JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 113, D09112, doi:10.1029/2007JD009216, 2008

Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model

Jeff Chun-Fung Lo,[1] Zong-Liang Yang,[1] and Roger A. Pielke Sr.[2]

Received 24 July 2007; revised 17 November 2007; accepted 7 January 2008; published 10 May 2008.

[1] The common methodology in dynamical regional climate downscaling employs a continuous integration of a limited-area model with a single initialization of the atmospheric fields and frequent updates of lateral boundary conditions based on general circulation model outputs or reanalysis data sets. This study suggests alternative methods that can be more skillful than the traditional one in obtaining high-resolution climate information. We use the Weather Research and Forecasting (WRF) model with a grid spacing at 36 km over the conterminous U.S. to dynamically downscale the 1-degree NCEP Global Final Analysis (FNL). We perform three types of experiments for the entire year of 2000: (1) continuous integrations with a single initialization as usually done, (2) consecutive integrations with frequent re-initializations, and (3) as (1) but with a 3-D nudging being applied. The simulations are evaluated in a high temporal scale (6-hourly) by comparison with the 32-km NCEP North American Regional Reanalysis (NARR). Compared to NARR, the downscaling simulation using the 3-D nudging shows the highest skill, and the continuous run produces the lowest skill. While the re-initialization runs give an intermediate skill, a run with a more frequent (e.g., weekly) re-initialization outperforms that with the less frequent re-initialization (e.g., monthly). Dynamical downscaling outperforms bi-linear interpolation, especially for meteorological fields near the surface over the mountainous regions. The 3-D nudging generates realistic regional-scale patterns that are not resolved by simply updating the lateral boundary conditions as done traditionally, therefore significantly improving the accuracy of generating regional climate information.

Citation: Lo, J. C.-F., Z.-L. Yang, and R. A. Pielke Sr. (2008), Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, *J. Geophys. Res.*, *113*, D09112, doi:10.1029/2007JD009216.

1. Introduction

[2] Coupled atmosphere-ocean general circulation models (AOGCMs) are numerical tools that are used to perform climate simulations and predict future climate change [*IPCC*, 2007], although skill at this spatial scale has not yet been achieved [*Trenberth*, 2007]. Because of limitations in computational resources, existing AOGCMs typically run at horizontal grid intervals on the order of 200 km, which is far too coarse for applications at regional or local scale regimes at scales of 10–50 km [*Leung et al.*, 2003; *Wang et al.*, 2004; *Giorgi*, 2006]. Thus the nested regional climate modeling technique, also referred to as dynamical downscaling, was developed to mitigate this problem, and currently it has become a common approach to obtain high-resolution regional climate information from AOGCMs [*Giorgi*, 2006].

[3] The starting point of dynamical downscaling is typically a set of coarse-resolution large-scale fields either from AOGCMs or from global reanalysis, which are used to provide the initial (ICs), and lateral meteorological and surface boundary conditions (LBCs) to the nested regional climate model (RCM). The RCM is not intended to modify/correct the large-scale circulation of the AOGCM but is intended to add regional detail in response to regional scale forcing (e.g., topography, coastlines, and land use/land cover) as it interacts with the larger-scale atmospheric circulations [*Giorgi*, 2006]. The purpose of downscaling is to obtain high-resolution detail as accurately as possible over the region of interest.

[4] During the past 20 years, the approaches to the simulation in nested RCM, along with their value-added, have often been debated. *Castro et al.* [2005], for example, concluded that the RCM cannot add skill to simulations of large-scale weather features beyond what is already in the parent global model or reanalysis, since the RCM is so strongly influenced by the parent model or reanalysis. B. Rockel et al. (Dynamical downscalling: Assessment of model system dependent retained and added variability for two different regional climate models, submitted to *Journal of Geophysical Research*, 2008) using a separate model

[1]Department of Geological Sciences, The Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas, USA.
[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.

0148-0227/08/2007JD009216$09.00

from that of *Castro et al.* [2005] confirmed this conclusion. The RCM simulated climate is necessarily strongly dominated by the parent model or reanalysis.

[5] *Castro et al.* [2005] also identified three types of regional climate modeling (defined as when the ICs are forgotten) in which skill becomes progressively more difficult as the parent global input goes from a reanalysis, to a global model with some aspects of the system prescribed (e.g., sea surface temperature) to a global prediction model in which all aspects of the climate system are predicted.

[6] From the experience of numerical weather prediction (NWP), the skill of limited area models diminish very rapidly with time, becoming useless for the simulation period beyond a week. Therefore the very first regional climate simulation was conducted in a short-term re-initialized weather forecasting mode [*Dickinson et al.*, 1989]. The regional climatology was obtained from the statistics of multiple short runs. Over the years, significant development has been achieved in the representation of physical processes in RCM, e.g., atmospheric radiation, cloud microphysics, shallow and deep convection, and land-surface processes [e.g., *Leung and Ghan*, 1998; *Liang et al.*, 2001], to improve the representation of these processes within RCMs in long simulations. The simulation approach of the RCM switches from re-initialization mode into a long-term continuous climate prediction mode, and the simulation length is much longer than few days as is the case in weather forecasting mode. Previous studies with nested RCMs have been run for integration times from a monthlong [*Giorgi and Bates*, 1989; *Giorgi*, 1990] to multidecadal simulations [*Machenhauer et al.*, 1998; *Deque et al.*, 2005].

[7] The long-term continuous simulation, with one single initialization of large-scale fields and frequent updates of LBCs, is currently the most common approach in regional climate simulation [*Leung et al.*, 2003], although the problems of this approach are well documented [*Davies*, 1976; *Warner et al.*, 1997]. For example, the flow developing within the RCM domain may become inconsistent with the driving LBCs [*von Storch et al.*, 2000], and the internal solution generated by RCMs appears to vary with size of the simulation domain [*Castro et al.*, 2005], as well as location and season [*Miguez-Macho et al.*, 2004].

[8] By subdividing the long-term continuous integration into short ones, re-initialization has been successful in weather forecasting to mitigate the problem of systematic error growth in long integrations. Although the re-initialization approach is widely accepted in weather forecasting and some studies did point out that this approach can be beneficial to RCM simulations [*Druyan et al.*, 2001; *Qian et al.*, 2003], it is rarely used in regional climate simulations. There are three reasons for this. First, the re-initialization approach may not be easily portable. Additional scripts are needed to handle the re-initialization process (e.g., file I/O operations). Second, the long spin-up time of RCMs constrains the re-initialization frequency. Third, there may be discontinuity points when results are applied to a transport model. Typically it takes a few hours to a few days for the driving ICs and LBCs to reach dynamical equilibrium with the internal model physics in RCMs. On the other hand, for the soil components, the spin-up time may take a few weeks to a year [*Chen et al.*, 1997].

[9] Recently, some RCMs also use nudging or relaxation of large-scale atmospheric circulations within the interior of the computational domain of the RCM. This has proved useful to prevent the RCMs from drifting away from the large-scale driving fields [*Mabuchi et al.*, 2002; *Miguez-Macho et al.*, 2004]. *von Storch et al.* [2000] also demonstrated that the spectral nudging method was successful in keeping the simulated states closes to the driving state at large scales, while generating small-scale features. Despite these useful features, nudging is not widely accepted in the RCM community.

[10] Both re-initialization and nudging essentially increase the utilization of the large-scale coarse-resolution data during the RCM integration, beyond the first initialization and frequent updates of the LBCs as done conventionally. For the purpose of regional climate downscaling, we hypothesize that such utilization in the RCM can generate realistic regional structures not resolved by the coarse-resolution forcing data. We introduce a new methodology to systematically evaluate the skill of three common regional climate downscaling approaches: 1) long-term continuous integration, 2) consecutive integrations with frequent re-initializations, and 3) as 1) but with a 3-D spatial nudging being applied. We use the $1^\circ \times 1^\circ$ National Centers for Environment Prediction (NCEP) Global Final Analysis (FNL) to drive a regional model WRF for the conterminous U.S. (CONUS) with a grid spacing at 36 km. We evaluate the simulations by comparing with the 32-km NCEP North American Regional Reanalysis (NARR) data set [*Mesinger et al.*, 2006]. This evaluation methodology is slightly different from the Big-Brother Experiment (BBE) raised by *Denis et al.* [2002]. In our study here, the Perfect Big Brother (PBB) is represented by the 32-km NARR; the NARR data set also serves as observations to which other RCM runs will be compared. Different from the BBE, our forcing of the RCM runs are not provided by the upscaled PBB. Instead, we use one-degree FNL as the Perfect Parent (PP) to drive the RCM runs, and generate the high-resolution Little-Brother (LB) simulation. The comparison of the climate statistics of PBB and LB simulations provides the performance assessment of the downscaling ability. Different from other RCM studies, the evaluation statistics are performed in a high temporal scale (6-hourly).

[11] We perform several sensitivity experiments, including different re-initialization frequencies and different extensions of applying nudging, to investigate the effects of different simulation approaches on the regional climate downscaling skill. We present an optimal method that can be used for increasing the accuracy of regional climate downscaling. Our study is demonstrated for the CONUS in the entire year of 2000, but it can be readily applied elsewhere and for other times.

