Objective Verification and Data Assimilation of the ARW WRF Model in the Complex Terrain of Southwest Asia

Jianjun Xu
Joint Center for Satellite Data Assimilation (JCSDA) & (UCAR)/AFWA

Lee Byerle and Steven Rugg
Air Force Weather Agency (AFWA)

Abstract

This paper firstly describes an objective verification of the National Center for Atmospheric Research (NCAR) mesoscale model (ARW WRF) used in the complex terrain of Southwest Asia from May 1 through 31 2006. The statistical evaluation is designed to assess the model's surface and upper-air forecast accuracy at nine specific locations during the transforming season (May) of Asian Monsoon. The evidence is observed that the model biases caused by inadequate parameterization of physical processes, except for the 2-m (meter) temperature forecasts, are relatively small compared to these nonsystematic errors resulted partially from uncertainty of initial condition. The model errors in surface air forecasts are closely related to the terrain configuration, but the performance of 2-m temperature forecasts is different from that for 10-m wind field.

Secondly, in order to understand the role of initial condition in accuracy of the model forecasts, this study then assimilated a kind of satellite radiance data into this model through the Joint Center for Satellite Data Assimilation (JCSDA) analysis system (GSI). The results show that on average for the 24- and 48-h (hour) forecasts of one month experiments, the satellite data provides useful information for improvement of initial condition, the model errors get reduced for most of location within the 24-h forecasts. However, the improvement is mainly limited to the forecasts over lower terrain or water areas. With radiance data assimilation, the thermal and dynamics structure of upper air have been changed substantially, an extra warm anticyclone center appears over the central Southwest Asia. The satellite data assimilation produces a positive impacting to improve the forecast in the complex terrain areas.

1. Introduction

Weather prediction in Southwest Asia (SWA) is often very complex because of mesoscale variations induced by the diverse topography. This is a predominately semi-arid to arid region surrounded by Black and Caspian Sea in the north, the Mediterranean in the west, the Arabian Sea and Persian Gulf in the south and Himalayas in the east, and crossed by the impressive Tauros and Zagros mountains. Although a few previous model studies (Even and Smith 2001, Even et al. 2004) provided some interesting results for the basic weather of simulation in SWA using a regional climate model (RegCM2) or MM5 model, the performance of new generation mesoscale forecast model (ARW WRF) developed by NCAR (Skamarock et al., 2005) employed in the operational forecasts over this region is not quite understood. One of purpose in this paper is to evaluate the predictability of ARW WRF model in SWA complex terrain.

The evaluation primarily concentrates on the forecasts of wind and temperature in this study since SWA is dominated by the hot dusty wind weather (Agrawala, et al., 2001). During May the transforming season from winter to summer, the temperature and dust wind increase substantially and the unstable local-scale weather occurs frequently, so that the accuracy of prediction is highly dependent upon the accuracy of the temperature and wind forecasts.

Some recent studies have evaluated ARW WRF model based on objective error statistics for precipitation forecasts over United States. Christopher et al. (2006) evaluated WRF forecasts over the continental United States against stage-IV observations for precipitation forecasts, pointed out that the model bias is associated with systems over the different location. Cheng and Steenburgh (2006) presented surface sensible weather forecasts by WRF and Eta model over the western United States. Their results suggest that improvements in initialization may be as or more important than improvements in physics for land surface processes. Gallus and Bresch (2006) compared the impacts of WRF dynamics core, physics package, and initial conditions on warm season rainfall forecasts over central United States. They found that the sensitivity of rainfall forecasts is a function of model physics, dynamics, and initial conditions and depends on case by case. For heavier rainfall, sensitivity to initial conditions is generally less substantial than the sensitivity to changes in dynamic core or physics. For light rainfall, the WRF model using NCAR physics is much more sensitive to a change in dynamic core than the WRF model using NCEP physics. It is very clear that the errors of weather forecasts are caused by many reasons. One of these reasons can be attributed to the imperfection of numerical model in representing the actual atmosphere. However, as Lorenz (1963) pointed out, the most fundamental cause of forecast failure is because the atmosphere is a chaotic system. A chaotic system is defined as one in which evolution is sensitive to initial conditions (ICs). It means that an arbitrarily small error in the analysis of the initial state of the atmosphere can have an overwhelming effect in a finite amount of time. Therefore, it is not surprising that considerable effort has focused on improving the estimates of the model initial states through advanced techniques. One of such techniques is the data assimilation. So that another purpose in this paper is to evaluate the impact of data assimilation on the weather forecasts over SWA areas.

The paper is organized as follows. Section 2 describes the real-time configuration of ARW WRF as run operationally within SWA. Section 3 explains the objective methodology used in the statistical verification. The results of the objective verification for the May 2006 are presented in section 4. Section 5 discusses the impact of data assimilation on the weather forecast. Finally, the summary and discussion are given in section 6.

2. ARW WRF model and forecasts

The weather model used in this study is the Advanced Weather Research and Forecasting (ARW WRF) model (Michalakes et al. 2001; Skamarock et al. 2005), which is a nonhydrostatic, fully compressible, primitive equation model. Lead institutions involved in the effort include the National Center for Atmospheric Research (NCAR), Air Force Weather Agency (AFWA), National Oceanic and Atmospheric Administration (NOAA), and other government agencies and universities. WRF is built around a software architectural framework in which different dynamical cores and model physics packages are presented under the same code. With the WRF model, it is possible to mix and match the dynamical cores and physics packages of different models to optimize performance since each model has strengths and weaknesses in different areas. It uses a terrain-following hydrostatic pressure coordinate and the Arakawa C grid staggering.

