Evaluation of WRF and HadRM Mesoscale Climate Simulations over the U.S. Pacific Northwest*

YONGXIN ZHANG AND VALÉRIE DULIÈRE

Climate Impacts Group, Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington

PHILIP W. MOTE

Climate Impacts Group, Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington, and Oregon Climate Change Research Institute, Oregon State University, Corvallis, Oregon

ERIC P. SALATHÉ JR.

Climate Impacts Group, Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington

(Manuscript received 3 October 2008, in final form 14 May 2009)

ABSTRACT

This work compares the Weather Research and Forecasting (WRF) and Hadley Centre Regional Model (HadRM) simulations with the observed daily maximum and minimum temperature (Tmax and Tmin) and precipitation at Historical Climatology Network (HCN) stations over the U.S. Pacific Northwest for 2003–07. The WRF and HadRM runs were driven by the NCEP/Department of Energy (DOE) Atmospheric Model Intercomparison Project (AMIP)-II Reanalysis (R-2) data. The simulated Tmax in WRF and HadRM as well as in R-2 compares well with the observations. Predominantly cold biases of Tmax are noted in WRF and HadRM in spring and summer, while in winter and fall more stations show warm biases, especially in HadRM. Large cold biases of Tmax are noted in R-2 at all times. The simulated Tmin compares reasonably well with the observations, although not as well as Tmax in both models and in the reanalysis R-2. Warm biases of Tmin prevail in both model simulations, while R-2 shows mainly cold biases. The R-2 data play a role in the model biases of Tmax, although there are also clear indications of resolution dependency. The model biases of Tmin originate mainly from the regional models. The temporal correlation between the simulated and observed daily precipitation is relatively low in both models and in the reanalysis; however, the correlation increases steadily for longer averaging times. The high-resolution models perform better than R-2, although the nested WRF domains do have the largest biases in precipitation during the winter and spring seasons.

1. Introduction

The U.S. Pacific Northwest is characterized by mountainous terrain and intricate land–sea contrasts (Fig. 1) resulting in a host of finescale weather systems such as sea and land breezes, rain shadows, and downslope windstorms that define the local weather and climate (Mass 2008). In a warming climate, such finescale weather systems can significantly alter the local temperature and precipitation trends (Salathé et al. 2008) and are essential to consider in climate simulations and climate change assessment at regional and local scales. Global models are generally able to resolve the large-scale weather systems that affect the Pacific Northwest but not the finescale processes associated with the local terrain. To capture these smaller features, a more realistic representation of the local complex terrain and the heterogeneous land surfaces is needed (Mass et al. 2002; Salathé et al. 2008). Therefore, the use of limited-area regional climate models with horizontal resolutions on the order of tens of kilometers is crucial for simulating the regional climate of the Pacific Northwest.

Recently, there have been increasing efforts over the Pacific Northwest in using limited-area mesoscale models for downscaling reanalysis data (Leung et al. 2003a,b)

* Joint Institute for the Study of the Atmosphere and Ocean Contribution Number 1602.

Corresponding author address: Yongxin Zhang, 3737 Brooklyn Ave. NE, Seattle, WA 98105.
E-mail: yongxin@u.washington.edu

DOI: 10.1175/2009JCLI2875.1

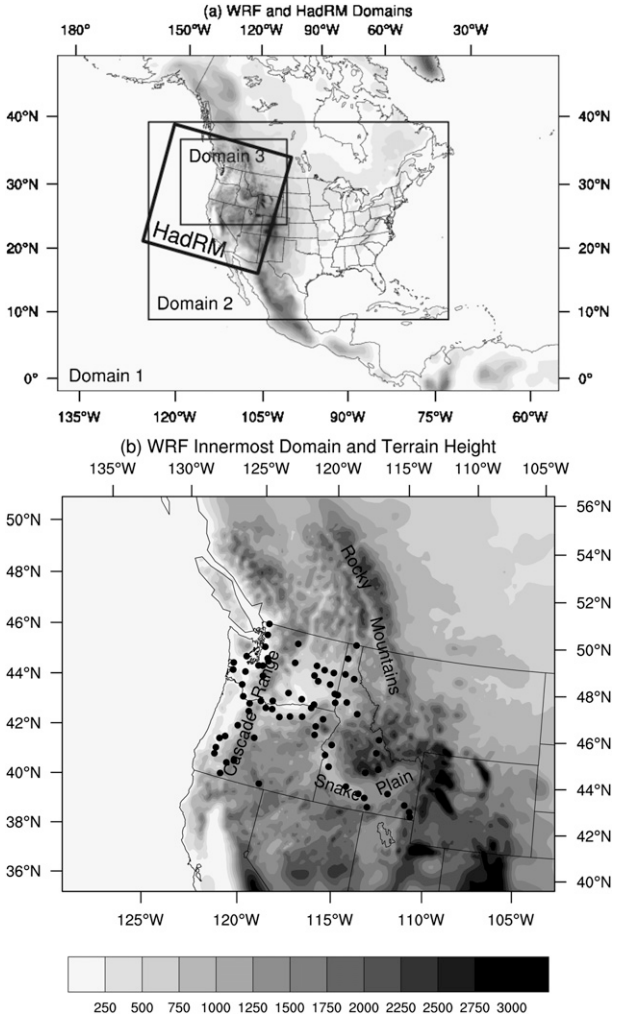

FIG. 1. (a) WRF model domains with two nests (domains 1, 2, and 3) and HadRM domain and (b) WRF innermost domain (domain 3) and terrain height (m). Shadings in (a) represent terrain height (m) for the corresponding WRF domain. Grid spacing for each domain is as follows: WRF domain 1 = 108 km; WRF domain 2 = 36 km; WRF domain 3 = 12 km; and HadRM domain = 25 km. HCN stations in the states of Washington, Oregon, and Idaho are represented by filled black circles in (b).

and global climate model simulations (Salathé et al. 2008) to study regional climate and climate change. Much of the work has been involved with the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5; Grell et al. 1993). As NCAR has been phasing out MM5 and has released the state-of-the-art, next-generation Weather Research and Forecasting Model (WRF; http://www.wrf-model.org/index.php), it is timely to switch from the MM5-based to the WRF-based mesoscale climate modeling and examine its performance over the Pacific Northwest. The Hadley Centre Regional Model (HadRM) is another limited-area regional climate model widely used worldwide as part of the Providing Regional Climates for Impacts Studies (PRECIS) system, which was developed at the Hadley Centre of the Met Office. The PRECIS system can be easily applied to any area of the globe to generate detailed climate change projections. Both the WRF and HadRM models are coupled land–atmosphere modeling systems that can be used to add finescale information to the large-scale projections of global climate models. What differs between these two models is that WRF, like MM5, was designed for short-term weather forecasts and refinements have to be applied in the model to perform long-term climate simulations (Salathé et al. 2008).

In this study, we apply the WRF-based and HadRM modeling systems over the Pacific Northwest in downscaling reanalysis data (Kalnay et al. 1996; Kanamitsu et al. 2002) for the period of 2003–07. The reanalysis data incorporate all the available observations at the time of processing and are generally regarded as best representing the 6-hourly large-scale state of the atmosphere. Reanalysis fields are particularly suitable for driving regional climate models and for model validation. Here, we present validation results based on comparisons to station observations over the Pacific Northwest from the Historical Climatology Network (HCN; Karl et al. 1990). Our purpose is to examine the models' performance in reproducing station observations at various time scales. This work is organized in the following manner. Section 2 contains a brief description of the models. Experimental design is discussed in section 3. Section 4 compares model simulations with the observations. Major conclusions and discussions are presented in section 5.

2. Model description

a. WRF model

The WRF model is a state-of-the-art, next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs (http://www.wrf-model.org). It is a nonhydrostatic model, with several available dynamic cores as well as many different choices for physical parameterizations suitable for a broad spectrum of applications across scales ranging from meters to thousands of kilometers. The dynamic cores in WRF include a fully mass- and scalar-conserving flux form mass coordinate version. The physics package includes microphysics, cumulus parameterization, planetary boundary layer (PBL), land surface models (LSM), and longwave and shortwave radiation (Skamarock et al. 2006).

In this work, the microphysics and convective parameterizations used were the WRF Single-Moment 5-class (WSM5) scheme (Hong et al. 2004) and the Kain–Fritsch scheme (Kain and Fritsch 1993), respectively. The WSM5 microphysics explicitly resolves water vapor, cloud water, rain, cloud ice, and snow. The Kain–Fritsch convective parameterization utilizes a simple cloud model with moist updrafts and downdrafts that includes the effects of detrainment and entrainment. The land surface model used was the Noah LSM 4-layer soil temperature and moisture model with canopy moisture and snow-cover prediction (Chen and Dudhia 2001). The LSM includes root zone, evapotranspiration, soil drainage, and runoff, taking into account vegetation categories, monthly vegetation fraction, and soil texture. The PBL parameterization used was the Yonsei University (YSU) scheme (Hong and Pan 1996). This scheme includes countergradient terms to represent heat and moisture fluxes due to both local and nonlocal gradients. Atmospheric shortwave and longwave radiations were computed by the NCAR Community Atmospheric Model (CAM) shortwave scheme and longwave scheme (Collins et al. 2004), respectively.