2. Model Description and Experimental Setup

[12] The regional meteorological model used in this study is the Weather Research and Forecasting (WRF) Model with Advanced Research WRF (ARW) dynamic core version 2.2 [*Skamarock et al.*, 2007]. WRF is a next-generation, limited-area, non-hydrostatic, with terrain following eta-coordinate mesoscale modeling system designed to serve both opera-

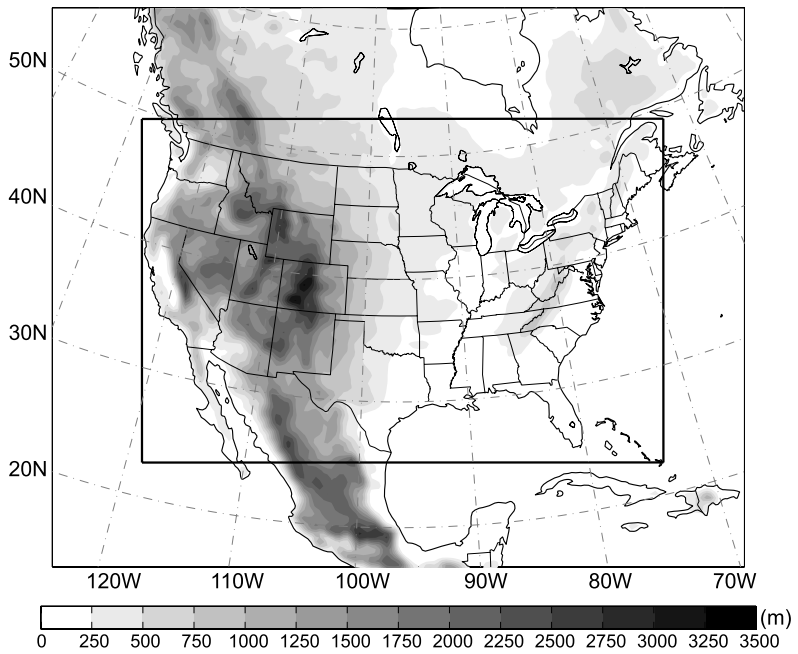

Figure 1. WRF model domain and topography map, with an outline of the extent of the verification region. Contours represent the terrain height, and contour interval is 250 m.

tional forecasting and atmospheric research needs. We choose WRF in this study because it is being developed and studied by a broad community of government and university researchers; however, there are limited studies of WRF for regional climate applications. The main physical options we used include WRF Single-Moment 5-class (WSM5) microphysical parameterization [*Hong et al.*, 1998, 2004]; the new Kain-Fritsch convective parameterization [*Kain*, 2004]; Dudhia shortwave radiation [*Dudhia*, 1989] and Rapid Radiative Transfer Model (RRTM) longwave radiation [*Mlawer et al.*, 1997]; the Yonsei University (YSU) planetary boundary layer (PBL) scheme [*Noh et al.*, 2003]; and the Noah land surface model [*Chen and Dudhia*, 2001].

[13] The model domain is centered at 39.24°N and 96.52°W with dimensions of 178 × 144 horizontal grids points with spacing of 36 km (Figure 1). It comprises most of North America, and the CONUS is contained in the interior of the domain. The Lambert conformal conic projection is used as the model horizontal coordinates with the standard parallel at 39.24°N. The domain boundaries are located mostly over flat land points or ocean points to avoid vertical interpolation problems due to the differences in topography between the forcing data and WRF. In the vertical, thirty-five terrain-following eta levels (eta level = 1.000, 0.995, 0.983, 0.970, 0.954, 0.934, 0.909, 0.880, 0.845, 0.807, 0.520, 0.468, 0.420, 0.376, 0.335, 0.298, 0.263, 0.231, 0.202, 0.175, 0.150, 0.127, 0.106, 0.088, 0.070, 0.055, 0.040, 0.026, 0.013, 0.000. The vertical coordinate η is defined as $\eta = \frac{p_h - p_{ht}}{p_{hs} - p_{ht}} = \frac{p_h - p_{ht}}{p^*}$, where p_h is hydrostatic component of the pressure, p_{hs} is surface pressure, and p_{ht} is a constant pressure at model top (50 hPa)) from surface to 50 mb are used.

[14] Initial and boundary conditions (ICBCs) for the large-scale atmospheric fields, sea surface temperature (SST), as well as initial soil parameters (soil water, moisture and temperature) are given by the 1° × 1° NCEP Global Final Analysis (FNL). Use of the FNL provides an upper bound on the skill that is achievable with global climate prediction models. The domain specified lateral boundary is composed of a 1-point specified zone and a 4-point relaxation zone. Boundary conditions at the specified zone are determined entirely by temporal interpolation from the 6-hourly FNL data. LBCs at the relaxation zone are nudged toward the FNL data following the method of *Davies and Turner* [1977], with higher nudging coefficients for grid points which are closer to the specified zone. Some previous studies suggested larger relaxation zones (10-15 grid points) to reduce the noise generation at the boundaries and details can be found by *Giorgi et al.* [1993]. The SST and green vegetation fraction updated during the simulations.

[15] The evaluation period is from 1 January 2000 to 1 January 2001 and the verification region is over the CONUS. The 32-km NCEP North American Regional Reanalysis (NARR) is used for model validation. The quality of NARR data has been evaluated with surface station and sounding measurements [*Mesinger et al.*, 2006].

3. Reinitialized Integrations Versus Continuous Integration

3.1. Experimental Design

[16] Three experiments using three different types of time integration approach are conducted to investigate the sensitivity of the downscaling skill to the frequency of model re-initialization. The first experiment is the control experiment; a traditional long-term continuous run (referred to as WRFS), where the model is integrated from 0000 UTC 31 December 1999 to 0000 UTC 1 January 2001 without interruption for re-initialization. The initial 24 h is considered as a spin-up period and the outputs during this period are

excluded from the analysis. For the second experiment (referred to as WRFM-30D), the model is re-initialized every 29 days, and each re-initialization runs for 30 days whose total integration time spans 1 year. The initial 24 h of each re-initialization is considered as a spin-up period and the outputs during this period are excluded from the analysis. In the third experiment (referred to as WRFM-7D), the model is reinitialized every 6 days, and each re-initialization runs for 7 days whose total integration time spans 1 year. Similar to WRFM-30D, the outputs of the preceding 24 h are excluded from the data analysis. The land model for the three experiments run with a continuous integration because the spin-up time for the land surface parameters generally longer than for atmospheric variables, and it takes a few weeks to a year of time for the land surface parameters to reach a dynamic equilibrium [*Chen et al.*, 1997]. The model output at every 6 h is used for model evaluation.

3.2. Mass-Field and Wind Field Verification

[17] Evaluation of the WRF downscaling experiments is performed by a variety of statistical verification techniques over the region shown in Figure 1 by comparing the gridded model values with NARR data. To facilitate comparisons, the 32-km NARR data are bilinearly interpolated onto the 36-km WRF grid. The evolution of root mean square error (RMSE) of the model-simulated surface pressure, 700-mb wind and 500-mb heights are displayed in Figure 2. The RMSE gives an overview of the accuracy of simulations and the formula is defined as [*Anthes*, 1983]

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2\right]^{1/2} \tag{1}$$

where N is the total number of verification grid point and p and o are the model simulation and observed values. The wind vector RMSE is given by [*Anthes*, 1983]

$$WRMSE = \frac{1}{N}\left[\sum_{i=1}^{N}\left[({}_up_i - {}_uo_i)^2 + ({}_vp_i - {}_vo_i)^2\right]^{1/2}\right] \tag{2}$$

where u and v subscripts denote the u and v wind.

[18] Figure 2 suggests that different integration approaches appear to have a significant effect on the skill of regional climate downscaling. The 7-day re-initialization run (WRFM-7D) generally performs better than the monthly (WRFM-30D) and the continuous (WRFS) runs in downscaling the mass-field and wind field atmospheric variables, and the continuous run shows the largest error. The monthly run (WRFM-30D), which involved more frequent re-initializations, generally performs better than WRFS for the modeled-surface pressure, 700-mb wind and 500-mb height. The mean WRFM-7D RMSE of the surface pressure (Figure 2a) was reduced by 9.2%, from 6.5 to 5.9 mb, for WRFS. The improvement of applying re-initialization is even greater in the upper atmosphere. The mean WRFM-7D RMSEs of 700-mb wind (Figure 2b) and 500-mb heights (Figure 2c) are reduced by 18.3% from 6.0 to 4.9 m s^{-1}, and 30.2% from 37.8 to 26.4 m, respectively, comparing with WRFS. We have also verified temperature and moisture fields and the conclusion is the same as that of mass and wind fields. This result generally implies that the more frequent the re-initialization, the more accurate the regional climate downscaling for the atmospheric variables pressure, temperature, wind, and moisture.

3.3. Precipitation Verification

[19] Here we focus on the model performance in simulating precipitation; a model diagnostic field which is more difficult to simulate. The accuracy of the simulated precipitation is determined by the following measures for 24-h intervals over the land areas of the CONUS (Figure 1) throughout the whole year of 2000. The first quantity is the bias score (*BS*), which indicates whether the model over or underpredicts the fractional areal coverage of precipitation for a certain threshold. The threat score (*TS*) is used to measure the skill of predicting the area of precipitation for a certain threshold. The *BS* and *TS* are defined as

$$BS = \frac{P}{O} \tag{3}$$

$$TS = \frac{H}{(P + O - H)} \tag{4}$$

where P is the number of grid points in which the threshold amount of precipitation was simulated, O is the number of grid points that the threshold amount was observed, and H is the number of grid points that threshold precipitation was both simulated and observed. The threshold amounts used is 10 mm. The mean error (*ME*), indicates whether the model over or under predicts the mean magnitude of precipitation, defined as

$$ME = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i) \tag{5}$$

where N is the total number of verification grid point and p and o are the model simulation and observed values. The fourth quantity is the correlation coefficient (*CC*), which indicates the strength and direction of a linear relationship between the simulation and observed values, and is given by

$$CC = \frac{\frac{1}{N-1}\left[\sum_{i=1}^{N}[(p_i - \bar{p})(o_i - \bar{o})]\right]}{\left[\frac{1}{N-1}\left[\sum_{i=1}^{N}[p_i - \bar{p}]^2\right]\right]^{1/2}\cdot\left[\frac{1}{N-1}\left[\sum_{i=1}^{N}[o_i - \bar{o}]^2\right]\right]^{1/2}} \tag{6}$$

where a bar denotes the mean of that variable.