A 15-km grid centered over the Southwest Asia (SWA) region (Fig. 1) is used to give better representation of the regions' complex topography and associated spatial variability in surface characteristics. To assess the predictability of the model, the forecasts are made for 48-h (hour) each day starting at 00Z during period of May 1 through 31 2006. In order to distinguish it from the following forecasts with data assimilation in section 5, the current forecast without data assimilation is named as NODA. The initial atmospheric and surface fields and boundary conditions, including soil moisture and temperature, are taken from the NCEP Global Forecast System (GFS) real time forecasts. The lateral boundary conditions are time dependent. The physics of the model are been chosen including the WRF Single Moment (WSM) 5-class microphysics scheme, Yosei University planetary boundary layer (PBL) scheme, Noah land surface scheme, Grell-Devenyi ensemble cumulus scheme, Rapid Radiative Transfer Model (RRTM) longwave radiation, and the Dudhia shortwave radiation scheme. The vertical 42-layer σ values are 0.995, 0.992, 0.983, 0.975, 0.961, 0.949, 0.932, 0.917, 0.897, 0.878, 0.855, 0.832, 0.806, 0.778, 0.749, 0.718, 0.687, 0.654, 0.623, 0.590, 0.559, 0.526, 0.495, 0.462, 0.431, 0.398, 0.367, 0.334, 0.304, 0.272, 0.244, 0.213, 0.187, 0.158, 0.134, 0.107, 0.085, 0.060, 0.040, and 0.018.

3. Topography and Evaluation method

To investigate the spatial heterogeneity of complex terrain in SWA region, nine representative sub-regions are depicted in Figure 1. They are defined as north of Iraq (A; 34°–36°N, 41°–43°E); northwest of Iran (B; 34°–36°N, 46°–48°E); north central of Iran (C; 34°–36°N, 54°–56°E); central of Afghanistan (D; 34°–36°N, 66°–68°E); west of Himalaya mountain (E; 34°–36°N, 74°–76°E); west of Saudi Arabia (F; 22°–24°N, 41°–43°E); east of Saudi Arabia (G; 22°–24°N, 51°–53°E); Arabian Sea (H; 22°–24°N, 63°–65°E); and west of India (I; 22°–24°N, 70°–72°E).

Table 1 displays the height of topography (Hgt), vegetation type (Veg) and soil type (Soil) over these nine regions. Except for the water type in Arabian Sea (marked H), the soil types in all other eight regions are loam, but the vegetation types show a little differences: the three high terrain over 2500 meter mountain regions (marked B, D and E) are covered by short plants with grass, shrubland and wooded tundra, respectively; the three low terrain under 1000 meter regions (marked A, C and F) and the two plain regions (marked G and I) almost do not have any plants. Obviously, the nine sub-regions represent significantly the heterogeneity of complex terrain in SWA region.

The objective evaluation is designed to present the model errors of winds and temperature for both 24-h (hour) and 48-h (hour) forecasts. The statistical measures used to quantify model

forecast errors are the bias (forecast − observation), root-mean-square (RMS) error, and error standard deviation (SD). For purposes of interpretation, the total error (RMS error) includes contributions from both systematic and random errors. Systematic error (Bias) can be caused by a consistent misrepresentation of physical parameters such as radiation or model convection. Nonsystematic or random errors, given by the error SD, are caused by uncertainties in the model initial condition or unresolveable differences in scales between the forecasts and observations (Nutter and Manobianco 1999). Note that due to lack of observations in SWA region, the analysis productions of GFS have been used to replace the observations.

If X represents any of the parameters under consideration for a given time and vertical level, then forecast error is defined as $X' = X_f - X_a$, where the subscripts *f* and *a* denote forecast and analyzed quantities, respectively. Given N valid pairs of forecasts and analysis, the bias is computed as

$$Bias = \overline{X'} = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} X_i' \qquad (1)$$

The RMS error is computed as

$$RMS = \left[\frac{1}{N} \sum_{i=1}^{N} (X_i')^2 \right]^{1/2} \qquad \mathbf{(2)}$$

The SD error is computed as

$$SD = \left[\frac{1}{N} \sum_{i=1}^{N} (X_i' - \overline{X'})^2 \right]^{1/2} \qquad \mathbf{(3)}$$

Note that if the model bias or systematic error is small, most of the RMS error is due to random, which is nonsystematic type variability in the errors.

4. Objective verification results

a. Surface

In the following section, ARW WRF model forecast error characteristics for 2-m (meter) temperature and 10-m wind speed are established. Results from the 30 days forecasts in the period of May 1 through 30 2006, the averages on both 24-h (hour) and 48-h forecasts are a function of location described in Fig. 1.

1) Temperature at 2-M

During May 2006, biases in 2-m temperature forecasts change with elevation of terrain (Fig. 2a). The amplitude of errors is larger at high terrain regions (E, B, D) with reaching −5°C of cold biases during the 24-h forecasts. Meanwhile, the biases are significantly smaller in low terrain regions (A, C, F, G, I) or water area (H) for both 24-h and 48-h forecasts.

Except that the SD error in the highest terrain region (E: west of Himalayas mountain) has an equivalent order of forecast bias, it is very small in the other areas. However, the forecast biases and corresponding RMS errors are comparable in magnitude in the most of mountain areas (Fig. 2a, b, c), the larger contribution to the total error for these locations evidently is derived from a systematic model error. The result indicates an apparent model deficit for description of surface temperature in high terrain areas.

2) Wind Speed at 10-M

Compared to forecasts of temperature at 2-m, the total forecasting errors in the 10-m wind speed are contributed largely by the nonsystematic error in the most of areas (not shown). The biggest of model biases occurs at northwest of Iran (B) with about −2.5 m s^{-1}. The biases at west of Himalayas mountain (E), east of Saudi Arabia (G) and west of India (I) are almost zero.