The design of the WRF-based mesoscale climate model over the Pacific Northwest follows that of the MM5-based mesoscale climate model described in Salathé et al. (2008). Basically, similar refinements as in the MM5 setup are made to the WRF configuration to perform long simulations and fully represent the climate system response to climate change forcing. First, nudging is applied throughout the interior of the outermost regional model domain using the forcing fields. Nudging relaxes the regional model simulations for wind, temperature, and moisture toward the driving global climate model simulations, which prevents possible drift of regional model solution from that of the driving global model over a long term. The inner nested domains are not nudged in the interior, allowing the mesoscale model to freely develop finescale features. Second, WRF is modified so that soil temperatures vary at the lower boundary (~8 m) in accordance with the evolving surface temperatures predicted by the model. The readers are referred to Salathé et al. (2008) for the rationale and procedures for updating deep soil temperature in MM5 and similarly in WRF.

b. HadRM model

HadRM (Jones et al. 2004) is the third-generation Hadley Centre regional climate model. It is a limited-area, high-resolution version of the Hadley Centre atmospheric general circulation model (HadAM3H), which is itself a high-resolution version of the atmospheric component of the atmosphere–ocean coupled general circulation model (HadCM3; Gordon et al. 2000; Johns et al. 2003). HadRM is a hydrostatic version of the fully primitive equations and includes dynamical flow, horizontal diffusion, clouds and precipitation, radiative processes, boundary layer and land surface, deep soil, and gravity wave drag.

The latitude–longitude grid is rotated in HadRM so that the equator lies inside the region of interest to obtain quasi-uniform gridbox area over that region. The available horizontal resolutions are $0.44^\circ \times 0.44^\circ$ and $0.22^\circ \times 0.22^\circ$. For the purpose of this study, we chose the higher resolution that corresponds to a minimum grid spacing of 25 km at the equator of the rotated grid.

The HadRM simulations were performed using the PRECIS package. This package also includes software to allow display and processing of the model output data (http://precis.metoffice.com). The PRECIS package is flexible, easy to use, and computationally inexpensive. It can be applied over the U.S. Pacific Northwest to provide detailed climate information for regional climate studies and climate impact assessment.

3. Experimental design

WRF was set up by using multiple nests at 108, 36, and 12 km horizontal grid spacing (Fig. 1a). The outermost WRF domain covers nearly the entire North American continent as well as much of the eastern Pacific Ocean and the western Atlantic Ocean. The use of this large domain ensures that synoptic weather systems approaching the United States are well represented by the time they reach the region. The 36-km model domain (domain 2) covers the continental United States and part of Canada and Mexico. The innermost model domain (domain 3) is centered on the Pacific Northwest and includes the states of Washington, Oregon, and Idaho (Fig. 1b). We used 31 vertical levels in the model with the highest resolution (~20–100 m) in the boundary layer. The model top was fixed at 50 mb. One-way nesting was used in this work.

The domain of HadRM (Fig. 1a) was chosen with the highest available horizontal resolution of ~25 km at the equator of the rotated grid. The HadRM model domain covers a large part of the eastern Pacific Ocean and part of Mexico and Canada to better resolve the synoptic weather systems that affect the Pacific Northwest. This model domain encompasses entirely the states of Arizona, California, Idaho, Nevada, Oregon, Utah, and Washington. There are 19 vertical hybrid levels in HadRM spanning from the surface to 0.5 hPa.

The WRF and HadRM runs were initialized at 0000 UTC 1 December 2002 and ended at 0000 UTC 1 January 2008. The first one-month simulations by WRF and

HadRM were regarded as model spinup. The initial and lateral boundary conditions for WRF and HadRM were interpolated from the National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) Atmospheric Model Intercomparison Project (AMIP)-II Reanalysis (R-2) data (Kanamitsu et al. 2002). The lateral boundary conditions were updated every six hours for both models. SST was updated every six hours in WRF using the Real-Time, Global, Sea Surface Temperature (RTG_SST) analysis (ftp://polar.ncep.noaa.gov/pub/history/sst) developed and archived at NCEP. In HadRM, SST was taken from a combination of the monthly Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST; http://badc.nerc.ac.uk/data/hadisst) and weekly NCEP observed datasets (http://www.cdc.noaa.gov/cdc/reanalysis/reanalysis.shtml). Nudging was applied to WRF domain 1 every six hours. The simulations from both WRF and HadRM models were output every hour.

4. Results

In this section, model simulations from WRF and HadRM are compared with observations at 72 HCN stations in the states of Washington, Oregon, and Idaho. We select only the stations from which 80% or more of the daily precipitation and temperature measurements are available during 2003–07 and the stations whose corresponding model grid points are land grid points in the WRF and HadRM model domains. The locations of these HCN stations are indicated in Fig. 1b.

Comparisons will be focused on daily maximum and minimum temperatures (Tmax and Tmin) and precipitation with averages over various time scales. The daily maximum and minimum temperatures are found from the hourly temperature simulations. We used the lapse rates obtained from the WRF and HadRM simulations to account for the difference in elevation of the model topography and the observing station. The lapse rates were found as the mean lapse rates at four neighboring points of the observing station and were set to be between 2°C km^{-1} and 7°C km^{-1}. P. W. Mote et al. (2009, unpublished manuscript) examined the surface and free-air lapse rates in the Cascade Mountains of Washington using data from many sources [National Weather Service (NWS) Cooperative Station Network, balloon soundings, i-Button senses, and MM5 simulations] over a course of one year. They obtained a mean lapse rate of 2.6°C km^{-1} for January Tmin and a mean lapse rate of 6.2°C km^{-1} for July Tmax. They noted that the observed lapse rates generally range between 2°C km^{-1} and 7°C km^{-1} and that the lapse rates in MM5 simulations match the observations rather well. Our examinations show that the differences between using the model generated lapse rates and the standard lapse rate of 6.5°C km^{-1} are rather small. For R-2, the standard lapse rate of 6.5°C km^{-1} was used.

No lapse rate was applied to precipitation, as a lapse rate over complex terrain would depend on several factors such as mountain width, buoyancy, and moisture fields (Smith and Barstad 2004) as well as winds (Esteban and Chen 2008).

a. Maximum temperature

Figure 2 shows scatterplots of the observed and simulated Tmax for 5-day and monthly averages at all HCN stations combined during the 5-yr period. Correlation coefficients between the observations and simulations are listed in Table 1. Notice in Table 1 that all values are significant at significance level 0.05 based on Student's *t* test. The R-2 reanalysis data, with the lapse-rate correction applied, resolve the observed spatial pattern of Tmax well (correlation coefficients are 0.92, 0.96, and 0.98 for daily values and 5-day and monthly averages, respectively). As a consequence, WRF and HadRM also simulate the observed Tmax well, following the skill in the forcing fields. High correlation coefficients (>0.90) between the observations and simulations are noted. There is slight indication of parabolic tendency in the scatterplots for the HadRM-simulated Tmax especially for Tmax ranging between 5° and 20°C (Figs. 2d,h). As will be shown later, this appears to be partly related to the driving data.

The normalized probability density functions (PDFs) of daily model biases of Tmax is presented in Fig. 2i. The PDF is a function that represents a probability distribution in terms of integrals. PDFs are characterized by their mean and standard deviation. The PDFs of daily model biases of Tmax (Fig. 2i) have a mean of −3.40°C for R-2, −0.74°C for WRF domain 2, −0.95°C for domain 3, and 0.24°C for HadRM; a standard deviation of 4.16°C for R-2, 4.09°C for WRF domain 2, 4.00°C for WRF domain 3, and 3.83°C for HadRM. The PDFs for the two nested WRF domains are virtually identical. Cold biases dominate in the R-2 reanalysis data. WRF domain 2 and 3 show mainly cold biases while HadRM exhibits nearly equal number of cold and warm biases.