[20] Figure 3 shows the model skill on simulating precipitation over the verification region [the land areas of the CONUS (Figure 1)]. It displays an annual oscillation in the precipitation skill score: the model generally overpredicts the precipitation throughout the year 2000, both in term of fractional areal coverage (*BS*) and mean magnitude (*ME*), except in summer (June–September). Since most of the precipitation in summer is associated with convection, it suggests the Kain-Fritsch convective parameterization produces too few cumulus clouds, thereby resulting in an underprediction of precipitation (Figures 3a and 3b) and a

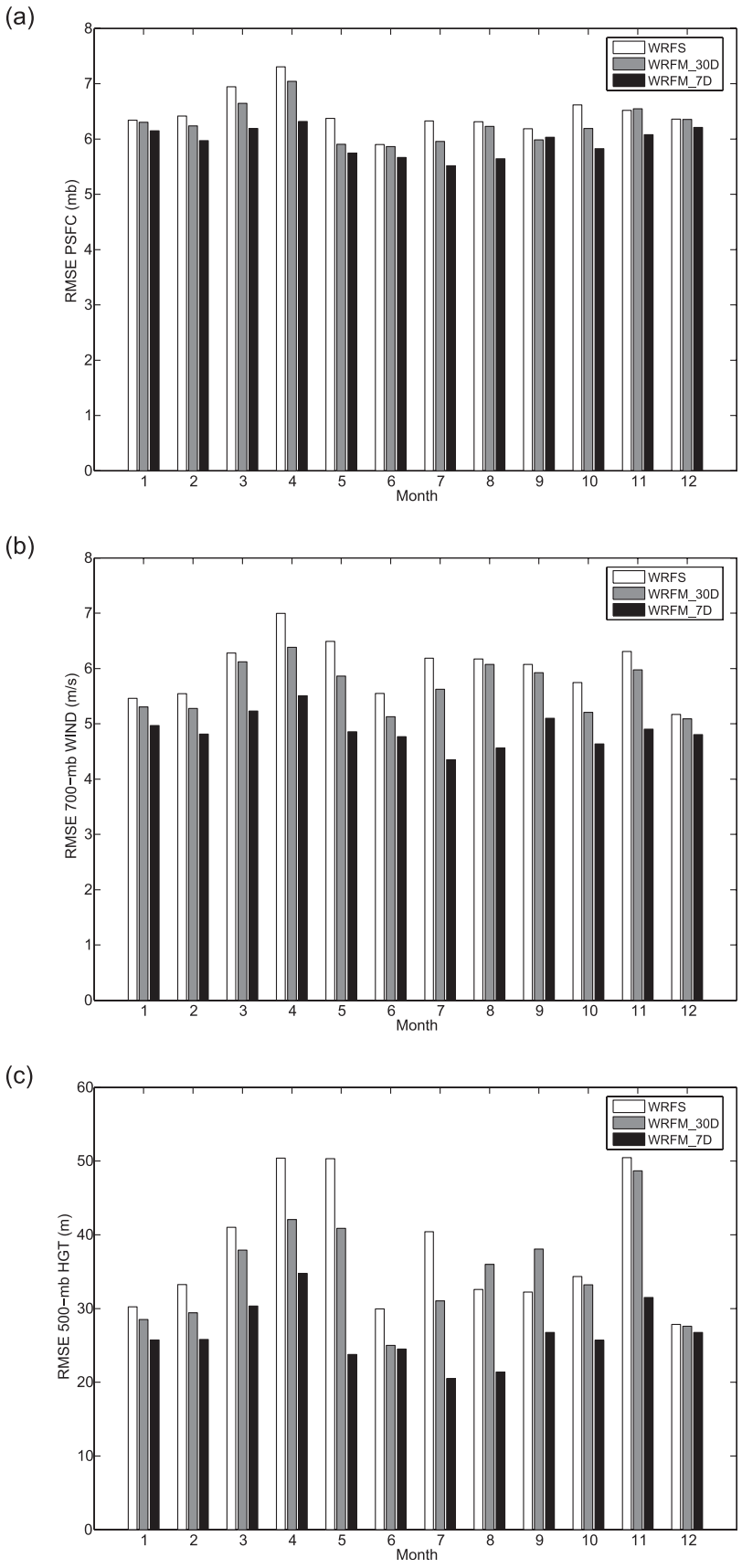

Figure 2. Monthly root-mean-square errors for re-initialization experiments during Jan–Dec 2000. (a) Surface pressure (mb), (b) 700-mb wind (m s^{-1}), (c) 500-mb geopotential height (m).

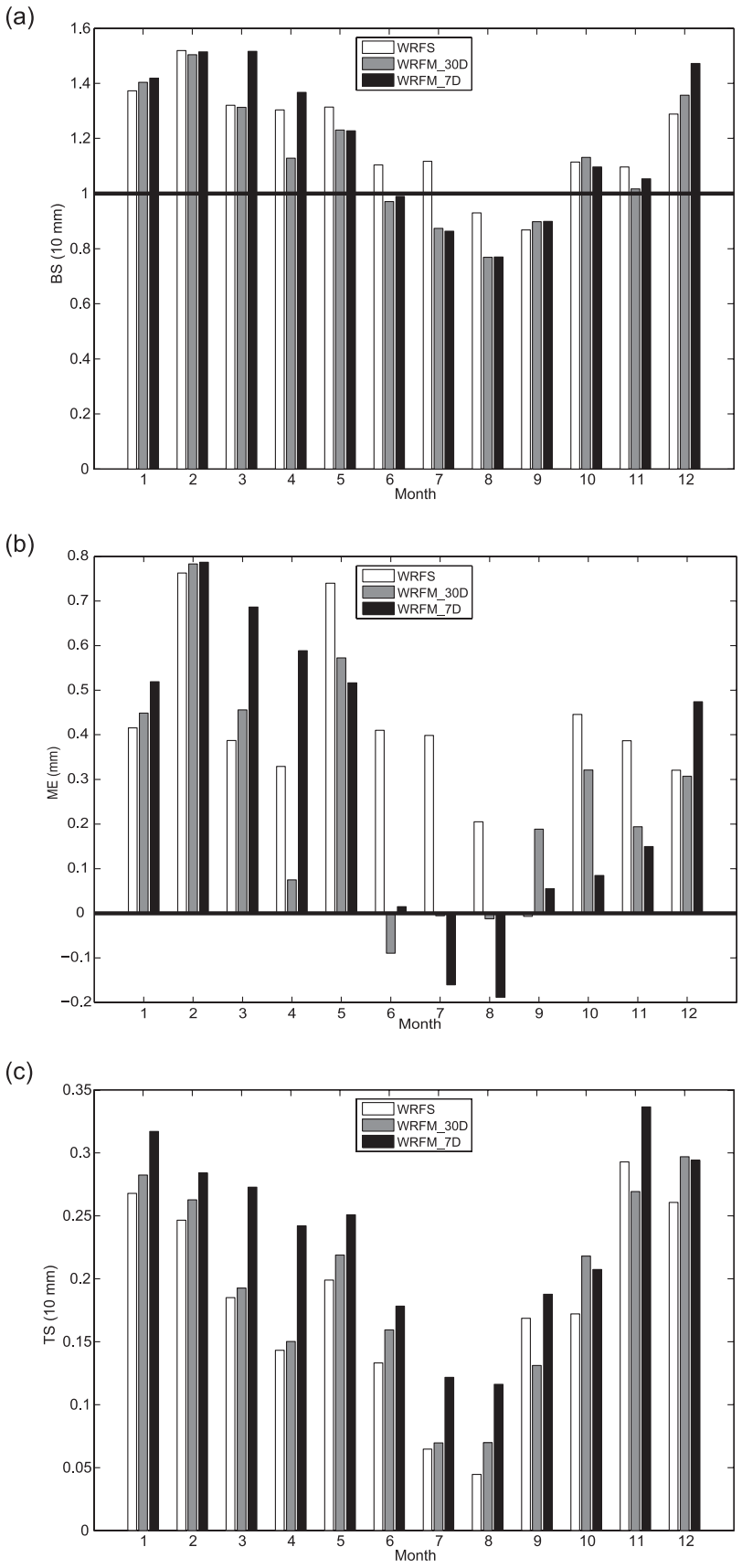

Figure 3. Precipitation skill scores for re-initialization experiments during Jan–Dec 2000. (a) Bias scores at the threshold $\geq$ 10 mm, (b) mean errors, (c) threat scores at the threshold $\geq$ 10 mm.