Based on above results, it is not difficult to find that the total forecast errors at surface show a substantial spatial heterogeneity, there is a relative bigger error in higher mountain areas. However, the resource of errors indicates a significant difference between temperature and wind speed, the errors in temperature at 2-m are mainly from systematic errors, which are controlled largely by the physical representation; meanwhile, the errors are positively correlation with elevation of terrain. In contrast, the errors in 10-m wind speed get more contribution from nonsystematic errors, which are probably related to uncertainty of initial condition. In general, the surface air does not apply to the upper air, how about the performance in upper air will be considered in the following section.

b. Upper air

1) Temperature at 500 hPa

The temperature biases at 500 hPa are range from -4°C to 1 for 24-h (Fig. 3) and a little colder with maximum -5°C for 48-h forecasts. The largest error appears over Himalaya mountain areas. The RMS errors range from about 1° to 2°C and the larger magnitude randomly distribute over central of Saudi Arabia, southeast of Iraq, northwest of Iran and west of Himalaya mountain. The corresponding error SD of 1°–2°C reveals that nonsystematic errors compose a substantial portion of the total error except for west of Himalaya mountain. Compared to the counterparts in 24-h forecast, the 48-h forecasts error is bigger in the most of study areas.

2) Winds at 200 hPa

Similar to the temperature forecasts at 500 hPa, the 24-h zonal wind forecasts at 200 hPa show (not shown) that the total forecast errors are dominated by the nonsystematic errors. But, note a substantial difference that the largest forecast error at Himalaya mountain is consistent with nonsystematic error and the Arabian Sea has also a strong nonsystematic error.

To further investigate the relation of mountain to the forecasting errors, we consider the meridional wind component. The result indicates that the larger forecast errors occur over a different place from the zonal wind forecasts. The largest forecast errors in land spread along the west slope of mountain

This difference is also evident looking at a comparison of zonal wind from meridional wind (not shown) that the larger forecast error in Himalaya mountain disappears in the meridional wind field.

To summarize, the performance of surface air is different from upper air; 2-m temperature forecast error caused by systematic error is more associated with the elevation of terrain, in contrast, similar to 10-m wind speed, upper air forecast error is dominated by the nonsystematic error, which is not closely related to configuration of terrain.

5. Impact of Satellite data Assimilation

In terms of the previous results, it is very clear that the model errors for most of forecast variables except for the 2-m temperature are dominated by the nonsystematic errors, which are caused by uncertainties in the model initial condition or unresolveable differences in scales between the forecasts and observations (Nutter and Manobianco 1999). So that the model initial condition is one of very important factors to impact on the model forecast error. By the many previous studies (Tracton et al. 1980; Halem et al. 1982; Andersson et al. 1991; Mo et al. 1995; Derber and Wu 1998; Bouttier and Kelly 2001) indicates that the assimilation of satellite radiance observations into a numerical weather prediction (NWP) system is an important path to improve weather forecasts by providing initial conditions representative of the true state of the atmosphere. For purpose of understanding the role of initial condition in the accuracy of forecasts, the satellite observation data assimilation will be considered in this section.

a. Data Assimilation system

There are two basic approaches to assimilate satellite information into a data assimilation system (DAS). The first is to assimilate retrieved data from radiances measured by satellite instruments. The second is to assimilate radiance measurements directly into a DAS. Direct radiance assimilation is theoretically superior to retrieval assimilation because the observational error statistics are more justified in direct radiance assimilation than in retrieval assimilation (Eyre et al. 1993; Derber and Wu 1998; Mcnally et al., 2000). The latter will be used in current study.

Developed by Joint Center for Satellite Data Assimilation (JCSDA), the Gridpoint Statistical Interpolation (GSI) analysis system is linked to the ARW WRF mesoscale system and the ATOVS (e.g., The Advanced Television and Infrared Observation Satellite (TIROS)-N Operational Vertical Sounder) observation used. The GSI analysis system is being developed based on NCEP current three-dimensional variational analysis (3DVAR) system known as Spectral Statistical Interpolation (SSI) (Parrish and Derber 1992).

b. Satellite data

The Satellite data is from ATOVS datasets supplied by NESDIS. The ATOVS is composed of Advanced Microwave Sounding Unit (AMSU) and High-Resolution Infrared Sounder (HIRS)/3. Two separate radiometers (AMSU-A and AMSU-B) compose the AMSU platform. The AMSU-A is a cross-track, stepped-line scanning total power radiometer. The instrument has an instantaneous field-of-view of 3.3° at the half-power points providing a nominal spatial resolution at nadir of 48 km. The AMSU-B is a cross-track, continuous line scanning, the total power radiometer has an instantaneous field-of-view of 1.1° (at the half-power points). Spatial resolution at nadir is nominally 16 km. The antenna provides a cross-track scan, scanning ±48.95° from nadir with a total of 90 earth fields-of-view per scan line.

These AMSU-A and AMSU-B radiance being used in current study have undergone substantial preprocessing by NESDIS before becoming available for usage. The data have been statistically limb corrected (adjusted to nadir) and surface emissivity corrected in the microwave channels. It is clear that NOAA-16 data cover the most of southwest Asia areas, the AMSU-B has more intensity of observation than the AMSU-A.

