Article

Lessons from Inter-Comparison of Decadal Climate Simulations and Observations for the Midwest U.S. and Great Lakes Region

Ashish Sharma 1,2,* , Alan F. Hamlet 3 and Harindra J.S. Fernando 3

1 Climate and Atmospheric Science Section, Illinois State Water Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, 2204 Griffith Dr., Champaign, IL 61820, USA
2 Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, 1301 W Green St., Urbana, IL 61801, USA
3 Department of Civil & Environmental Engineering and Earth Sciences (CEEES), University of Notre Dame, 156 Fitzpatrick Hall, Notre Dame, IN 46556, USA; Alan.Hamlet.1@nd.edu (A.F.H.); Harindra.J.Fernando.10@nd.edu (H.J.S.F.)

* Correspondence: sharmaa@illinois.edu

Received: 21 April 2019; Accepted: 11 May 2019; Published: 13 May 2019

Abstract: Even with advances in climate modeling, meteorological impact assessment remains elusive, and decision-makers are forced to operate with potentially malinformed predictions. In this article, we investigate the dependence of the Weather Research and Forecasting (WRF) model simulated precipitation and temperature at 12- and 4-km horizontal resolutions and compare it with 32-km NARR data and 1/16th-degree gridded observations for the Midwest U.S. and Great Lakes region from 1991 to 2000. We used daily climatology, inter-annual variability, percentile, and dry days as metrics for inter-comparison for precipitation. We also calculated the summer and winter daily seasonal minimum, maximum, and average temperature to delineate the temperature trends. Results showed that NARR data is a useful precipitation product for mean warm season and summer climatological studies, but performs extremely poorly for winter and cold seasons for this region. WRF model simulations at 12- and 4-km horizontal resolutions were able to capture the lake-effect precipitation successfully when driven by observed lake surface temperatures. Simulations at 4-km showed negative bias in capturing precipitation without convective parameterization but captured the number of dry days and 99th percentile precipitation extremes well. Overall, our study cautions against hastily pushing for increasingly higher resolution in climate studies, and highlights the need for the careful selection of large-scale boundary forcing data.

Keywords: regional climate modeling; climatology; climate extremes; WRF model; temperature; precipitation

1. Introduction

The Midwest United States (U.S.) and the Great Lakes region are one of the world’s most productive agricultural landscapes with the world’s largest freshwater ecosystems and several large urban centers. The regional climate of the region is influenced by large-scale atmospheric circulations (e.g., moisture from the Gulf of Mexico in the south and arctic air outbreaks from the north), convective activity, and ice cover/water temperature in the Great Lakes (e.g., [1]). This region is usually fertile and has been extensively drained for agriculture. These factors contribute to large seasonal and daily variations in air temperature, including hot summers, cold winters, and occasional periods of drought and deluge. In recent decades, a statistically significant warming trend has been observed over the region [2–5]. The region exhibits warming with increasing latitudes.
The Midwest U.S. also has experienced two record-breaking floods in the last 20 years [6], strong heat waves have been more frequent [7], and extreme precipitation frequency has doubled relative to the last century [8]. Based on a special report for the Environmental Law and Policy Center and the Chicago Council on Global Affairs, there is a generally positive trend in annual precipitation for U.S. states bordering the Great Lakes for present-day (1986–2016) relative to 1901–1960, but with strong local variations in the trend across the states (http://elpc.org/glclimatechange/). There is an approximate 10% increase in annual precipitation averaged over the Great Lakes states [9]. The region has also recently witnessed unprecedented extreme changes in the timing of precipitation and runoff, with important implications for flooding, soil erosion, nutrient export, and agricultural practices [10,11].

Warm, wet winters are producing extensive early-season flooding, which threatens people and infrastructure [12]. In addition, an increase in population and associated infrastructure have made the region more vulnerable to flooding, irrespective of how humans are altering the climate itself. For example, the extreme storms experienced by the Great Lakes cities in the last 30 years or so have far exceeded the current design standards for infrastructure and have caused extensive damage. Damages in and around Calgary, Alberta exceeded $1.7B (CA $), and Toronto sustained more than $900M (CA $). In February 2018, the St. Joseph River near South Bend experienced a 2500-year flood, due to an intense atmospheric river and rain-on-snow event.