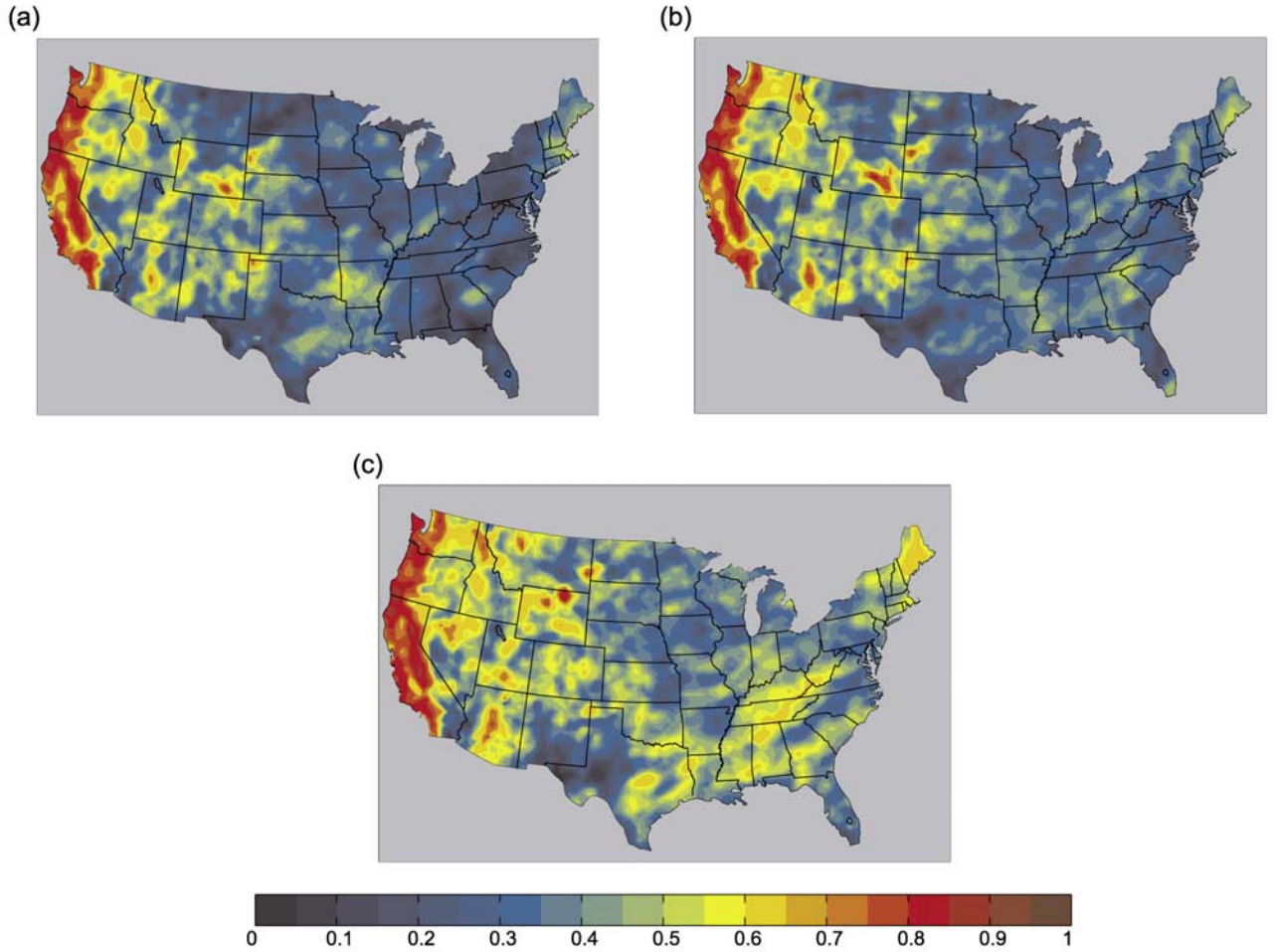

Figure 4. Correlation between time series of 24-h-accumulated precipitation simulations and observed values from NARR at every grid point during Jan–Dec 2000. (a) WRFS, (b) WRFM-30D, and (c) WRFM-7D.

relative low *TS* (Figure 3c) during summer. Comparing the three re-initialization experiments, as a result of better simulating the atmospheric variables of pressure, temperature, wind, and moisture, WRFM-7D generally shows a relatively higher *TS* (Figure 3c) than that produced by WRFM-30D and WRFS. However, for *BS* (Figure 3a, perfect = 1) and *ME* (Figure 3b, perfect = 0), there is no strong overall advantage for WRFM-7D.

[21] The correlation coefficient map displayed in Figure 4 gives a spatial overview of the model performance in simulating precipitation. WRFS performs very well ($CC > 0.8$) for the precipitation simulation over the Pacific Northwest and the western coast (Figure 4a), while the simulation is not good ($CC < 0.4$) over the Midwestern and eastern coast. The corresponding cumulative probability of the correlation coefficient maps in Figure 4 is shown in Figure 5. The curves in Figure 5 plot the probability that are equal to or less than a given *CC*, i.e., the value on the vertical axis for a point P is the probability ($CC \leq P$). From Figure 5, 80% of the verification region ($P = 0.8$) has $CC \leq 0.48$ for WRFS. WRFM-30D performs slightly better than WRFS and 80% of the verification region has $CC \leq 0.49$. WRFM-7D performs the best, with a high *CC* over the eastern and southeastern United States, and 80% of the verification region with $CC \leq 0.56$.

[22] In summary, more accurate downscaling simulations can be obtained using re-initializations to constrain the large-scale circulation by limiting the error growth in the model during the long-term simulation. The orographic precipitation over the Pacific Northwest appears to be captured easier than the large-scale and convective precipitation over the Midwestern and eastern coast. This suggests that the precipitation over the Northwest is dominated by large-scale weather features which are well resolved by the FNL, which when imposed on the higher spatial resolution of terrain in WRF, produces an improved precipitation simulation.

4. Experiments With Analysis Nudging

4.1. Experimental Design

[23] The method of Newtonian relation or nudging, as first introduced by *Charney et al.* [1969], relaxes the model state toward the observed state by adding, to one or more of the prognostic equations, an artificial tendency terms based on the difference between the two states [*Stauffer and*

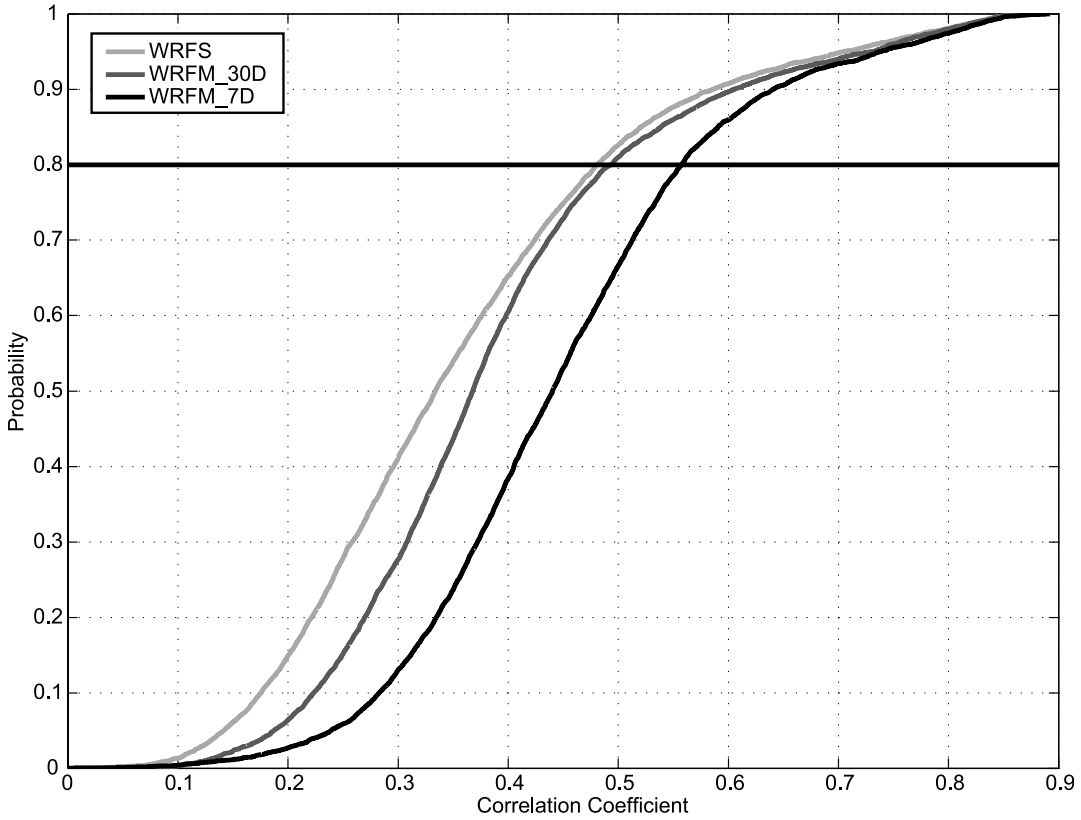

Figure 5. Cumulative probability of the distribution of correlation coefficients depicted in Figure 4 for re-initialization experiments during Jan–Dec 2000.

Seaman, 1990]. The model solution can be nudged toward either gridded analyses (analysis nudging) or individual observations (observational nudging) during the period of time surrounding the observations. Here we apply analysis nudging to the downscaling simulations to investigate the effect of assimilating the large-scale driving fields throughout the integration. The analysis-nudging term for a given variable is proportional to the difference between the model simulation and FNL analysis calculated at every grid point within the model domain. Different from the nudging method used by *von Storch et al.* [2000], the nudging method in this paper is formulated in the 3-D spatial domain instead of the spectral domain. In WRF, the predictive equation of variable $\alpha(x, t)$ mass weighted by pressure p^* is written as

$$\frac{\partial p^*\alpha}{\alpha t} = F(\alpha, x, t) + G_\alpha \cdot W_\alpha \cdot \varepsilon_\alpha(x) \cdot p^*(\hat{\alpha}_0 - \alpha) \qquad (7)$$

[24] The model's physical forcing terms (advection, Coriolis effects, etc.) are represented by F, where α represents the model's dependent variables, x represents the independent spatial variables, and t is time. The nudging coefficient G_α determines the magnitude of the term relative to all the other model processes in F. Its spatial and temporal variation is determined mostly by the four-dimensional weighting function, W, which specifies the horizontal, vertical and time weighting applied to the analysis, where $W = w_{xy}w_\sigma w_t$. The analysis quality factor, ε, which ranges between 0 and 1, is based on the quality and distribution of the data used to produce the gridded analysis. The estimate of the observation for α analyzed to the grid is $\hat{\alpha}_0$. The variables nudged are horizontal wind, temperature, and moisture. The pressure and geopotential height are not assimilated explicitly in the experiments. The nudging coefficient $G_\alpha = 3 \times 10^{-4}$ s^{-1} [*Stauffer et al.*, 1985; *Stauffer and Seaman*, 1990] is chosen.