Derber and Wu (1998) pointed out that a presence of a single data point containing large errors can result in substantial degradation of the analysis and subsequent forecast. For this reason, a simple quality control has been developed and the observed brightness temperature data have been modified empirically with various parameters for different instruments. In the GSI analysis

system, the check will include two steps. First, a location check (including removal of observations outside the domain) and thinning procedure (excluding location/time duplicates and incomplete observations) will be performed to ensure vertical consistency of upper-air profiles. Secondly, numerous quality control (QC) checks are redone based on various quality parameters after the modeled brightness temperature was obtained through the radiative transfer model. These quality parameters are used in terms of the expected observational error variance as a function of channels and have been adjusted by the position across the track of the scan, whether it is over land, sea, snow, sea ice, a transition region, elevation, the difference between the model and the real topography, and the latitude. The statistics show that the number used in the GSI regional data assimilation system is quite different. The AMSU-B has much more than the other two AMSU-A. Firstly, we will look at the details for each day. In NOAA-15, the maximum number of AMSU-B for all 30 days is from 50000 to 150000 pixels, and for AMSU-A, the number is only around 40000 pixels. In NOAA-16, the number in AMSU-B exceeds 150000 pixels, while the AMSU-A is under 60000 pixels. Secondly, on average for 30days, the evidence shows that through two QC steps, the final assimilated number of radiance reduces substantially. The used percentage in AMSU-A is over 40%, but in AMSU-B is only 16%.

It is obvious that bias correction and quality control toss out many bad data. It is a benefit to the minimization procedure in data assimilation systems, however, due to the imperfections inherent in bias correction and quality control schemes, a lot of valuable observation data has been tossed out. Future studies should continue to refine good bias correction and quality control schemes.

c. Experiments design

If the previous forecasts described in section 2 excluding satellite data assimilation is expressed as NODA, the current forecasts including satellite data assimilation is referred as DA, which is the same as the NODA except for the initial condition updated by the AMSU-A and AMSU-B radiance data assimilation using the GSI analysis system.

d. Results analysis

To investigate the effect of ATOVS radiance data assimilation on the forecasts over the study areas, two statistical variables - mean square error skill scores and standardized Z statistic – are calculated against the analysis data.

1) Mean-square-error skill scores

By the previous study (Murphy 1988), forecasting skill scores are generally defined as measures of the relative accuracy of two forecasts, one of the two forecasting system is a "reference system". For current experiments, the NODA forecasts can be taken as the reference system. Based on the mean-square-error, the skill score (SS) can be expressed as follows:

$$SS\ (d, r, a) = 1 - [RMS^{2}(d, a) / RMS^{2}(r, a)] \qquad (4)$$

Note that SS in (4) is a function of the DA forecasts (d), the NODA reference forecasts (r), and the analyzed quantity (a); the *RMS (d, a) and RMS (r, a)* are defined as equation (2) indicating the root mean-square-error of DA and NODA forecast relative to the analysis, respectively. To the extent, the positive skill is reflected the better performance than the reference forecasts.

Fig. 4 contains the results for the 2-m temperature, and 10-m wind speed over the nine specific locations. For the 2-m temperature, the results indicate that all SS in the different specific

locations is positive for the 24-h and 48-h forecasts, but the SS for 48-h forecasts in the most of areas is much smaller than that for 24-h forecasts. The SS in the north of Iran mountain (B) and west of Himalaya mountain (E) is about 10-20% smaller than that in the lower mountains or plain areas. Compared to Fig. 2, we can find that the forecasts errors in the high mountain areas are mainly from the model systematic errors, the nonsystematic errors make a relative smaller contribution to the total forecasts; the satellite data assimilation, at least for the AMSU-A and AMSU-B radiance, seems not to make a significant contribution to the improve accuracy of surface temperature forecasts in the higher mountain areas.

In contrast, the 10-m wind speed shows a different SS value from the surface temperature, six of nine locations including all high mountain areas (B, D, E) appears negative skill score, it means the satellite data assimilation produces negative impacts, but the SS in Arabian Sea gets by 45% and 35% for 24-h and 48-h forecasts, respectively.

From these statistical results, it is suggested that assimilation of satellite radiance produces better forecasts of surface temperature and wind speed. However, the improvement is mainly limited to the forecasts over lower terrain or water areas.

2) Standardized Z statistic

In order to determine if the initial condition updated by satellite data assimilation leads to a statistically significant change in forecast accuracy, a standardized Z statistic (Walpole and Meyers 1989) is calculated for a given parameter and compared with the normal distribution using a 99% confidence level. The Z statistic is defined as follows:

$$Z = \frac{\sum_{i=1}^{N} (X_{2i} - X_{1i})}{\{N[(SD_1)^2 + (SD_2)^2 - (SD_{12})^2]\}^{1/2}} \qquad (5)$$

$$SD_{12}^2 = \frac{1}{N}\sum_{i=1}^{N} (X'_{1i} - \overline{X'_1})(X'_{2i} - \overline{X'_2}) \qquad (6)$$

Where, X, X' and SD represents the same as equation (1) ~ (3), the subscripts 1 and 2 denote variables from the NODA and DA forecasts verifying at 24-hour and 48-hour, respectively. Other notation is that SD_{12} denotes the covariance of two forecasts.

For purposes of understanding the statistical significant of data assimilation influence, we firstly look what the impacting is from the data assimilation. The difference between DA and NODA forecast at 24-h and 48-h forecasting has been calculated to represent the impacting. For the 24-h forecasts (Fig. 4), the 2-m temperature DA forecasts in the most of Southwest Asia region including Saudi Arabia, Syria, Iraq, Iran, Afghanistan and Pakistan are on average 0° to 2°C warmer than the NODA forecasts, the larger impacting occurs at south or southwest slope off the Hindu Kush mountain areas (compared against Fig. 1). Over the 48-h forecast, the positive difference reduces substantially (not shown). Secondly, using a 99% confidence level, value of the standardized Z statistic that lie outside the critical value of ±2.58 indicates that the data are able to support a statistically significant. Measured by this standard, the Z (Fig. 5) displays that the positive difference only in southeast of Iran are not statistically significant. The nonsystematic errors limited within Afghanistan and Pakistan areas contribute to a substantial portion of the total model error.

