained by assimilating the AMV observations. However, the effectiveness of the assimilation of this *GMS-5* AMV data appears to vary with the quantity of the assimilated AMVs. Improved forecasts tend to occur for cases with more AMVs, especially if these observations are available below 400 hPa.

It is worth noting that the AMV observations from the GMS satellite used in this paper have limitation in data coverage. However, this limitation now can be overcome with AMVs from the Geostationary Operational Environmental Satellite (GOES), the Japanese geostationary satellite *Multifunctional Transport Satellite-1R* (*MTSAT-1R*), which are providing much improved coverage of the lower levels. For example, high-density IR winds from the 3.9-μm channel now available on the GOES imagers can provide excellent low-level coverage even at night (Velden et al. 2005). The assimilation system for TC track forecasts is anticipated to operate better with these high-density and high quality AMV data.

Acknowledgments. The authors thank the National Satellite Meteorological Center of the China Meteorological Administration for providing the AMV data.

REFERENCES

Blackadar, A. K., 1979: High resolution models for the planetary layer. *Advances in Environmental Science and Engineering,* J. R. Pfafflin and E. N. Ziegler, Eds., Vol. 1, Gordon and Breach Science Publishers, 50–85.

Bormann, N., S. Saarinen, G. Kelly, and J. Thepaut, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.,* **131,** 706–718.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

Errico, R. M., and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU–NCAR Mesoscale Model. *Mon. Wea. Rev.,* **120,** 1644–1660.

Goerss, J. S., C. S. Velden, and J. D. Hawkins, 1998: The impact of multispectral *GOES-8* wind information on Atlantic tropical cyclone forecasts in 1995. Part II: NOGAPS forecasts. *Mon. Wea. Rev.,* **126,** 1219–1227.

Kuo, H. L., 1974: Further studies of the parameterization of the influence of cumulus convection on large scale flow. *J. Atmos. Sci.,* **31,** 1232–1240.

Nieman, S. J., W. P. Menzel, C. M. Hayden, D. Gray, S. Wanzong, C. S. Velden, and J. Daniels, 1997: Fully automated cloud drift winds in NESDIS operations. *Bull. Amer. Meteor. Soc.,* **78,** 1121–1134.

Soden, B. J., C. S. Velden, and R. E. Tuleya, 2001: The impact of satellite winds on experimental GFDL hurricane model forecasts. *Mon. Wea. Rev.,* **129,** 835–852.

Velden, C. S., C. M. Hayden, W. P. Menzel, J. L. Franklin, and J. S. Lynch, 1992: The impact of satellite-derived winds on the hurricane track forecasting. *Wea. Forecasting,* **7,** 107–119.

——, T. L. Olander, and S. Wazong, 1998: The impact of multispectral *GOES-8* wind information on Atlantic tropical cyclone forecasts in 1995. Part I: Dataset methodology, description, and case analysis. *Mon. Wea. Rev.,* **126,** 1202–1218.

——, and Coauthors, 2005: Recent innovations in deriving tropospheric winds from meteorological satellites. *Bull. Amer. Meteor. Soc.,* **86,** 205–223.

Vukicevic, T., and K. Raeder, 1995: Use of an adjoint model for finding triggers for Alpine lee cyclogenesis. *Mon. Wea. Rev.,* **123,** 800–816.

Wang, Z., J. Chen, and W. Zeng, 1997: Cloud motion wind from CWIS in comparison to others (in Chinese). *J. Nanjing Inst. Meteor.,* **20,** 382–386.

Xiao, Q., X. Zou, and Y. H. Kuo, 2000: Incorporating the SSM/I-derived precipitable water and rainfall rate into a numerical model: A case study for ERICA IOP-4 cyclone. *Mon. Wea. Rev.,* **128,** 87–108.

——, ——, and M. Pondeca, 2002: Impact of *GMS-5* and *GOES-9* satellite-derived winds on the prediction of a NORPEX extratropical cyclone. *Mon. Wea. Rev.,* **130,** 507–528.

Zhang, S. F., and S. W. Wang, 1999: Numerical experiments of the typhoon tracks by using satellite cloud-derived wind (in Chinese). *J. Trop. Meteor.,* **15,** 347–355.

Zou, X., and Q. Xiao, 2000: Studies on the initialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.,* **57,** 836–860.

——, A. Barcilon, I. M. Navon, J. Whitaker, and D. G. Cacuci, 1993: An adjoint sensitivity study of blocking in a two-layer isentropic model. *Mon. Wea. Rev.,* **121,** 2833–2857.