[25] Two sensitivity experiments are conducted to investigate the impact of nudging to constrain the large-scale circulation within the RCM. The first experiment applies the full 3-D nudging throughout the whole atmospheric column (referred to as WRFS-FDDA-ALL). The second one applies the nudging throughout the whole atmospheric column excluded from the boundary layer (referred to as WRFS-FDDA-NOPBL). The other model settings including the model domain, relaxation zone, physics options, initial conditions and boundary conditions, and the simulation period are identical to the control experiment. The continuous integration approach is used for all the nudging experiments.

4.2. Verification of Mass and Wind Fields

[26] Figure 6 shows the evolution of RMSE of the model-simulated surface pressure, 700-mb wind and 500-mb heights for the nudging experiments. Compared with the traditional approach (WRFS), all the nudging experiments show significant improvements in the downscaling skill. For example, the mean WRFS-FDDA-ALL RMSE of the surface pressure (Figure 6a) was reduced by 20.0%, from 6.5 to 5.2 mb, for WRFS. The overall advantage of applying

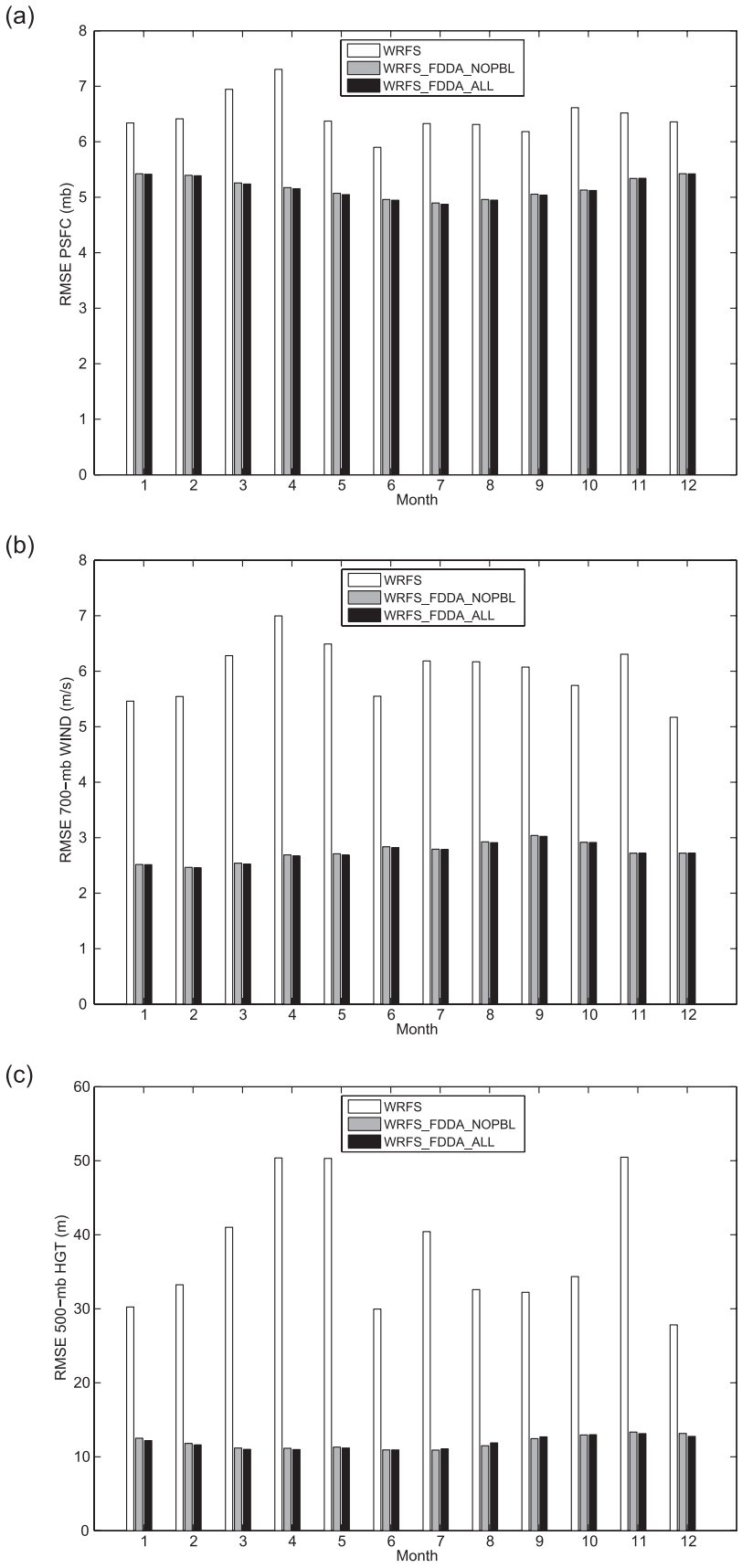

Figure 6. Monthly root-mean-square errors for nudging experiments during Jan–Dec 2000. (a) Surface pressure (mb), (b) 700-mb wind (m s^{-1}), (c) 500-mb geopotential height (m).

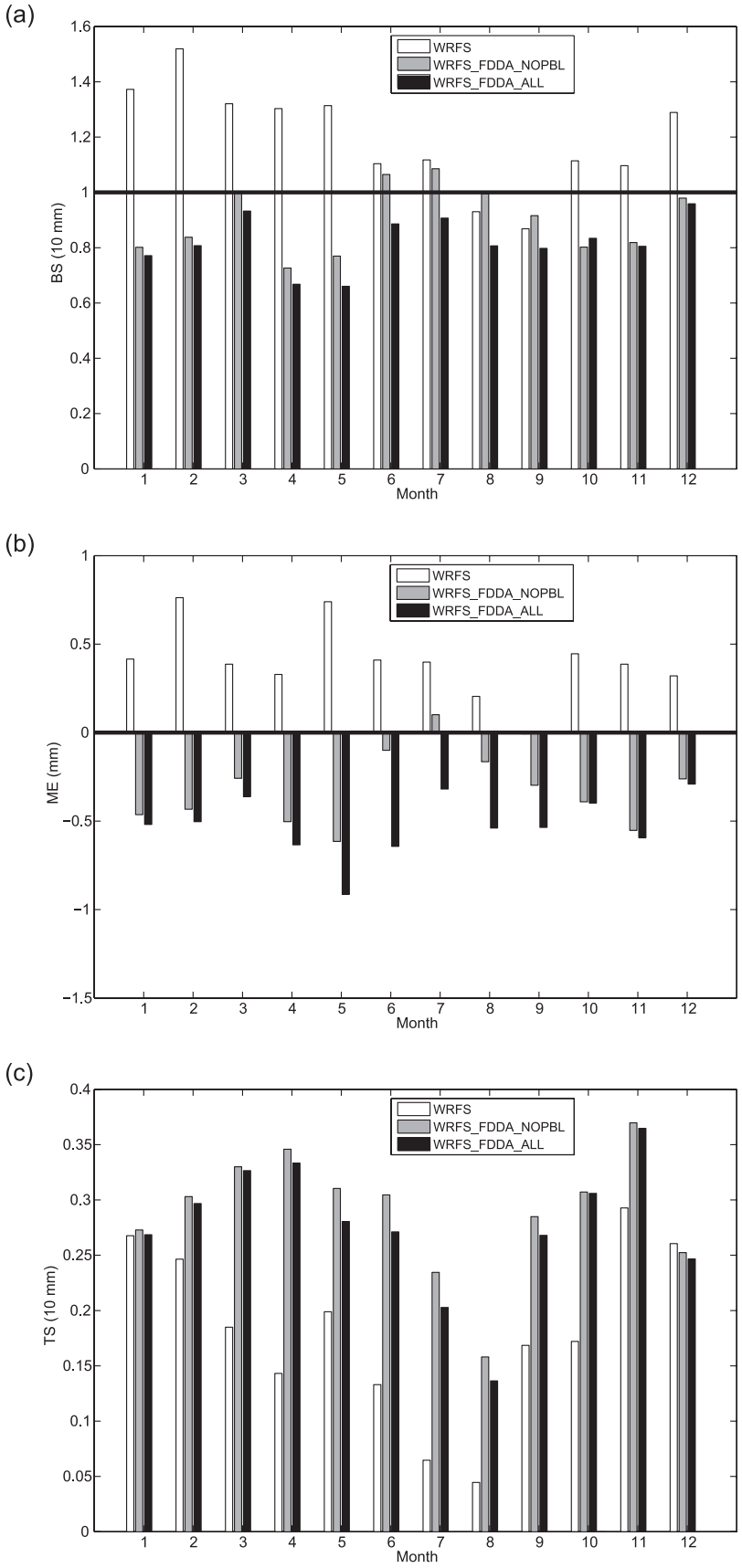

Figure 7. Precipitation skill scores for nudging experiments during Jan–Dec 2000. (a) Bias scores at the threshold $\geq$ 10 mm, (b) mean errors, (c) threat scores at the threshold $\geq$ 10 mm.