Comparison of the difference in 10-m wind speed for 24-h and 48-h forecasts (not shown) reveals that the large impacting locates over Arabian Sea, Red Sea, Persian Gulf and border of Pakistan and Afghanistan, these places are beyond areas of high terrain mountain. However, the statistical significant regions are only over a few places in the northwest Arabian Sea, borders of Pakistan and Afghanistan for 24-h forecasts. There is almost no place getting over statistical significant for 48-h forecasts. If the impacting of the satellite data assimilation on 2-m temperature occurs nearby the mountain region, the impacting on the 10-m wind speed happens on the places far away from these mountains, especially over water areas.

For upper air, the detailed statistical results of 500 hPa temperature and geopotential height forecasts are presented that the radiance data assimilation increases both the temperature and geopotential height over the central of Southwest Asia, the standardized Z statistic shows the statistical significant over these region. It is obvious that the upper temperature and geopotential height forecasts impacted by satellite data assimilation are not associated with configuration of terrain. However, the impacting covers the most of portion of the study areas in spite of the impacted region reduced substantially for the 48-h forecasts (not shown).

Consistent with the changes of the temperature forecast, the 24-h forecast of wind fields at 200 hPa are modulated by the radiance assimilation with a same way. The wind forecast shows (not shown) that the westerly (easterly) increases in the north (south); the southerly (northerly) in the west (east). The standardized Z statistic shows the lager anomaly areas exceed the significant statistical test. These results indicate that with radiance data assimilation, the thermal and dynamics structure of upper air has been changed substantially, an extra warm anticyclone center appears over the central Southwest Asia.

6. Summary and Discussion

a. Summary

This paper presented an objective verification and impacting of radiance data assimilation on the weather forecast over the complex terrain areas of Southwest Asia using the National Center for Atmospheric Research (NCAR) mesoscale model (ARW WRF) and Joint Center for Satellite Data Assimilation (JCSDA) analysis system (GSI). The experiment period is from May 1 through 31 2006. The results are summarized as follows:

The model biases caused by inadequate parameterization of physical processes, except for the temperature at 2 meter, are relatively small compared to these nonsystematic errors results partially from initial condition uncertainty. The total forecast errors at surface show a substantial spatial heterogeneity, there is a relative bigger error in higher mountain areas. However, the resource of error indicates a significant difference between temperature and wind speed, the error in temperature at 2-m is mainly from systematic error, which is controlled largely by the physical representation; meanwhile, the errors are positively correlation with elevation of terrain. In contrast, the errors in 10-m wind speed get more contribution from nonsystematic error, which is probably related to uncertainty of initial condition.

The performance of upper air is different from surface temperature. Similar to the 10-m wind speed, upper air forecast error is dominated by the nonsystematic error, which is not closely related to configuration of terrain.

The ATOVS satellite data provides useful information for improvement of initial conditions, and the model error got reduced substantially within the 24 hour forecasts. The mean square error skill score (SS) shows that assimilation of satellite radiance produce better forecasts of

surface temperature and wind speed. However, the improvement is mainly limited to the forecasts over lower terrain or water areas.

The standardized Z statistic indicates that with radiance data assimilation, the thermal and dynamics structure of upper air has been changed substantially, an extra warm anticyclone center appears over the central Southwest Asia. However, the upper air forecasts impacted by satellite data assimilation are not associated with geographic location of high terrain.

b. Discussion

In this study, the weather forecasts using the ARW WRF system is evaluated over the Southwest Asia mountain areas. Due to the complexity of the high terrain and lack of knowledge in the estimation of physical processes, the limitations of forecasts for this model should be paid more attention.

First of all, parameterization of physical processes plays a significant role in forecasting of surface temperature. For the 2-m temperature forecasts, the systematic error component is not only larger than the random errors, but also indicates it is substantial related to the elevation of terrain. In contrast, the random errors are responsible for the forecasting of upper air or 10-m wind fields. The random errors prevent perfect forecast guidance and are caused by a combination of uncertainty of initial condition and unreasonable model scales. The detailed statistical results presented in sections 4 are specific to surface and upper-air at nine specific locations. The basic error characteristics for one forecasting variable vary by the selected locations, and may not be representative of errors at other forecast variables. For example, in a preliminary investigation of temperature errors, the results demonstrated that the maximum 2-m temperature biases occurred over the high mountain areas while the temperature bias on 500 hPa were found over the most of southwest Asia, it do not related to configuration of terrain.

Note that the results presented here are referred to the NCEP global analysis system, the value of forecasts performance need to be verified fatherly by the real-time station observation. As expressed by Manning and Davis (1997), "These statistics would provide additional information to model users and alert model developers to those research areas that need more attention." The additional and complementary need for verification strategies in ARW WRF models is discussed in following papers.

Secondly, the random error is very complicated, it is possible partially from the uncertainty of initial condition, a good initial condition might help users to compare the latest forecast guidance with current observations and make appropriate adjustments in real time. The assimilation of satellite radiance observations into a numerical weather prediction (NWP) provide initial conditions representative of the true state of the atmosphere. Current results show that positive impacts of satellite data on weather prediction in the most of Southwest Asia areas, but the impacts are not as obvious in the high terrain areas, such as Himalaya Mountain and Iran Mountain regions. This feature implies that the random error is not only from the uncertainty of initial condition, the other reason like the resolution of model horizontal scale is necessary to be considered. This issue will be discussed in the other future work.

Acknowledgements. The authors thank Dr. Barker, Dr. Liu from NCAR and Dr. Russ Treadon from JCSDA for answering many questions regarding the ARW WRF and GSI data assimilation system.