Figure 3 shows the seasonal and annual mean model biases of Tmax at HCN stations averaged over the 5-yr period. Biases from both models are generally smaller than 5°C in magnitude for all seasons. During winter and fall, the nested WRF domains exhibit predominantly warm biases along the coast of Oregon and Washington and cold biases in the interior. During spring, virtually all stations show cold biases in the nested WRF domains. Summer is characterized by mainly cold biases in the

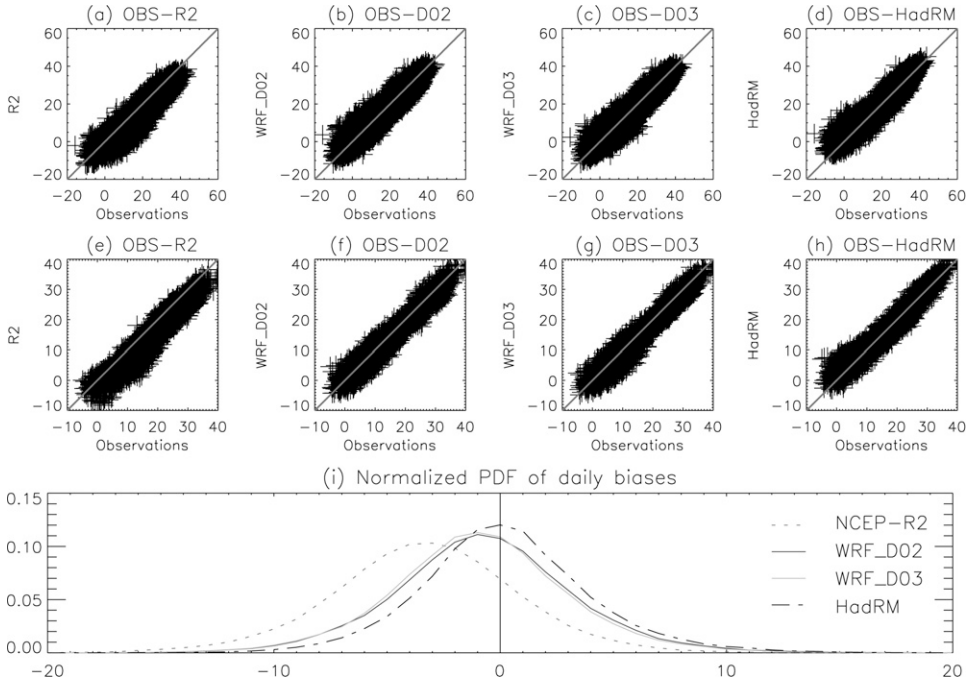

FIG. 2. Scatterplots of 5-day averaged Tmax (°C) at HCN stations between observations and (a) the R-2 reanalysis data, (b) WRF domain 2 simulations, (c) WRF domain 3 simulations, and (d) HadRM simulations. (e)–(h) The same as (a)–(d), respectively, but for monthly averaged Tmax. (i) The normalized PDF of daily model biases of Tmax.

nested WRF simulations with scattered warm biases noted along the coast of Oregon and Washington and the southern part of Idaho. For HadRM, the distribution of model biases appears similar to those of WRF except in the interior, where mainly warm biases are noted in winter and fall.

Annual mean model biases of Tmax at HCN stations indicate large cold biases on the order of 3°–6°C in the R-2 reanalysis data over the entire model domain. For both WRF domain 2 and domain 3, similar but small biases on the order of −1°–1°C are identified along the coast of the Pacific Northwest while cold biases on the order of 2°–3°C are noted in the interior. Cold biases scatter around the mean of −1.9°C are also noted in high-resolution MM5 simulations over the Pacific Northwest driven by the NCEP–NCAR reanalysis project (NNRP) data (Salathé et al. 2008). The annual mean biases of Tmax in the HadRM simulations range predominantly between −2° and +2°C at HCN stations and are generally smaller in magnitude when compared to the WRF simulations.

Similar seasonal and annual distributions in the model biases of Tmax between WRF and HadRM seem to suggest that the large-scale driving data play a role in bringing the biases. This is supported by examining the model biases of Tmax in WRF domain 1 at 108-km resolution that exhibit distributions sitting between the driving data and the mesoscale domains (Fig. 3). This may suggest that the cold biases in the R-2 reanalysis data are passed onto the high-resolution WRF and HadRM domains, thereby partially explaining the cold biases in those simulations. Meanwhile, high-resolution simulations also appear to partially offset the cold biases

TABLE 1. Correlation coefficients of daily, 5-day mean, and monthly mean Tmax, Tmin, and precipitation between observations and model simulations from R-2 reanalysis data, WRF domain 2, WRF domain 3 and HadRM. Note: All values are significant at significance level 0.05 based on Student's *t* test.

		Tmax	Tmin	Precipitation
Obs-R-2	Daily	0.92	0.83	0.46
	5-day	0.96	0.92	0.65
	Monthly	0.98	0.96	0.74
Obs-D02	Daily	0.91	0.83	0.43
	5-day	0.96	0.91	0.64
	Monthly	0.99	0.97	0.74
Obs-D03	Daily	0.92	0.84	0.43
	5-day	0.97	0.91	0.65
	Monthly	0.98	0.97	0.76
Obs-HadRM	Daily	0.93	0.87	0.42
	5-day	0.96	0.94	0.64
	Monthly	0.98	0.97	0.76

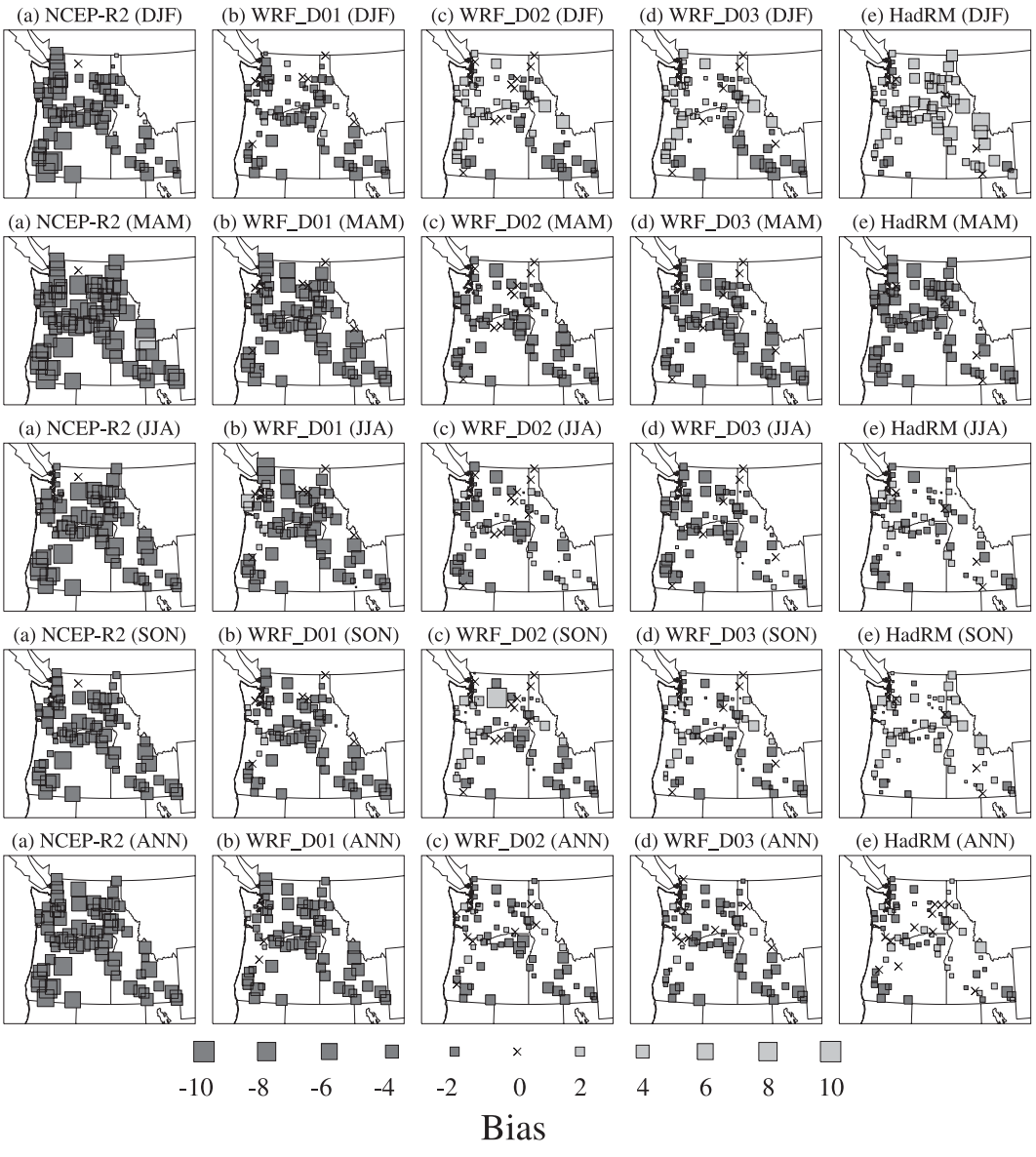

FIG. 3. Geographical distributions of seasonal (DJF, MAM, JJA, and SON) and annual (ANN) mean model biases of Tmax (°C) for (a) R-2 reanalysis data, (b) WRF domain 1, (c) WRF domain 2, (d) WRF domain 3, and (e) HadRM averaged over the 5-yr period. Each horizontal panel corresponds to model biases for one season and the annual mean and is arranged in the order for DJF, MAM, JJA, SON, and ANN from top to bottom. Size of the filled squares corresponds to the magnitude of the represented model biases. Cross represents a magnitude range of −0.2°–0.2°C.

inherited from the R-2 reanalysis data, especially for the winter and autumn periods.