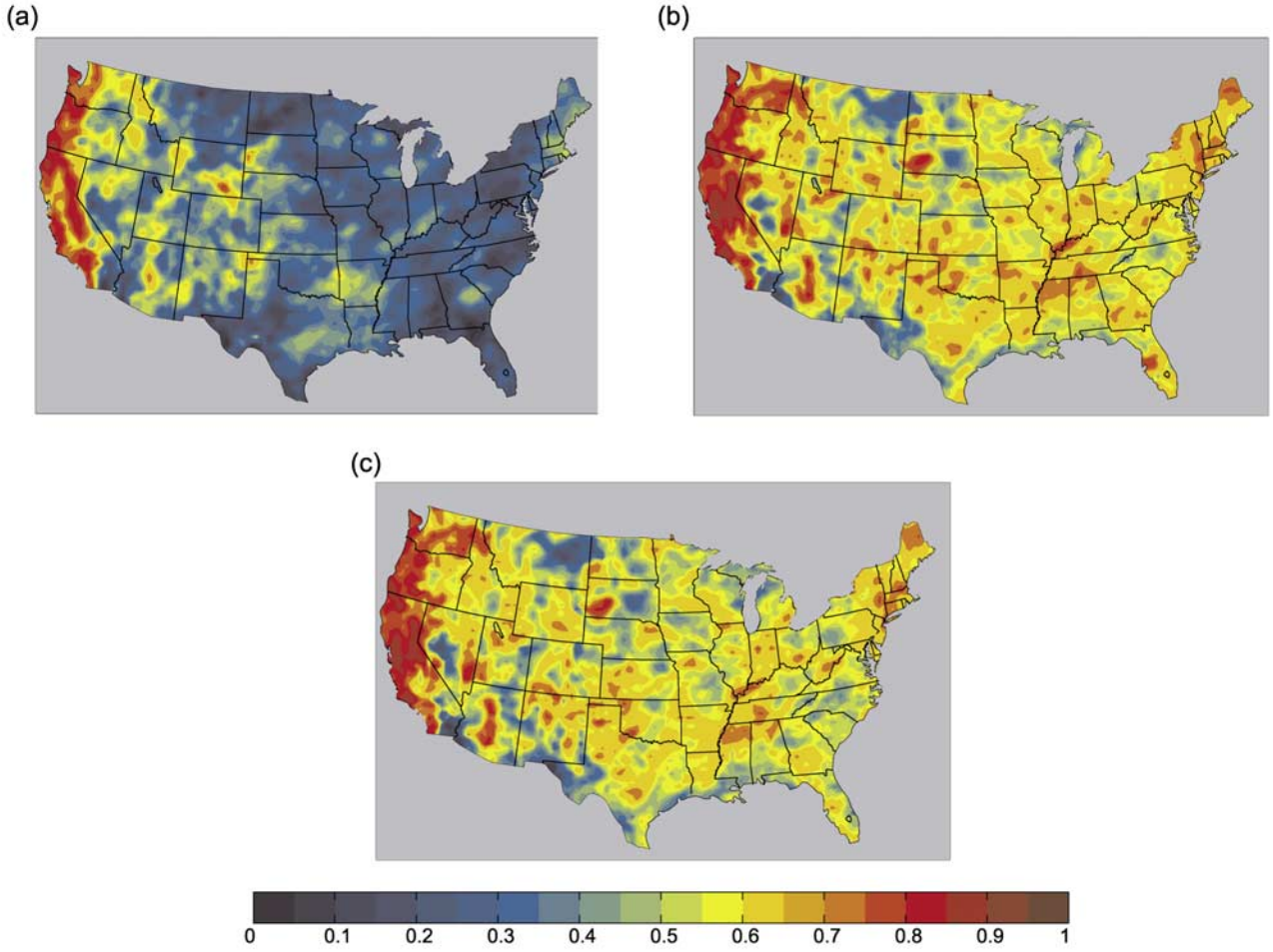

Figure 8. Correlation between time series of 24-h-accumulated precipitation simulations and observed values from NARR at every grid point for Jan–Dec 2000. (a) WRFS, (b) WRFS-FDDA-NOPBL, and (c) WRFS-FDDA-ALL.

nudging is even stronger in the upper atmosphere than at the surface. The mean WRFS-FDDA-ALL RMSEs of 700-mb wind (Figure 6b) and 500-mb heights (Figure 6c) are significantly reduced by 55.0% from 6.0 to 2.7 m s^{-1}, and 68.5% from 37.8 to 11.9 m, respectively, compared with WRFS. There is no significant difference between the accuracy of the two nudging experiments, of which, WRFS-FDDA-ALL performs slightly better than WRFS-FDDA-NOPBL.

[27] We also evaluated the temperature and moisture fields and the conclusion is the same as that of the mass and wind fields. This result generally implies that applying nudging from the coarse resolution reanalysis can limit the large-scale error growth and significantly improve the skill of regional climate downscaling for the atmospheric pressure, temperature, wind, and moisture.

4.3. Precipitation Verification

[28] Figure 7 shows the downscaling skill of the nudging experiments on simulating precipitation. The nudging experiments generally improve the model-simulated areal coverage of precipitation (Figures 7a and 7c). However, applying nudging does not improve significantly on the simulation of the mean precipitation magnitude (Figure 7b); the WRFS generally overpredicts the mean precipitation magnitude, while the nudging experiments generally underpredict the mean precipitation magnitude over the CONUS. As reported in the previous section, the two nudging experiments do not result in much difference in simulating pressure, temperature and moisture; however, the effect on the precipitation simulation can be quite large. WRFS-FDDA-NOPBL performs better than WRFS-FDDA-ALL; showing a lower bias in the precipitation simulation both in the fractional areal coverage (Figure 7a) and the mean magnitude (Figure 7b), and also a higher *TS* (Figure 7c). The deterioration in precipitation simulation due to nudging toward surface variables within the model PBL is possible because of large differences in topography between the coarse-resolution FNL and the WRF downscaling domain. WRFS-FDDA-NOPBL does not nudge the model toward the unrealistic surface variables and produces a better result in precipitation simulation.

[29] The correlation coefficient maps displayed in Figure 8 provide a spatial overview of the model performance of the precipitation simulation for the nudging experiments. Both nudging experiments generally perform much

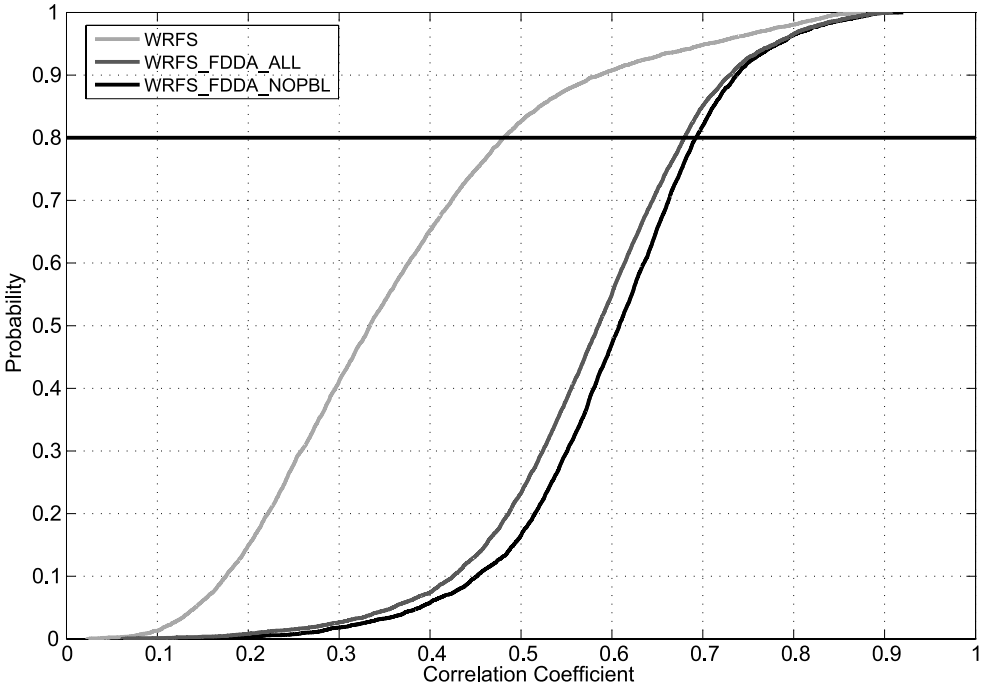

Figure 9. Cumulative probability of the distribution of correlation coefficients depicted in Figure 8 for nudging experiments during Jan–Dec 2000.

better than the WRFS on precipitation simulation over the CONUS. The corresponding cumulative probability of the correlation coefficient map is shown in Figure 9. WRFS-FDDA-NOPBL produces a significant improvement (Figure 8b, 80% area of the verification region with $CC \leq 0.69$) for WRFS (Figure 8a, 80% $CC \leq 0.48$).

5. Skill Enhancement of the WRF Downscaling

[30] One interesting issue we can see from Figures 8b–8c is that although the nudging simulations are constrained by the perfect large-scale circulation and provide an excellent downscaling for the pressure, temperature and moisture, there are some zones where the performance is not good in simulating precipitation (e.g., North American Monsoon area over the south-western US). This result indicates that although using nudging could provide a better simulation of pressure, temperature and moisture, the model physics is still playing an important role in regional climate downscaling especially for simulating precipitation.

[31] In the following, we will examine the skill enhancement of the regional climate downscaling using the WRF model. We compare the model-simulated result with the downscaling result given by directly interpolating the coarse FNL to the 36-km WRF grid (referred to as FNL-INT). Figure 10 shows the evolution of the RMSE for these downscaling experiments. For WRFS, with a higher resolution to resolve the topography and land surface characteristics, the model simulations are better than FNL-INT for the surface variables (Figure 10a). However, this advantage is diminished higher in the atmosphere and the errors in downscaling are larger than directly using the FNL data (Figures 10b–10c). The downscaling skill increases when nudging is applied. The model not only generates realistic regional structures not resolved by the FNL data at the surface (Figure 10a), but also maintains the consistency with large-scale behavior of the FNL forcing data higher in the atmosphere (Figures 10b–10c).