Reference

Agrawala, S., M. Barlow, H. Cullen, and B. Lyon, 2001: The drought and humanitarian crisis in Central and Southwest Asia: a climate perspective. *Published by International Research Institute for climate prediction (IRI).* Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, 10964, USA, IRI Special Report 01-11.

Andersson, E., A. Hollingsworth, G. Kelly, P. Lönnberg, J. Pailleux, and Z. Zhang, 1991: Global observing system experiments on operational statistical retrievals of satellite sounding data. *Mon. Wea. Rev.,* **119,** 1851–1864.

Bouttier F. and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **127**, Part B, 1469-1488.

Cheng Y. Y. and W. James Steenburgh, 2005: Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the western United States. *Wea. Forecasting*, **20**, 812–821.

Christopher Davis, Barbara Brown, and Randy Bullock, 2006: Object-based verification of precipitation forecasts. Part I: methodology and application to mesoscale rain areas. *Mon. Wea. Rev*., **134**, 1772–1784.

Derber J. C. and W-S Wu. 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299.

Evans, J.P., Smith, R., 2001. Modelling the climate of South West Asia. *Proceedings International Congress on Modelling and Simulation*, MODSIM01, F. Ghassemi et al. (eds.) Australian National University, Canberra, Australia, December 10-13.

Evans, J.P., Smith, R.B., and Oglesby, R.J., 2004. Middle East climate simulation and dominant precipitation processes. *Inter. J Climat.,* **24**, 1671-1694.

Eyre, J. R., G. Kelly, A. P. McNally, E. Andersson, and A. Persson, 1993:Assimilation of TOVS radiances through one dimensional variational analysis. *Quart. J. Roy. Meteor. Soc.,* **119,** 1427–1463.

Gallus W. A. and James F. Bresch, 2006: Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts. *Mon. Wea. Rev*., **134,** 2632–2641

Halem, M., E. Kalnay, W. E. Baker, and R. Atlas, 1982: An assessment of the FGGE satellite observing system during SOP-1. *Bull. Amer. Meteor. Soc.,* **63,** 407–429.

Lorenz, Edward N., 1963: The mechanics of vacillation. *J. Atmo. Sci*., **20**, 448–465.

Manning, K. W., and C. A. Davis, 1997: Verification and sensitivity experiments for the WISP95 MM5 forecasts. *Wea. Forecasting*, **12**, 719–735.

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, 2001: Development of a next generation regional weather research and forecast model. *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology,* W. Zwieflhofer and N. Kreitz, Eds., World Scientific, 269–276.

McNally, A. P., J. C. Derber, W. Wu, and B. B. Katz, 2000: The use of TOVS level-1b radiances in the NCEP SSI analysis system. *Quart. J. Roy. Meteor. Soc.,* **126**, 689–724.

Mo, K. C., X. L. Wang, R. Kistler, M. Kanamitsu, and E. Kalnay, 1995: Impact of satellite data on the CDAS–reanalysis system. *Mon. Wea. Rev.,* **123,** 124–139.

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev*., **116**, 2417–2424.

Nutter P. A. and J. Manobianco, 1999: Evaluation of the 29-km Eta model. Part I: objective verification at three selected stations. *Wea. Forecasting*, **14**, 5–17.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

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 Tech. Note NCAR/TN-468+STR, 94 pp.

Tracton, M. S., A. J. Desmarais, R. J. van Haaren, and R. D. McPherson, 1980: The impact of satellite soundings on the National Meteorological Center's analysis and forecast system—The Data Systems Test results. *Mon. Wea. Rev.,* **108,** 543–586.

Walpole, R. E., and R. H. Meyers, 1989: Probability and statistics for engineers and scientists. Macmillan, 765 pp.

Table 1 The height of topography (Hgt: meter), vegetation type (Veg) and soil type (Soil) in the specific nine subregions (defined as Fig. 1) over SWA.

	A	B	C	D	E	F	G	H	I
Hgt	328	2557	737	3833	4839	958	67	0	75
Veg	barren	grass	barren	Shrub land	wooded tundra	barren	barren	water	Dry land
Soil	loam	loam	clay loam	loam	loam	sandy loam	loam	water	loam

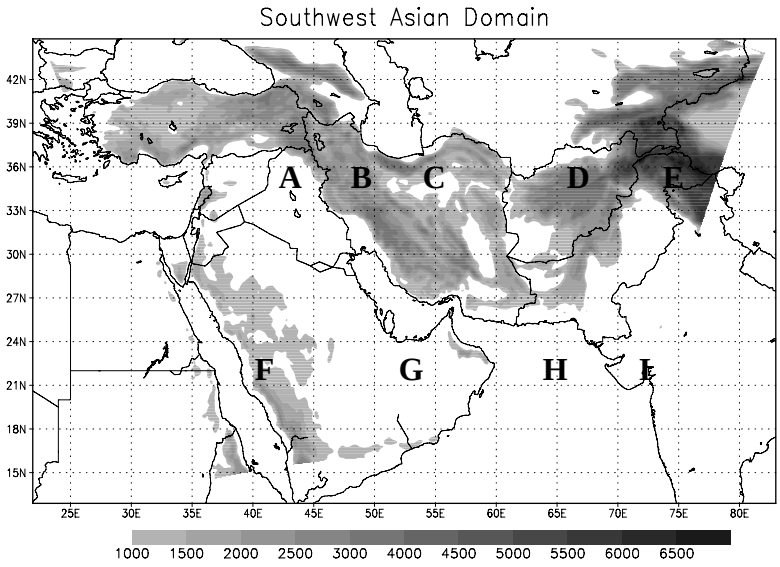

Figure 1 Domain of model and subregion definition. Shaded indicates the elevation of terrain. The subregions are defined as north of Iraq (A; 34°–36°N, 41°–43°E); northwest of Iran (B; 34°–36°N, 46°–48°E); north central Iran (C; 34°–36°N, 54°–56°E); central Afghanistan (D; 34°–36°N, 66°–68°E); west of Himalaya Mountain (E; 34°–36°N, 74°–76°E); west of Saudi Arabia (F; 22°–24°N, 41°–43°E); east of Saudi Arabia (G; 22°–24°N, 51°–53°E); Arabian Sea (H; 22°–24°N, 63°–65°E); and west of India (I; 22°–24°N, 70°–72°E)