Future climate change projections for the Midwest region suggest increasing temperature and weather extremes, with regional warming up to 10 °C by 2100 [13]. Thus, the changes in temperature and precipitation, as well as future changes in the extremes of the distribution of temperature and precipitation, pose challenges for energy production, ecological conservation, human health, water resource management, and agricultural productivity of this region (IPCC [14,15]).

These adverse changes create challenges for water resource management [16,17], sustainable and resilient design of infrastructure [18], design of urban systems [19–26], ecosystem management [27–30], and regional impacts on agricultural land management and production [31]. These diverse climate change impact pathways highlight the need to evaluate the most commonly available observational and simulated datasets to determine the effectiveness of climate models in realistically capturing observed climate trends, variability, and extremes that contribute to impacts.

Generally, observational meteorological datasets are sparse and consist of near-surface observations of precipitation and temperature. The measurement of other variables, such as humidity and solar radiation, is relatively rare, and direct observations are, in isolation, often insufficient to characterize conditions with useful spatial coverage. Likewise, experimental field data (e.g., from radiosondes) are available in a few locations, but these kinds of observations are very limited. Therefore, climate models have proven to be a complementary tool to fill in the data gaps in space and time. Nonetheless, the realistic simulation of rainfall and temperature variation, especially at local scales, in climate models remains a significant challenge. For example, the current global climate models (GCMs) lack the capability of representing the multiscale cloud and precipitation [32]. RCMs have more realistic rainfall than global climate models [33–35], providing future projections with greater confidence. Thus, regional simulations constrained with large-scale forcing are crucial for determining local climate extremes, and hence they are imperative for climate change planning for impact pathways related to human and natural systems [36,37]. Dynamical and statistical downscaling methods provide finer spatial resolution information for long-and short-term climate impact assessments. However, the dynamically downscaled regional climate simulations have shown a higher skill in capturing atmospheric variability [38–41]. Furthermore, literature shows that data assimilation (nudging), especially the spectral nudging technique within the dynamical downscaling framework, adds more variability and improves the model’s skill at smaller scales [42,43].

Interestingly, limited effort has been invested in validating the input large-scale forcing datasets for regional climate models, especially in the context of extremes. This oversight may lead to the inadequate understanding of biases in model outputs. At the same time, there is a debate in the research community on the importance of increasing the spatial resolution of RCMs in simulating climate. Thus, the overall
goal of this article is to evaluate multiple RCM model outputs at different resolutions (12- and 4-km) and NARR datasets at 32-km resolution (a high-resolution NCEP model dataset with assimilated observations) with commonly available gridded station observational data. Note, henceforth ‘horizontal grid increments’ from the RCM will be called ‘resolution’ in the paper [44–47].

2. Methods

2.1. Gridded Observations

We used a gridded observational meteorological dataset at daily timescales and 1/16th-degree (approx. 5 km) resolution over the Midwest and Great Lakes region. This dataset used a hybrid product based on interpolated station records from the Daily Global Historical Climatology Network (GHCN-Daily) station records, bias adjusted station records in Canada and regridded National Center for Atmospheric Research (NCAR) Reanalysis [48] wind speed data. Gridded temperature and precipitation products are based on techniques first developed by Maurer et al. [49] and Hamlet and Lettenmaier [50]. Irregularly spaced meteorological observation station records were regridded using the SyMAP (Synteny Mapping and Analysis Program) algorithm [51], which is an inverse distance squared weighting scheme. The scheme also considered the relative position of stations that are potentially cross-correlated with each other. In addition, adjustments for precipitation gauge undercatch in the U.S. data were applied to the gridded precipitation products based on empirical estimates of the catch ratio (CR) (defined as the ratio of long-term measurements of precipitation from a standard 8-in U.S. precipitation gauge to those from an accurate reference precipitation gauge) [52]. The estimate of CR varies as a function of precipitation type (snow, mixed rain, and rain) and wind speed [53]. Having established the CR for the meteorological conditions in each day, daily precipitation values were then adjusted by multiplying by 1/CR. Precipitation undercatch in Canada was corrected by interpolating daily station records with undercatch corrections already applied in previous efforts [54]. Additional details of this gridded meteorological dataset are presented in Byun and Hamlet [13].