[32] Figure 11 displays the RMSE vertical profile for wind and height given by the downscaling experiments. The traditional downscaling approach (WRFS) shows no skill throughout the atmospheric column. The RMSE is larger than directly using the interpolated FNL data. However, by using nudging, WRFS-FDDA-NOPBL performs even better than FNL over all vertical levels for the nudged variable (wind, Figure 11a). For the non-nudged variable (geopotential height, Figure 11b), downscaling results from WRFS-FDDA-NOPBL are comparable with the quality of the FNL data.

6. Summary and Discussion

[33] Assessment of the methodologies of regional climate downscaling has been performed using a regional model WRF, with initial conditions and boundary conditions driven by coarse-resolution FNL reanalysis. We introduced a new modified Big-Brother Experiment (BBE) [*Denis et al.*, 2002]. Various aspects of the regional-scales downscaling variables have been evaluated against another set of high resolution reanalysis (NARR). Our evaluation has focused over the CONUS during entire 2000, and the evaluation statistics are performed at a high temporal scale (6-hourly). We compare the long-term continuous integration with short-term reinitialized simulations. In addition, we apply

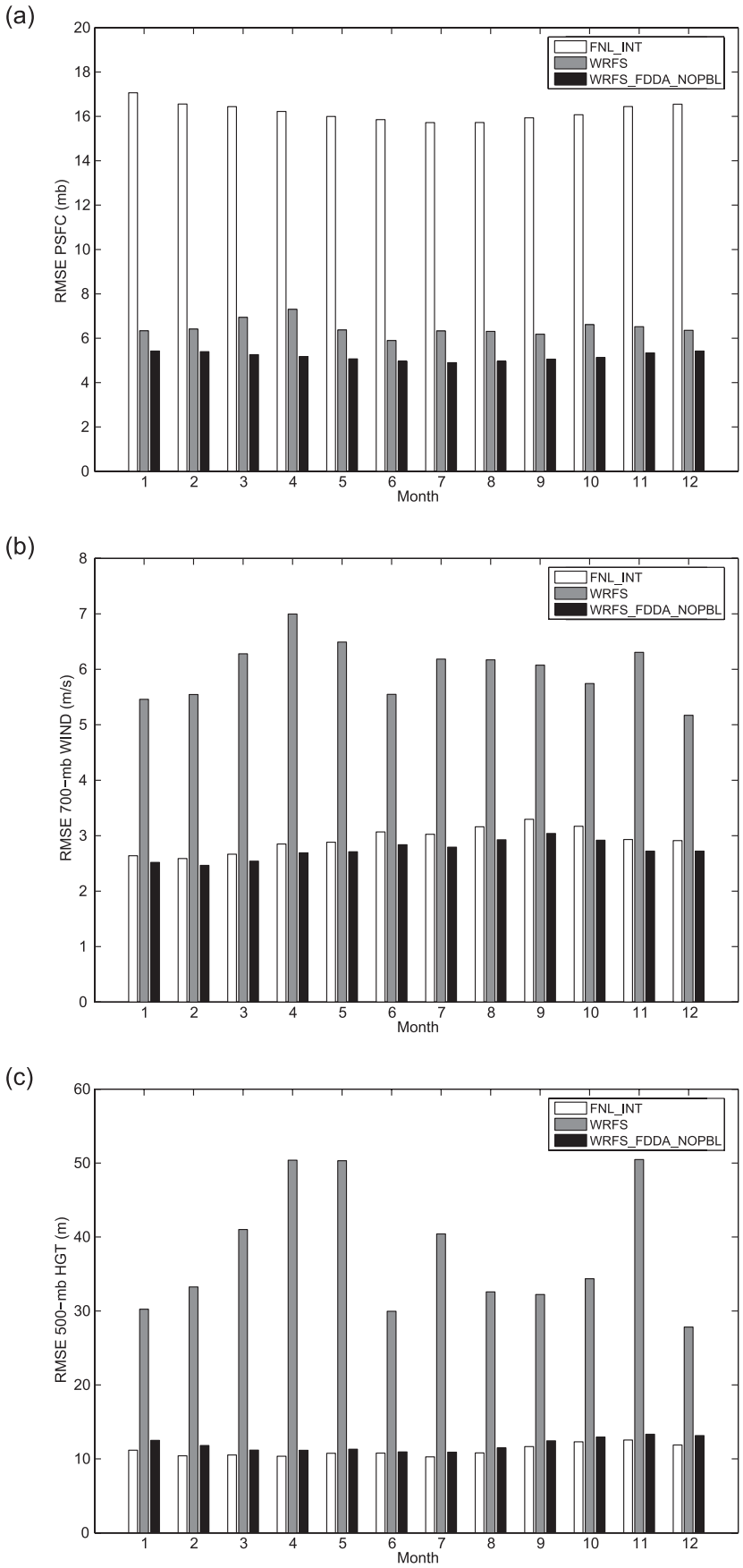

Figure 10. Monthly root-mean-square errors for downscaling experiments during Jan–Dec 2000. (a) Surface pressure (mb), (b) 700-mb wind (m s^{-1}), (c) 500-mb geopotential height (m).

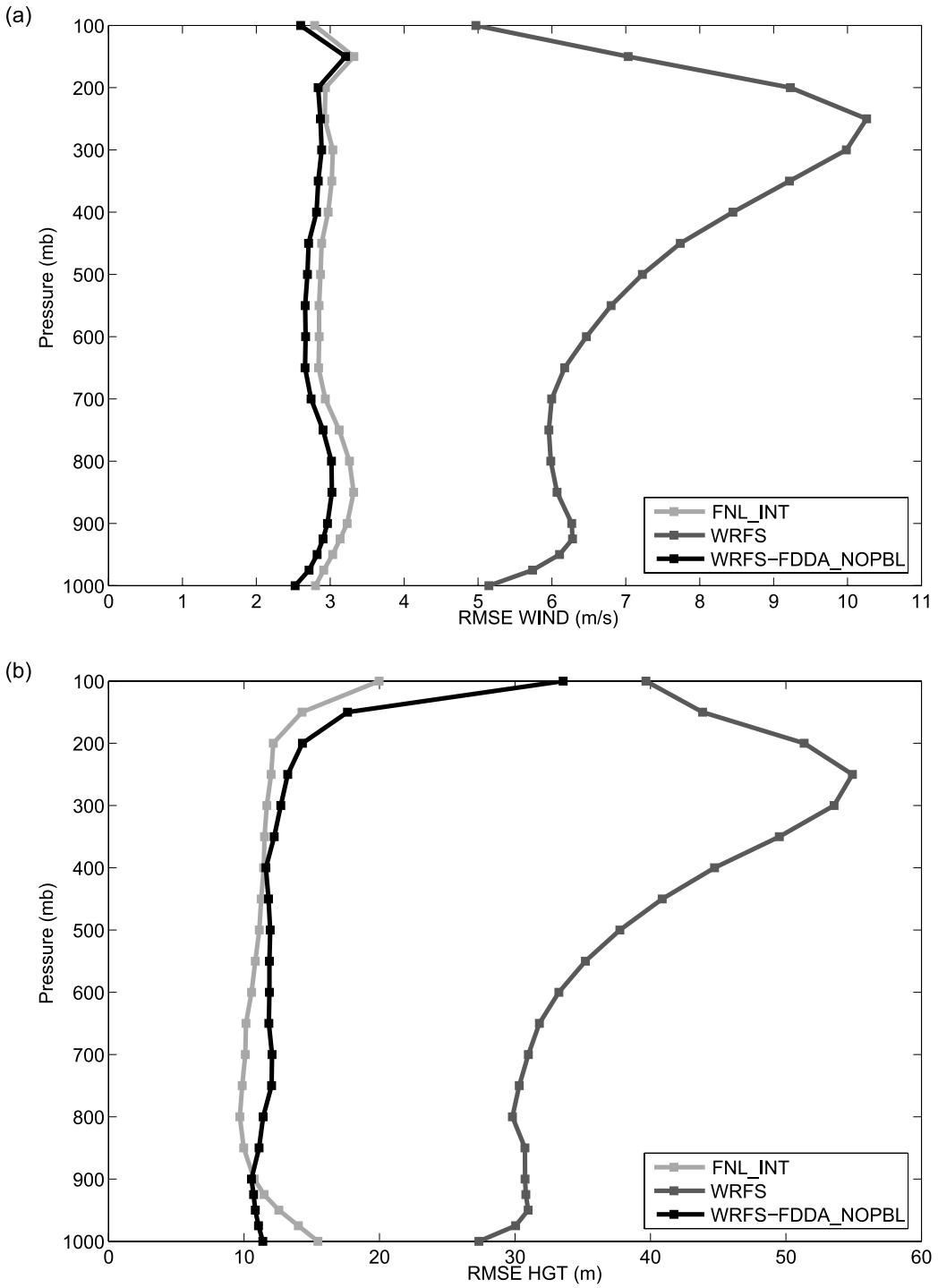

Figure 11. Vertical profile of the root-mean-square errors for downscaling experiments during Jan–Dec 2000. (a) Wind (m s^{-1}), and (b) geopotential height (m).

analysis nudging to the downscaling experiments to study the sensitivity of the downscaling skill to the degree of constraint of the large-scale circulation.