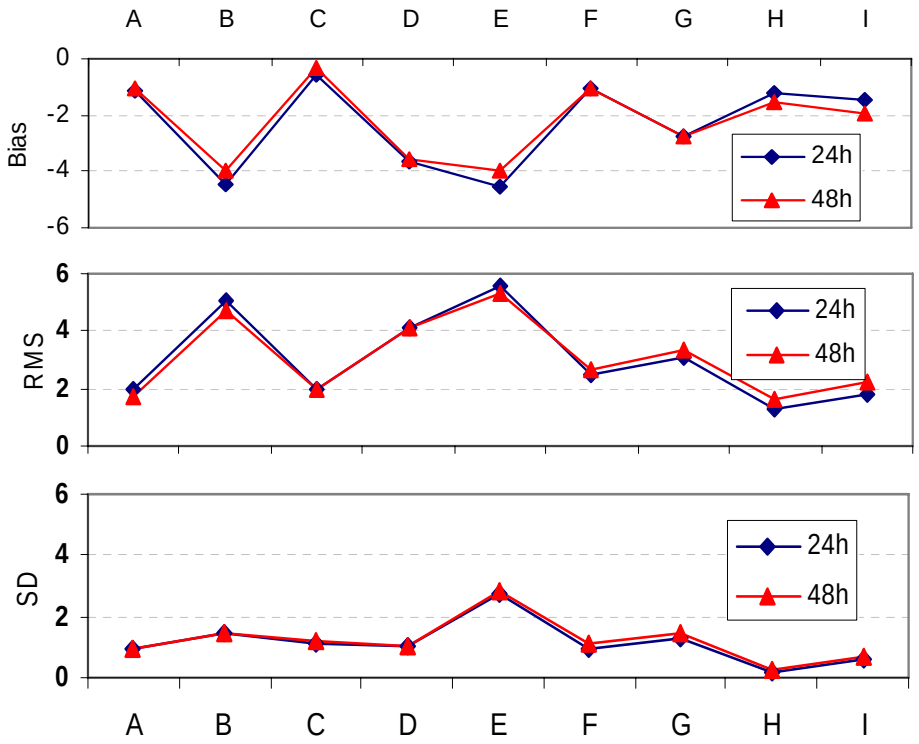

Figure 2 Bias, root mean square (RMS) error and standard deviation(SD) error for 2-m temperature (°C) forecasts from 0Z for May 1 through 31, 2006. Results are plotted for averaged 24-h and 48-h forecasts as a function of specific location (marked as A, B, C, D, E, F, G, H, I) defined in Fig.1