We further examined the surface radiation budget at the time of the maximum temperature from WRF domains 1, 2, and 3 and HadRM. Since the model biases of Tmax for WRF domain 1 lie between the biases for the driving data and the mesoscale domains, it is worthwhile to examine the surface radiation budget among the different model domains to see where the model biases might come from. We noted similar radiation budget in all model domains without significant differences that could explain the biases. This would imply that the model biases of Tmax are related more to the driving data than to the regional models.

It is intriguing to notice that the large cold biases in the R-2 reanalysis data are partially ''corrected'' in the high-resolution simulations (Fig. 3). This may indicate that finescale processes resolved in regional climate models are important in reducing large biases in the driving data; however, similarities between WRF

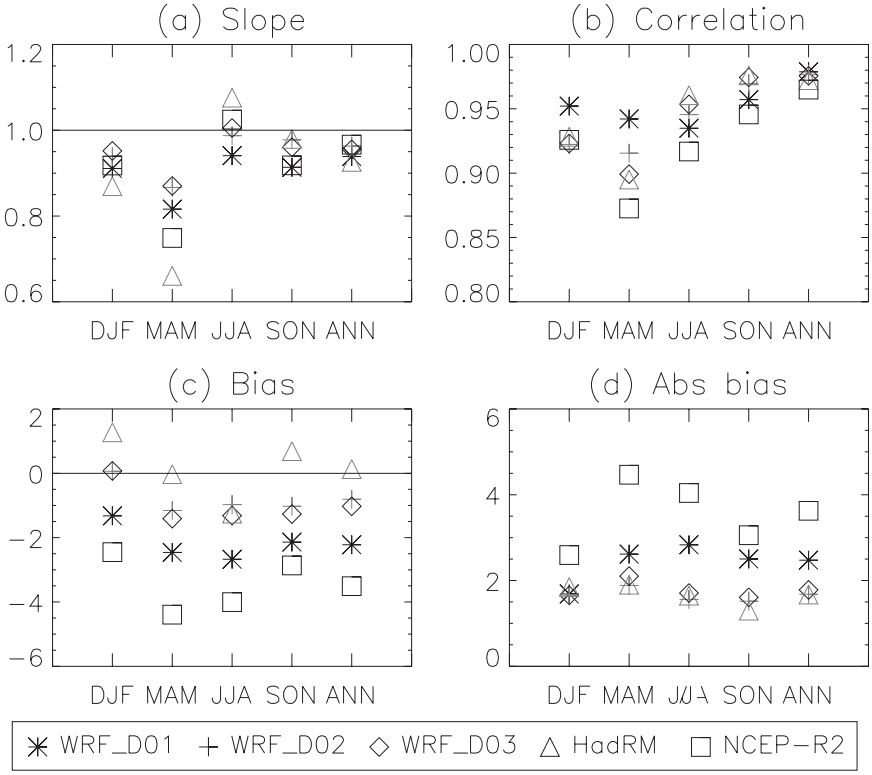

FIG. 4. Statistical performance of Tmax in the R-2 reanalysis data and WRF and HadRM simulations at seasonal (DJF, MAM, JJA, SON) and annual (ANN) time scales averaged over the HCN stations and over the 5-yr period: (a) linear regression slope, (b) correlation coefficient, (c) mean biases, and (d) absolute mean biases.

domain 2 and domain 3 also indicate that this resolution-correction has no effect beyond a certain horizontal resolution.

To further evaluate the performance of WRF and HadRM in representing the observed Tmax, we computed the statistical variables (linear regression slope, correlation coefficients, mean biases, and absolute mean biases) at seasonal [December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON)] and annual time scales averaged over the HCN stations and over the 5-yr period. Absolute mean biases are used instead of the root-mean-square errors (RMSE) since RMSE tends to overweight large biases. The results are presented in Fig. 4. Except for MAM, the slopes range around 1.0 and do not differ appreciably from each other among the R-2 reanalysis data and the WRF and HadRM models (Fig. 4a). This is expected since variations in temperature are largely dictated by large-scale systems. For MAM, all the slopes are smaller than 0.9 and the slope for HadRM is the smallest (0.65).

Correlation coefficients for the WRF domain 1 are the highest during DJF and MAM and, for the regional simulations, are the highest in HadRM during JJA and SON (Fig. 4b). WRF domain 2 and 3 show identical correlation coefficients at all times. Correlation coefficients corresponding to the R-2 reanalysis data are always lower than the WRF and HadRM simulations for MAM, JJA, SON, and annual means. It is evident that higher horizontal resolutions do not necessarily correspond to higher correlation coefficients, and hence it does not mean that Tmax is better simulated by increasing the horizontal resolution.

The R-2 reanalysis data always exhibit the largest model biases and absolute model biases at all time scales followed by the WRF domain 1 (Figs. 4c,d). WRF domains 2 and 3 show similar biases to those for the correlation coefficients. HadRM appears to show the smallest model biases among the regional simulations. In all cases, the biases in the regional models are reduced substantially compared to those in the R-2 forcing data.

MAM is the season when the regional models generally show smaller slopes, lower correlation coefficients, and higher absolute model biases when compared to the other seasons (Figs. 4a,b,d). This deficiency can be traced to the driving data since the R-2 reanalysis data

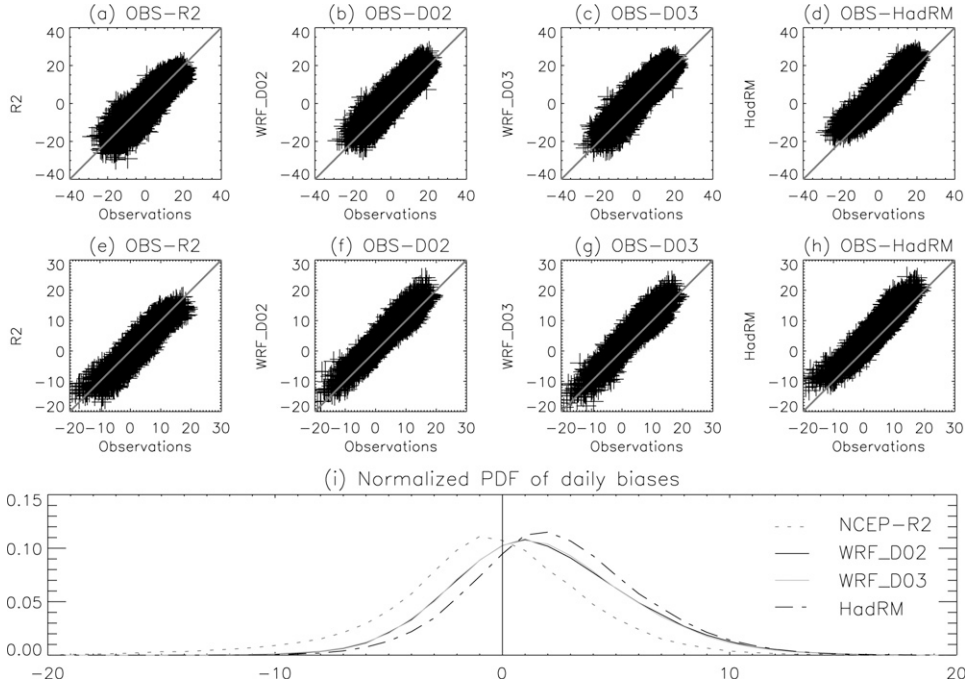

FIG. 5. As in Fig. 2, but for Tmin (°C).

also exhibit the smallest slope, the lowest correlation coefficient, and the highest mean biases and absolution mean biases in MAM when compared to the other seasons.

b. Minimum temperature

Figure 5 shows scatterplots of the observed and simulated Tmin for 5-day and monthly averages at all HCN stations combined during the 5-yr period. Correlation coefficients between the observations and simulations are presented in Table 1. A good correspondence is evident between the observations and model simulations with the correlation coefficients for daily values being 0.83, 0.84, and 0.87 for WRF domain 2, WRF domain 3, and the HadRM domain, respectively. It is noted that the correlation coefficients of Tmin between the observations and simulations are always smaller than those of Tmax (see Table 1). There are possibly two causes that could account for these differences. First, the R-2 reanalysis data show a correlation coefficient of 0.83 for daily minimum temperature and 0.92 for daily maximum temperature, indicating that the large-scale driving data at a 6-hourly temporal resolution do not represent the observed daily minimum temperatures as well as the maximum temperature. Second, WRF and HadRM may contain model deficiencies in the parameterization of nocturnal boundary layer physics and dynamics that have a large impact on nighttime (Tmin) temperatures. Warm biases are suggested in the scatterplots especially for the HadRM simulations in that the majority of points are located above the 1:1 line (Fig. 5).

The normalized PDFs of daily model biases of Tmin have a mean of −0.76°C for R-2, 1.69°C for WRF domain 2, 1.62°C for WRF domain 3, and 2.32°C for HadRM; a standard deviation of 4.50°C for R-2, 4.01°C for WRF domain 2, 3.98°C for WRF domain 3, and 3.75°C for HadRM. The PDFs for the two nested WRF domains are identical as in the case for Tmax. Mainly cold biases are indicated in the R-2 reanalysis data while the nested WRF and HadRM domains contain mainly warm biases.