[34] The traditional continuous integration approach, in all cases, shows the worst performance among the downscaling experiments. The model drifts from the forcing FNL reanalysis during the course of long integrations. It poorly simulates not only the forcing variables, (e.g., pressure, temperature, wind, and moisture), but also the model diagnostics variables (e.g., precipitation). Compared with the coarse-resolution FNL, the continuous run performed a reasonable job in downscaling surface parameters because of the more detailed topography. However, for the atmospheric variables above the surface, its performance is even worse than directly using the interpolated FNL.

[35] The re-initialization runs retain sufficient long-term forcing and yield intermediate skill among the downscaling experiments. The re-initialization runs outperform the con-

tinuous simulation runs, while a run with a more frequent (weekly) re-initialization outperforms that with the less frequent re-initialization (monthly). However, the model spin-up problems introduced by each re-initialization should be addressed in the design of an optimal re-initialization frequency. This problem can be eliminated by using a short overlap period. On the other hand, soil parameters which generally have long memory should not be subdivided from a long simulation into shorter ones.

[36] The downscaling simulations using the 3-D analysis nudging, which constrains the error growth in large-scale circulation during the long simulation, show the highest skill. The 3-D nudging simulation generates realistic regional-scale patterns that are not resolved by the coarse resolution FNL which are used to nudge the model, especially for meteorological fields near the surface. It also maintains consistency with the large-scale behavior of the FNL forcing data higher in the atmosphere. We have also performed an additional experiment which is same as the weekly re-initialization experiment (WRFM-7D), except applying full 3-D nudging in each re-reinitialization. The results are the same as the continuous run with nudging, implying it is not necessary to subdivide long integrations into short-term periodic re-initialization when nudging is applied. As a result, 3-D nudging not only improves the accuracy, but also the portability of regional climate downscaling.

[37] Our findings also point out the importance of choosing a suitable downscaling approach. If we choose an unsuitable integration approach (e.g., WRFS), the downscaling simulation may be valueless because the errors of the downscaling simulation could even be larger than those which directly using the interpolated forcing data. The impact of choosing the downscaling approach could thus be larger than the choice of the model physical schemes.

[38] The results show that applying nudging can significantly improve the precipitation simulation. On the other hand, as seen from Figures 8b–8c, nudging does not entirely solve the downscaling problem. There are still some zones where the nudging simulation performs very poor with respect to the precipitation simulation. To improve the downscaling performance in those areas, improvements in the parameterized physics and also the dynamical core of the model are required.

[39] In addition, the strong influence of the nudging on the regional results, as well as the sensitivity of the regional results to the time period between re-initializations, demonstrates the dominance of the larger-scale reanalysis in controlling regional climate simulations. It should also be emphasized that the above conclusions are valid when "perfect" large-scale forcing data (reanalysis) is used, but that they are not necessarily valid for imperfect GCM multidecadal climate predictions. Regional climate simulations that rely on those predictions for LBCs and nudging are thus dominated by the global model information.

[40] **Acknowledgments.** This material is based on work supported by the U.S. Environmental Protection Agency grant RD83145201, the National Aeronautics & Space Administration grant NNX07AL79G, and the John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin. The views expressed in this paper are those of the authors and do not necessarily reflect the views or polices of the EPA and NASA. The authors want to thank the Texas Advanced Computing Center for providing computer resources. R. A. Pielke Sr. received support to complete this study also from the University of Colorado at Boulder (CIRES/ATOC). Dallas Staley is thanked for proofing.

References

Anthes, R. A. (1983), Regional models of the atmosphere in middle latitudes, *Mon. Weather Rev.*, *111*, 1306–1335.

Castro, C. L., R. A. Pielke Sr., and G. Leoncini (2005), Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS), *J. Geophys. Res.*, *110*, D05108, doi:10.1029/2004JD004721.

Charney, J., M. Halem, and R. Jastrow (1969), Use of incomplete historical data to infer the present state of the atmosphere, *J. Atmos. Sci.*, *26*, 1160–1163.

Chen, F., and J. Dudhia (2001), Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, *Mon. Weather Rev.*, *129*, 569–585.

Chen, F., Z. I. Janjic, and K. Mitchell (1997), Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP Mesoscale Eta model, *Boundary Layer Meteorol.*, *85*, 391–421.

Davies, H. C. (1976), A lateral boundary formulation for multi-levels prediction models, *Q. J. R. Meteorol. Soc.*, *102*, 405–418.

Davies, H. C., and R. E. Turner (1977), Updating prediction models by dynamical relaxation: An examination of the technique, *Q. J. R. Meteorol. Soc.*, *103*, 225–245.

Denis, B., R. Laprise, D. Caya, and J. Cote (2002), Downscaling ability of one-way nested regional climate models: The Big-Brother Experiment, *Clim. Dyn.*, *18*, 627–646.

Deque, M., et al. (2005), Global high resolution vs. regional climate model climate change scenarios over Europe: Quantifying confidence level from PRUDENCE results, *Clim. Dyn.*, *25*, 653–670.

Dickinson, R. E., R. M. Errico, F. Giorgi, and G. T. Bates (1989), A regional climate model for the western United States, *Clim. Change*, *15*, 383–422.

Druyan, L. M., M. Fulakeza, P. Lonergan, and M. Saloum (2001), A regional model study of synoptic features over West Africa, *Mon. Weather Rev.*, *129*, 1564–1577.

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.

Giorgi, F. (1990), Simulation of regional climate using a limited area model nested in a general circulation model, *J. Clim.*, *3*, 941–963.

Giorgi, F. (2006), Regional climate modeling: Status and perspectives, *J. Phys. IV*, *139*, 101–118.

Giorgi, F., and G. T. Bates (1989), The climatological skill of a regional model over complex terrain, *Mon. Weather Rev.*, *117*, 2325–2347.

Giorgi, F., M. R. Marinucci, G. T. Bates, and G. D. Canio (1993), Development of a second-generation Regional Climate Model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions, *Mon. Weather Rev.*, *121*, 2814–2832.

Hong, S.-Y., H.-M. H. Juang, and Q. Zhao (1998), Implementation of prognostic cloud scheme for a regional spectral model, *Mon. Weather Rev.*, *126*, 2621–2639.

Hong, S.-Y., J. Dudhia, and S.-H. Chen (2004), A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, *Mon. Weather Rev.*, *132*, 103–120.

IPCC (2007), Climate change 2007: The physical science basis, in *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., 996 pp., Cambridge Univ. Press, Cambridge, U.K.

Kain, J. S. (2004), The Kain–Fritsch convective parameterization: An update, *J. Appl. Meteorol.*, *43*, 170–181.

Leung, L. R., and S. J. Ghan (1998), Parameterizing subgrid orographic precipitation and surface cover in climate models, *Mon. Wea. Rev.*, *126*, 3271–3291.

Leung, L. R., L. O. Mearns, F. Giorgi, and R. L. Wilby (2003), Regional climate research, *Bull. Am. Meteorol. Soc.*, *84*, 89–95.

Liang, X.-Z., K. E. Kunkel, and A. N. Samel (2001), Development of a regional climate model for US midwest applications. Part I: Sensitivity to buffer zone treatment, *J. Climate*, *14*, 4363–4378.

Mabuchi, K., Y. Sato, and H. Kida (2002), Verification of the climatic features of a regional climate model with BAIM, *J. Meteorol. Soc. Jpn.*, *80*, 621–644.

Machenhauer, B., et al. (1998), Validation and analysis of regional present-day climate and climate change simulations over Europe, *MPI Report 275*, *MPI*, Hamburg, Germany.

Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*, 343–360.

Miguez-Macho, G., G. L. Stenchikov, and A. Robock (2004), Spectral nudging to eliminate the effects of domain position and geometry in

regional climate model simulations, *J. Geophys. Res.*, *109*, D13104, doi:10.1029/2003JD004495.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.*, *102*(D14), 16,663–16,682.

Noh, Y., W. G. Cheon, S.-Y. Hong, and S. Raasch (2003), Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data, *Boundary Layer Meteorol.*, *107*, 401–427.

Qian, J.-H., A. Seth, and S. Zebiak (2003), Reinitialized versus continuous simulations for regional climate downscaling, *Mon. Weather Rev.*, *131*, 2857–2874.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers (2007), A description of the advanced research WRF version 2, 100 pp., *NCAR Tech. Note NCAR/TN–468+STR*.

Stauffer, D. R., and N. L. Seaman (1990), Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data, *Mon. Weather Rev.*, *118*, 1250–1277.

Stauffer, D. R., T. T. Warner, and N. L. Seaman (1985), A Newtonian "nudging" approach to four-dimensional data assimilation: Use of SESAME-IV data in a mesoscale model, *Preprints, Seveth Conf. on Numerical Weather Prediction*, Montreal, Amer. Meteor. Soc., 77–82.

Trenberth, K. (2007), Predictions of climate. The climate change blog http://blogs.nature.com/climatefeedback/.

von Storch, H., H. Langenberg, and F. Feser (2000), A spectral nudging technique for dynamical downscaling purposes, *Mon. Weather Rev.*, *128*, 3664–3673.

Wang, Y., et al. (2004), Regional climate modeling: Progress challenges and prospects, *J. Meteorol. Soc. Jpn.*, *82*, 1599–1628.

Warner, T. T., R. A. Peterson, and R. E. Treadon (1997), A tutorial on lateral conditions as a basic and potentially serious limitation to regional numerical weather prediction, *Bull. Am. Meteorol. Soc.*, *78*, 2599–2617.

J. C.-F. Lo and Z.-L. Yang, Department of Geological Sciences, The Jackson School of Geosciences, The University of Texas at Austin, Austin, TX 78712-0254, USA. (liang@mail.utexas.edu)

R. A. Pielke Sr., Cooperative Institute for Research in Environmental Sciences, University of Colorado, Stadium 255-16, Boulder, CO 80309, USA.