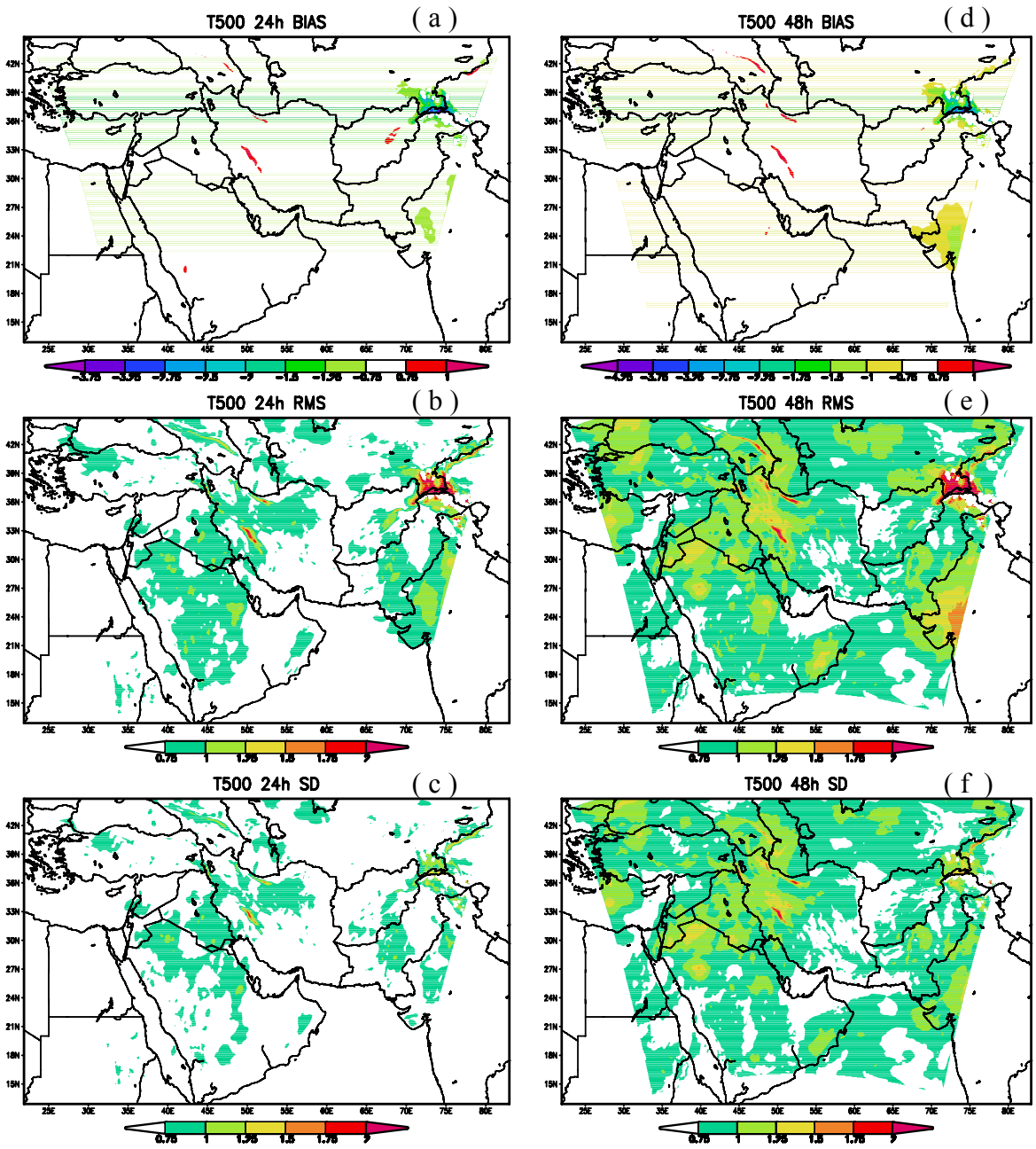

Figure 3 500 hPa temperature. Shaded region indicates the error exceeds 0.75°C.

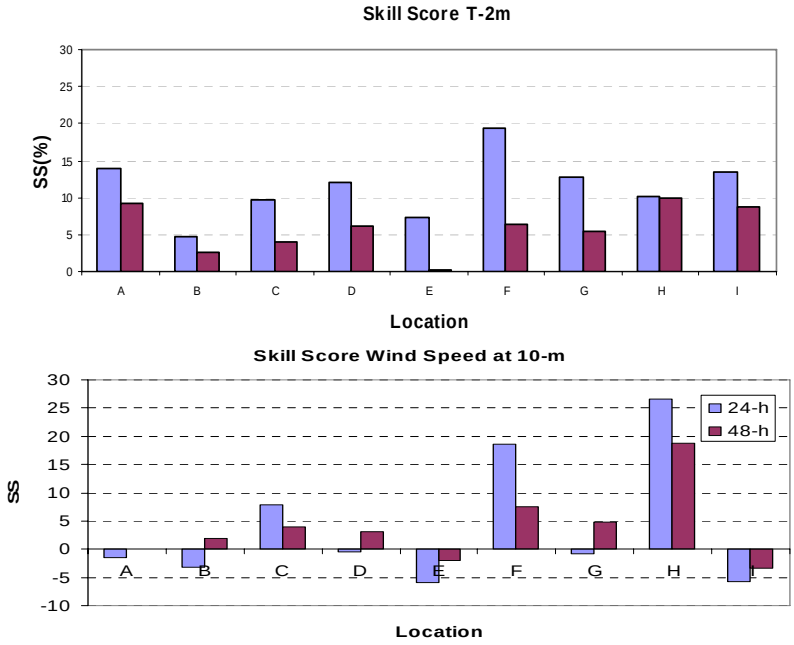

Figure 4 Mean square error skill scores (SS) for 2-m temperature (upper panel) and 10-m wind speed (lower panel). Results are plotted for averaged 24-h and 48-h forecasts as a function of specific location.

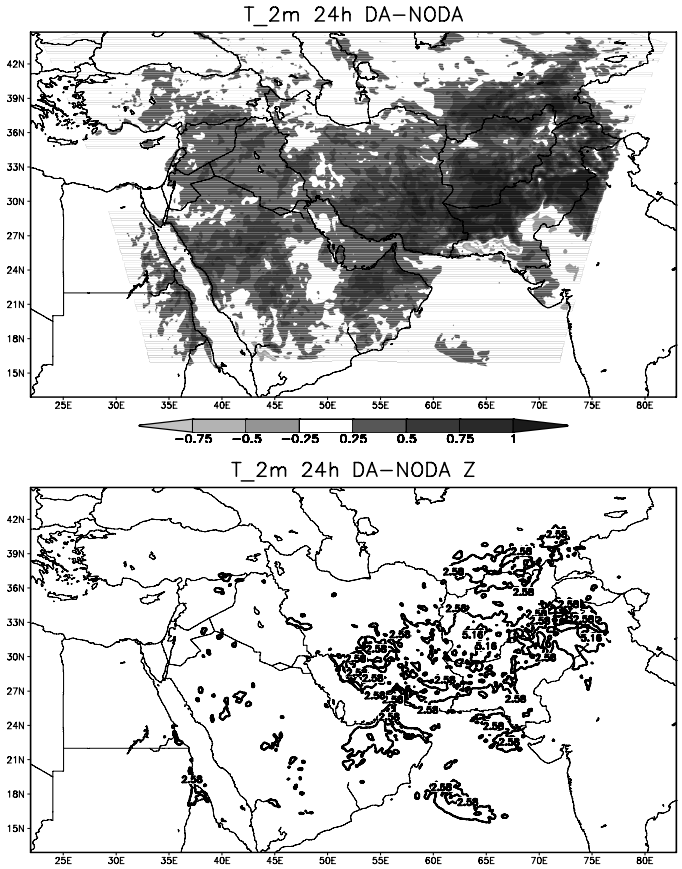

Figure 5 Statistical results in mean 24-h forecasts for 2-m temperature (°C) May 1 through 31 2006. Upper panel: averaged difference between data assimilation (DA) forecasts and Control (CTRL) forecast; Lower panel: standardized Z statistics, the value of Z is outside ±2.58 indicating that the changes caused by data assimilation are considered statistically significant at the 99% confidence level.