Figure 6 shows the seasonal and annual mean model biases of Tmin at HCN stations. The R-2 reanalysis data exhibit predominantly cold biases at all times (Fig. 6). In contrast, the nested WRF and HadRM simulations show mainly warm biases on the order of 0°–5°C during each season, with a few cold biases along the coast of the Pacific Northwest. WRF domain 1 also shows predominantly warm biases even though nudging was applied to WRF domain 1 every 6 hours. This discrepancy implies that the model biases of Tmin in the nested WRF and HadRM simulations cannot be attributable to the driving data, which suggests that the warm biases are most likely associated with the regional models.

We examined the surface radiation budget at the time of the minimum temperature from WRF domains 1, 2,

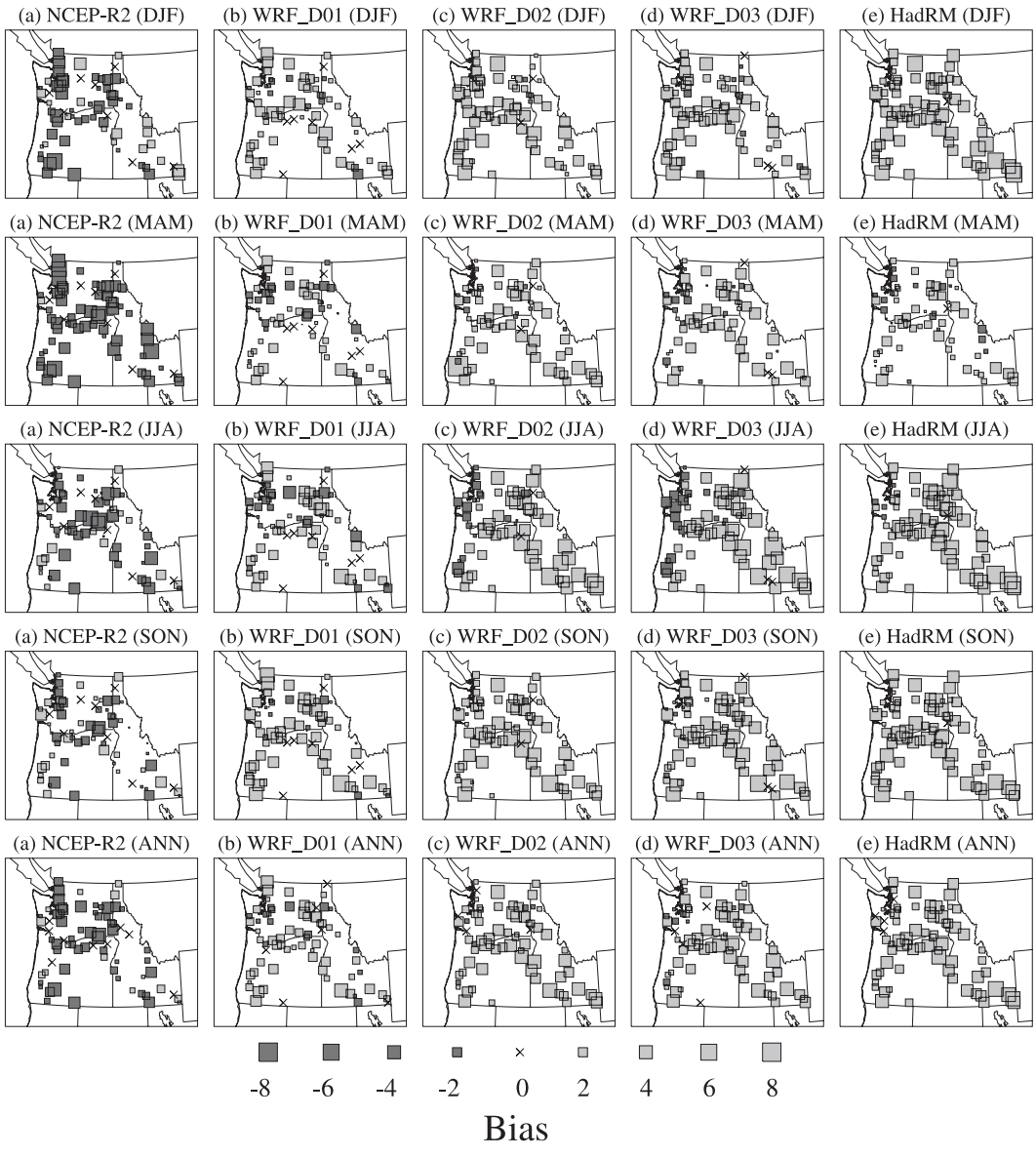

FIG. 6. As in Fig. 3, but for Tmin (°C).

and 3 and HadRM. We noted similar radiation budget in both models except for the downward longwave flux at surface, which is larger in HadRM than in WRF. This may partially explain the relatively larger warm biases of Tmin in HadRM than in WRF. We also examined the PBL height and outgoing longwave radiation at the top of the atmosphere in all the model domains and we did not notice significant differences between WRF and HadRM.

The statistical variables (linear regression slope, correlation coefficients, mean biases, and absolute mean biases) at seasonal (DJF, MAM, JJA, and SON) and annual time scales averaged over the HCN stations and over the 5-yr period are presented in Fig. 7. A relatively large spread is noted in the slopes with that for the R-2 reanalysis data situated in middle (Fig. 7a). The slopes for the WRF domains are nearly identical to each other most of the time, indicating that horizontal resolutions do not play a large role in bringing about the spread of the simulated Tmin. Notice that the slopes generally vary around 1.0 except for MAM when the slopes for R-2 and WRF domains 1 and 2 are about 0.85 and the slope for HadRM is 0.75.

WRF domain 1 shows the highest correlation coefficients at all times and is followed by HadRM (Fig. 7b). Both WRF domains 2 and 3 exhibit the lowest correlation

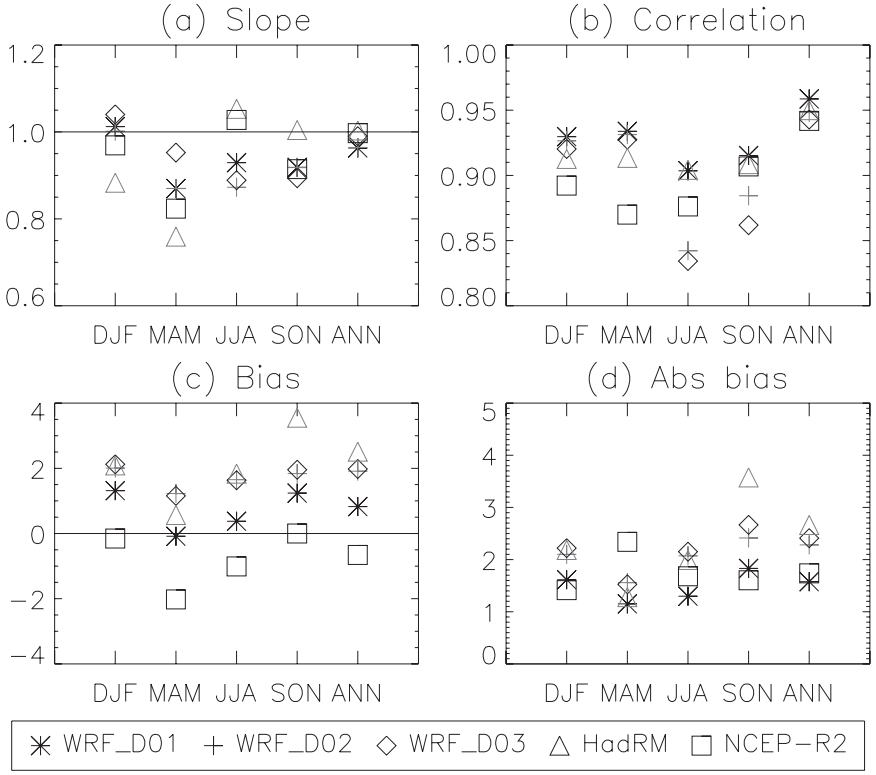

FIG. 7. As in Fig. 4, but for Tmin.

coefficients for JJA and SON. The correlation coefficients corresponding to the R-2 reanalysis data are the lowest for DJF and MAM. It is evident that higher resolutions do not necessarily correspond to higher correlation coefficients as in the case for Tmax.

Small biases of Tmin are identified in the R-2 reanalysis data for DJF and SON while cold biases are noted for MAM and JJA (Fig. 7c). The regional models all show warm biases in consistence with the spatial distributions of the model biases (see Fig. 6). Among them, WRF domain 1 has the smallest biases while the biases in the other domains are close to each other, except for SON where the bias in HadRM is larger than in the other domains (~3.5°C). In terms of the mean absolute biases (Fig. 7d), the R-2 reanalysis data and WRF domain 1 show the smallest biases at all times, except for MAM when the R-2 reanalysis data display the largest absolute biases because of large cold biases (Fig. 7c). HadRM shows the largest absolute biases for SON that are related to the large warm biases during the same time period.

We can draw several conclusions from the above discussions. First, the regional models do not necessarily perform better in representing the observed Tmin than the R-2 reanalysis data. Second, warm biases of Tmin find their origin in the regional models. This is in contrast to Tmax where the model biases are related to the R-2 reanalysis data to some extent. Finally, higher resolution does not necessarily improve model performance in resolving the observed Tmin.

c. *Precipitation*

At daily time scale, the scatterplot shows considerable spread between the observed and simulated precipitation (not shown). The correlations between the observed and simulated precipitation are 0.43 for WRF domains 2 and 3 and 0.42 for the HadRM domain (Table 1). The correlation coefficients steadily increase when averaging over increasing number of days (Fig. 8). For each time scale, the corresponding running mean of the observed and simulated precipitation time series at each station was formed and the computed correlation was averaged over all HCN stations. For example, the correlation coefficients of 5-day mean precipitation are 0.64 for WRF domain 2, 0.65 for WRF domain 3, and 0.64 for the HadRM domain and they become 0.74, 0.76, and 0.76 for monthly mean precipitation. The improvement in correlation over the range of the averaging days is the largest between 1- and 20-day means but saturates substantially after the 30-day mean (Fig. 8). Averaging was not performed beyond 90 days since the correlation coefficients may become misleading for averaging

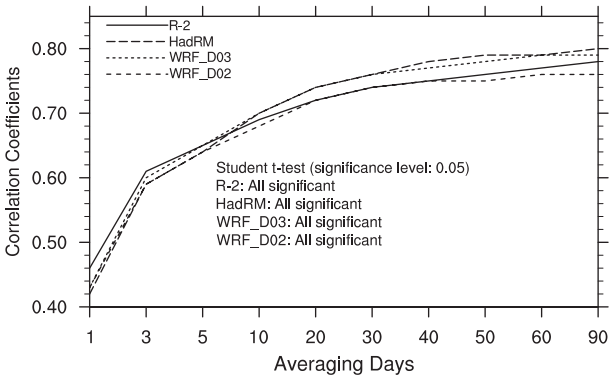

FIG. 8. Correlation coefficients of precipitation between observations and model simulations at various time scales averaged over the HCN stations. The result of Student's *t* test is included in the plot.

periods approaching the 5-yr simulated time period. Also notice in Fig. 8 that all values are significant at significance level 0.05 based on Student's *t* test.

It is noted in Fig. 8 that the R-2 reanalysis data exhibit slightly higher correlation coefficients than the regional models for 7 averaging days and shorter time periods. This suggests that the reanalysis data represent the rain-bearing storms reasonably well at synoptic time scales as has also been noted in Widmann et al. (2003). This is important since the assumption in dynamic downscaling at regional and local scales is that the driving data are able to resolve the large-scale weather systems. The timing of rain-bearing storms is determined by the reanalysis, and the regional models cannot alter this timing. Note that WRF domain 3 and HadRM show slightly higher correlation coefficients for 10 averaging days and beyond when compared to the R-2 reanalysis data and the other domains. The improvement in capturing precipitation in the high-resolution models is presumably due to the improved representation of the local terrain. Leung et al. (2003b) found that it is important to capture the regional orography correctly when determining variability at longer time scales.

To establish how well the models capture the observed daily weather sequences at different locations across the region, temporal correlation coefficients of daily precipitation at all HCN stations are illustrated in Fig. 9. Relatively high correlation coefficients ($\geq$0.6) are noted along the coast of Oregon and Washington and the windward sides of the mountains in the simulations of both models. Relatively low correlation coefficients (0.2 $\sim$ 0.4) are identified on the lee sides of the mountains and the Snake River plain. These regional differences in model performance may be related to the modeled terrain effects as air masses move across the Cascade and Rocky Mountains and are subject to larger modifications in the interior domain and lee sides than the coastal areas and windward sides. While the timing of large-scale systems may be well represented, the modulation of precipitation by orography, depending on wind direction, is not well represented and reduces correlations largely in the lee of the Cascade Range. Notice that the R-2 reanalysis data show rather similar distributions to those of WRF and HadRM, so higher spatial resolution does not automatically increase the correlations.

Figure 10 shows scatterplots of the observed and simulated precipitation for 5-day and monthly averages at all HCN stations combined during the 5-yr period. Correlation coefficients between the observations and simulations are listed in Table 1. A relatively large spread of points is identified in the 5-day and monthly mean precipitation when compared to those of Tmax

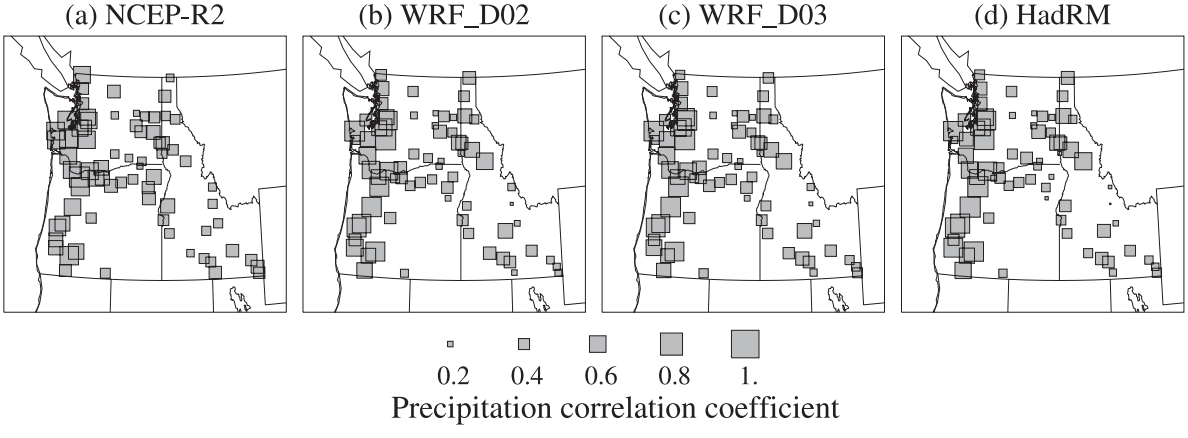

FIG. 9. Correlation coefficients of daily precipitation between observations and model simulations at HCN stations for (a) R-2 reanalysis data, (b) WRF domain 2, (c) WRF domain 3, and (d) HadRM. Magnitudes of the correlation coefficients are represented by the size of the filled squares.

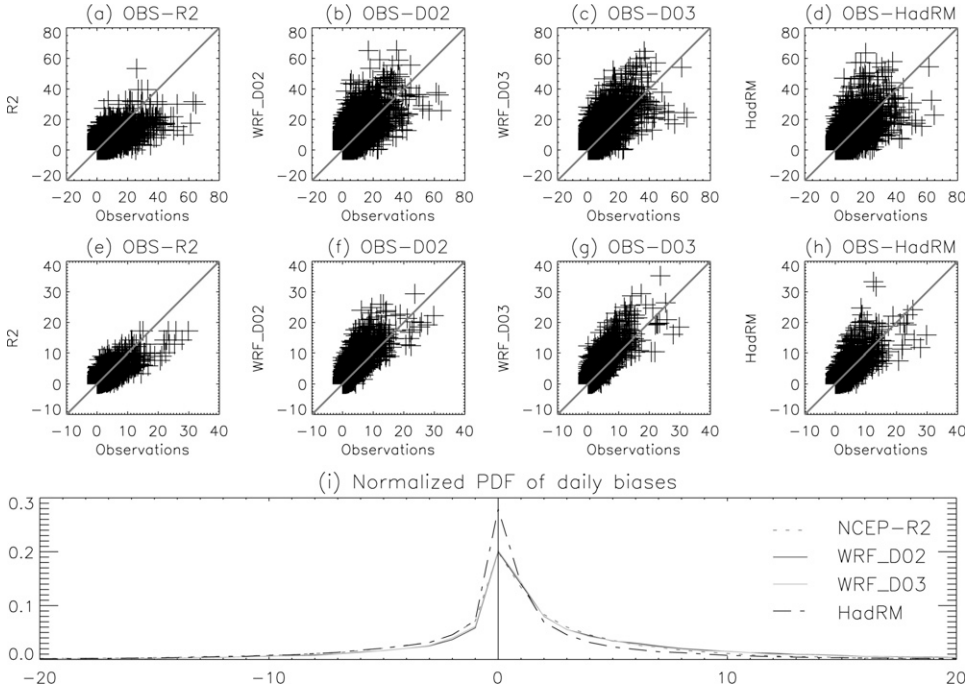

FIG. 10. As in Fig. 2, but for precipitation (mm).

and Tmin (Figs. 2 and 5), as precipitation generally occurs as a result of interaction between large-scale weather systems and local terrain and thus is more difficult to observe and to match with the observations.

The normalized PDFs of daily model biases of precipitation shows a mean of 0.28 mm for R-2, 2.32 mm for WRF domain 2, 2.09 mm for WRF domain 3, and 0.35 mm for HadRM; a standard deviation of 7.60 mm for R-2, 9.03 mm for WRF domain 2, 8.78 mm for WRF domain 3, and 8.29 mm for HadRM (Fig. 10). Wet biases are indicated in the regional models. Model biases in HadRM are rather small. Daily model biases of precipitation between −0.1 and 0.1 mm were omitted in constructing the PDFs, which would otherwise show huge peaks at the 0-mm line with similar distributions elsewhere.

Figure 11 shows the 5-yr seasonal and annual mean model biases of precipitation normalized by the observed 5-yr seasonal and annual mean precipitation at each HCN station. Normalization was employed on precipitation to make meaningful comparisons between dry and wet stations. The WRF and HadRM simulations show a large variability in the spatial distributions of the normalized biases. This reflects the strong influence of complex terrain on local precipitation. The nested WRF simulations exhibit predominantly wet biases on the order of 100%–400% in the interior at all times. The wet biases in WRF domain 3 tend to be smaller than in WRF domain 2. Dry biases are noted at a number of stations along the coast in the nested WRF simulations. Model biases in HadRM tend to be small at all times. Wet biases on the order of 100%–300% are noted in the southern parts of Oregon and Idaho in the HadRM simulations.

For precipitation, the R-2 reanalysis data show mainly dry biases on the windward side of the mountains and wet biases on the lee side (Fig. 11). The wet biases can be as high as 700% at some of the stations. This result reflects the poor simulation of the rain shadow in the coarse-grid reanalysis. The model biases in the WRF domain 1 simulations bear much resemblance to those in the R-2 reanalysis data except over the southern part of Idaho where wet biases on the order of 200%–600% are noted in WRF domain 1.

The statistical variables (linear regression slope, correlation coefficients, mean biases, and absolute mean biases) at seasonal (DJF, MAM, JJA, and SON) and annual time scales averaged over the HCN stations and over the 5-yr period are presented in Fig. 12. The slopes corresponding to the R-2 reanalysis data and WRF domain 1 range between 0.5 and 0.6 and are the furthest from the 1.0 line among all the simulations (Fig. 12a). Such a small slope in the R-2 reanalysis data is consistent with the scatterplots (Figs. 10a,e), which show appreciable deviations from the 1:1 line. The slopes corresponding to HadRM are also small for SON (0.6) and DJF (0.7). The slopes corresponding to WRF domains 2 and 3 fluctuate around 1.0, suggesting a best match to the

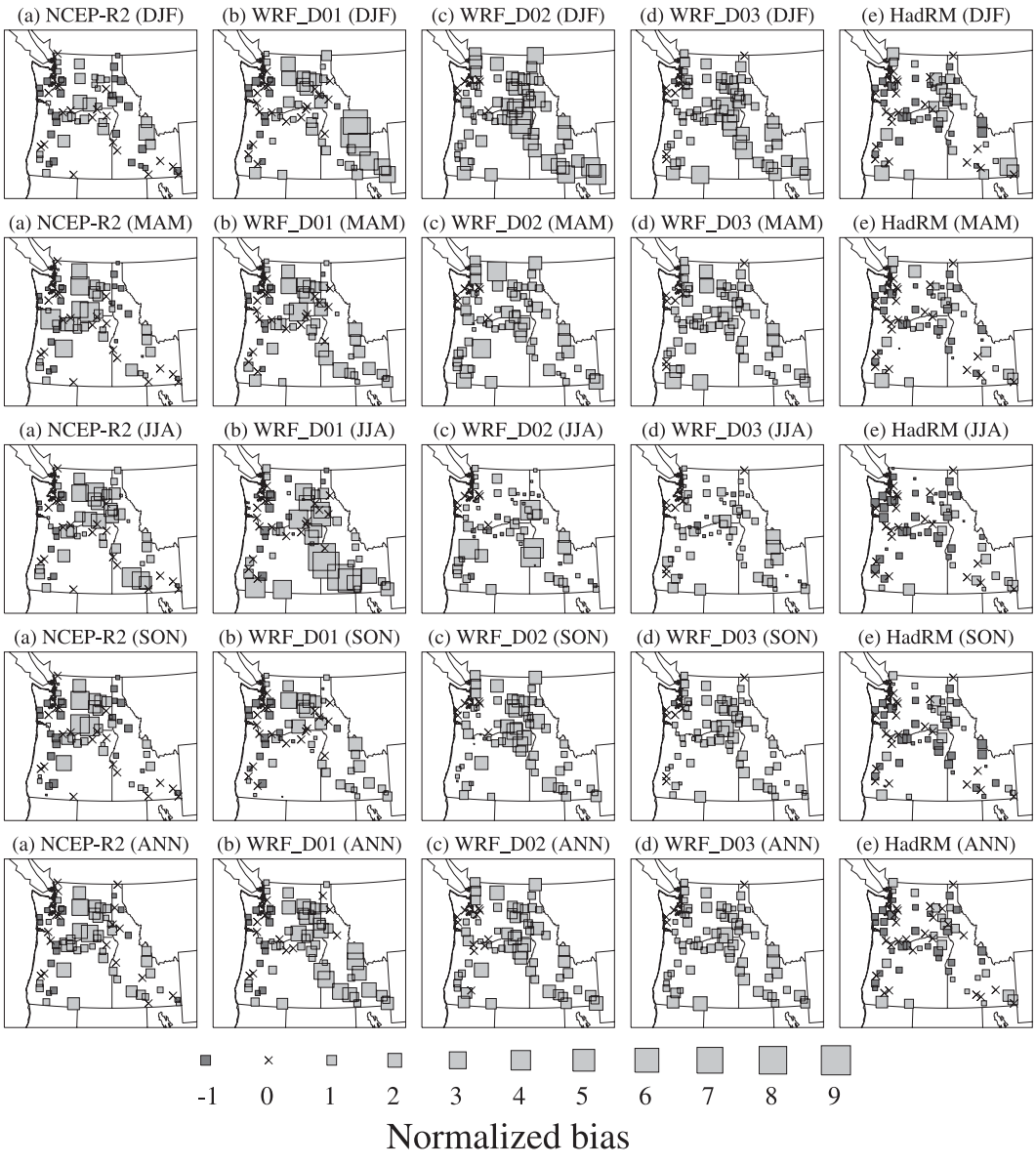

FIG. 11. As in Fig. 3, but for seasonal and annual mean model biases of precipitation normalized by seasonal and annual mean precipitation at each HCN station over the 5-yr period. Cross represents a magnitude range of −0.1%–0.1%.

observations. The slopes relate directly to the simulated spatial gradient in precipitation, which depends on model resolution. Thus, the benefits of fine-grid spacing are clear in the results for WRF domain 3 and HadRM.

The correlation coefficients for WRF domain 1 are the lowest at all times and are followed by the R-2 reanalysis data and HadRM (Fig. 12b). HadRM shows slightly higher correlation coefficients than the R-2 reanalysis data for JJA and SON. The correlation coefficients corresponding to WRF domain 3 are the highest at all times.

Mean biases in the R-2 reanalysis data, WRF domain 1, and HadRM simulations are small (~0.5–0.5 mm; Fig. 12c). Large wet biases (~2 mm) are noted in WRF domains 2 and 3 for DJF and MAM. As a result, absolute mean biases are also the largest in WRF domains 2 and 3 for DJF and MAM (Fig. 12d). JJA and SON correspond to small mean biases and absolute mean biases in the R-2 reanalysis data and all the regional domains. Notice that JJA is the dry season over the Pacific Northwest.

The above analyses suggest that WRF domains 2 and 3 simulate precipitation better than the R-2 reanalysis and WRF domain 1; however, the biases are also the largest in WRF domain 2 and 3 for DJF and MAM.

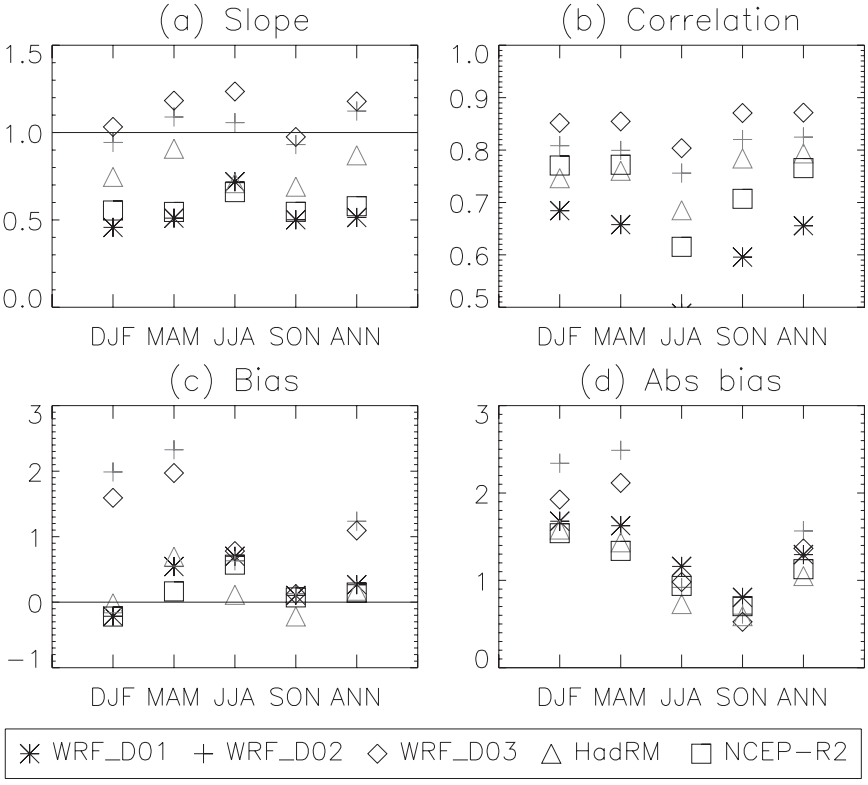

FIG. 12. As in Fig. 4, but for precipitation.

HadRM shows appreciable improvement over the R-2 reanalysis data in terms of slope, correlation coefficients, and biases. Between WRF domains 2 and 3, the correlation coefficient is slightly higher in WRF domain 3 with slightly lower biases than WRF domain 2. WRF domain 1 does not show improvement over the R-2 reanalysis data.

5. Conclusions and discussion

This work examines the performance of two regional models, WRF and HadRM, in simulating the observed Tmax, Tmin, and precipitation at HCN stations over the U.S. Pacific Northwest over a 5-yr period (2003–07). The large-scale driving data (i.e., the R-2 reanalysis data) are also verified against the station observations. The analyses are focused mainly on daily, monthly, and seasonal time scales.

The simulated Tmax from WRF and HadRM as well as in the R-2 reanalysis data compares well with the observations as reflected by high correlation coefficients at all times. Large cold biases of Tmax on the order of 4°–8°C are noted in the R-2 reanalysis data over the HCN stations when averaged both seasonally and annually. The cold biases in the R-2 reanalysis data are larger in spring and summer than in winter and fall. Predominantly cold biases of Tmax on the order of 2°–4°C are noted in the nested WRF and HadRM domains in spring and summer, while in winter and fall more stations show warm biases, especially in the HadRM simulations. There are clear indications that the large cold biases in the R-2 reanalysis data are reduced in the high-resolution regional domains. It is suggested that the cold biases in the R-2 reanalysis data are passed onto the WRF and HadRM domains, but high-resolution simulations can partially offset the cold biases.

The correlation coefficients of Tmin between the observations and model simulations are reasonably good although not as high as those of Tmax in both models. The differences in model performance for Tmin and Tmax are likely due to a combination of deficiencies in the large-scale driving data and model parameterizations of nocturnal physics and dynamics, which specifically affect Tmin. Predominantly warm biases of Tmin on the order of 2°–6°C are noted in both model simulations when averaged seasonally and annually. Model biases of Tmin tend to be small along the coast of Oregon and Washington in the WRF simulations when compared to the model interior. Cold biases of Tmin on the order of 2°–4°C are evident in the R-2 reanalysis data

at all times. It is suggested that the model biases of Tmin in WRF and HadRM are related to deficiencies in the regional models.

The correlation coefficients between the simulated and observed daily precipitation are relatively low. However, correlation steadily increases when averaging increasing number of days. This increase in the correlation coefficients levels out considerably around 30 days. The temporal correlation between the observed and simulated precipitation is higher at stations along the coast of Oregon and Washington and windward sides of the mountains than elsewhere in both models. A large variability in the spatial distributions of the normalized seasonal and annual biases of precipitation is evident in both model simulations. Model overestimation of the observed precipitation is indicated in the WRF simulations. The seasonal and annual mean biases in the HadRM simulations are rather small. For the R-2 reanalysis data, dry biases on the windward sides of the mountains and wet biases on the lee sides are noted.

Our results suggest that the WRF simulation at 12-km grid spacing does not clearly outperform 36-km simulation in terms of Tmax and Tmin. Variations in local temperatures are dictated by large-scale weather systems, with mesoscale processes associated with the underlying surface playing a secondary role; thus grid spacing is not critical to the temperature simulation even over complex terrain. The primary local effect, namely decreasing temperature with elevation, has been taken into account through the lapse-rate correction applied to model output. The value of fine-grid spacing may be more significant in simulating extreme weather events, climate variability, and climate change, which we shall address in future work. Over interannual and decadal time scales, regional alterations in snow cover, soil parameters, cloudiness, and circulation associated with interactions between the large-scale climate change and the regional topography and land–water contrasts become increasingly important (Salathé et al. 2008). In contrast to the temperature results, an appreciable improvement is found in the precipitation simulation for the 12-km WRF when compared to the 36-km WRF and 25-km HadRM simulations. Precipitation in the Pacific Northwest is generally localized over complex terrain because of the interaction of airflow with local topography. Thus, the magnitude and distribution of precipitation are closely related to the topography, and accurate simulation depends on a realistic representation of topographic effects.

Our analyses also suggest that HadRM and WRF are generally comparable in their performance in resolving the observed Tmax, Tmin, and precipitation at horizontal resolutions on the order of tens of kilometers. It is clear that both models are suitable for climate simulations. Note that HadRM is a hydrostatic model with less flexibility in horizontal resolution but is computationally efficient, while WRF is a complex modeling system with a broad spectrum of applications across scales ranging from meters to thousands of kilometers but is computationally expensive to run. Different sources of model biases as reflected in this analysis strongly suggest that the quality of the driving data is as much a limiting factor in climate simulation as the regional models.

Acknowledgments. This work is funded by an Environmental Protection Agency STAR Grant, by the National Science Foundation (ATM0709856), and by a Microsoft Corporation gift to the Climate Impacts Group. We thank the PRECIS team, especially Richard Jones, David Hein, David Hassell, and Simon Wilson for providing the PRECIS package and for their guidance in using the package. We are grateful to Prof. Clifford Mass and Richard Steed for their assistance and insights in running the WRF model. Three anonymous reviewers are acknowledged for their constructive comments and suggestions in improving the manuscript. The WRF simulations were performed at the NCAR Computational and Information System Laboratory (CISL). NCAR is sponsored by the National Science Foundation. This publication is partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA17RJ1232.

REFERENCES

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

Collins, W. D., and Coauthors, 2004: Description of the NCAR Community Atmospheric Model (CAM 3.0). NCAR Tech. Note, NCAR/TN-464+STR, 226 pp.

Esteban, M. A., and Y.-L. Chen, 2008: The impact of trade wind strength on precipitation over the windward side of the Island of Hawaii. *Mon. Wea. Rev.,* **136,** 913–928.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.,* **16,** 147–168.

Grell, G., J. Dudhia, and D. R. Stauffer, 1993: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). NCAR Tech. Note NCAR/TN-398 + IA, National Center for Atmospheric Research, 107 pp.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124,** 2322–2339.

——, J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.,* **132,** 103–120.

Johns, T. C., and Coauthors, 2003: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Climate Dyn.,* **20,** 583–612.

Jones, R. G., M. Noguer, D. C. Hassel, D. Hudson, S. S. Wilson, G. J. Jenkins, and J. F. B. Mitchell, 2004: Generating high resolution climate change scenarios using PRECIS. Met Office Hadley Centre, 40 pp.

Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 246 pp.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1631–1643.

Karl, T. R., J. C. N. Williams, F. T. Quinlan, and T. A. Boden, 1990: United States Historical Climatology Network (HCN) serial temperature and precipitation data. Carbon Dioxide Information and Analysis Center, Oak Ridge National Laboratory, Environmental Science Division Publication 3404, 389 pp.

Leung, L. R., Y. Qian, and X. Bian, 2003a: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part I: Seasonal statistics. *J. Climate,* **16,** 1892–1911.

——, ——, ——, and A. Hunt, 2003b: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part II: Mesoscale ENSO anomalies. *J. Climate,* **16,** 1912–1928.

Mass, C. F., 2008: *The Weather of the Pacific Northwest*. University of Washington Press, 336 pp.

——, D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.,* **83,** 407–430.

Salathé, E. P., R. Steed, C. F. Mass, and P. H. Zahn, 2008: A high-resolution climate model for the United States Pacific Northwest: Mesoscale feedbacks and local responses to climate change. *J. Climate,* **21,** 5708–5726.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2006: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp.

Smith, R. B., and I. Barstad, 2004: A linear theory of orographic precipitation. *J. Atmos. Sci.,* **61,** 1377–1391.

Widmann, M., C. S. Bretherton, and E. P. Salathé Jr., 2003: Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *J. Climate,* **16,** 799–816.