2.2. NARR Datasets

We used time-varying, large-scale, 3-h National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR) simulations at 32-km resolution (http://rda.ucar.edu/datasets/ds608.0/) as one of the datasets for comparison. NARR is a common dataset used by planners to make impact decisions in the United States and Canada. The research community uses NARR for creating climate and hydrology products. NARR model output is often downscaled to higher resolutions (e.g., 1–4 km) for assessing local climate vulnerabilities. The quality of NARR data has been evaluated with surface stations and sounding measurements [55]. Hence, it has been used in numerous studies for the validation of regional climate simulations in North America (e.g., [56]).

2.3. WRF Model Implementation

This study used the Weather Research and Forecasting model (WRF model, Version 3.6.1) [57]. This is a non-hydrostatic, compressible model used extensively at multiple resolutions ranging from meso-gamma (2–20 km) to local scales (~50 m) [22,24,58–63]. For the experiments in this study, the model was configured such that the outer domain covering most of North America at 12-km resolution (600 x 516 grid points) is centered at 52.24°N and 105.5°W. The inner domain at 4-km resolution (420 x 366 grid points) covers the Great Lakes megaregion. Note, the outer 12-km domain is big enough to account for all large-scale processes impacting the inner 4-km domain of our interest. Note, we settled for the innermost domain’s resolution to be 4-km and not higher than that, as a tradeoff between computational requirements for our large-scale and long-term climate simulations. The model had 50 pressure-based terrain-following vertical levels from the surface to 100 hPa. High-resolution static fields were used to the corresponding nested domains. The two-way nested model domains with terrain height are shown in Figure 1a. Figure 1b shows eight U.S. states surrounding the Great Lakes.
region (viz., Illinois, Indiana, Michigan, Minnesota, New York, Ohio, Pennsylvania and Wisconsin) along with the Canadian province Ontario in the north of the Great Lakes (Figure 1b).

Time-varying, large-scale, 6-hourly National Center for Environmental Prediction (NCEP)-U.S. Department of Energy (DOE) Atmospheric Model Intercomparison Project II reanalysis datasets (NCEP-R2) [64] at T62 Gaussian grids and 17 pressure levels were applied as lateral boundary conditions. Instantaneous model outputs at 3-hour intervals were used for analysis. Since the Great Lakes themselves have a strong influence on the regional and local climate, sea surface temperature (SST) and lake surface temperature (LST) were updated at the same interval as large-scale NCEP-R2 boundary conditions. This helped to provide a more accurate lower-boundary condition and capture the possible lake effects in the WRF model. In addition, we employed spectral nudging for horizontal winds, temperature, and geopotential height at the outer domain for the whole time-period to retain the representation of large-scale dynamics. Note, the variables in the planetary boundary layer (PBL), such as pressure and moisture, were not nudged because there is a strong coupling of the atmosphere and land surface, and we allowed the atmospheric state at lower levels to adjust freely to surface properties and forcings. For a better representation of landuse, we used 30-m 2006 National Land Cover Data set (NLCD 2006) [65] regridded for both respective domain resolutions.

WRF model has multiple parameterizations for microphysics, convection, radiation, boundary layer, and surface. Based on antecedent research over the region [20,24,43], we utilized the WRF Single-Moment six-class graupel scheme (WSM6) [66] for microphysics for its robustness in accounting for different heat and moisture tendencies in the atmosphere. We incorporated the Dudhia scheme [67] for shortwave and the Rapid Radiative Transfer Model (RRTM) for longwave radiation parameterizations [68]. The Monin–Obukhov similarity scheme based on Carlson–Boland viscous sub-layer and standard similarity functions from look-up tables was used for the surface layer, and the Yonsei University scheme [69] for PBL, and a four-layer Noah land surface model (LSM) [70] for surface physics. Subgrid-scale cumulus convective parameterization was turned on only for the outer domain (12-km) invoking the Kain–Fritsch scheme [71]. Since many cumulus parameterizations are poorly posed at 4-km grid spacing, the rationale behind the choice of the Kain–Fritsch scheme was that it activates for most of the cumulus events and yields more realistic precipitation features than fully explicit simulations for the Great Lakes region [72–74] and other regions (e.g., [59,75]).

2.4. Experimental Design

We performed WRF simulations for a period of 10 years from 1991 to 2000. The model was reinitialized for large-scale forcings (both atmospheric and surface) every year as previous studies have
shown that reinitializations for long-term climate improve not only the forcing variables (e.g., pressure, temperature, wind, and moisture) but also the model diagnostic variables (e.g., precipitation) [56,76–81]. A 24-h spin up time was used for each year. For the experiment, publicly available NARR datasets were cropped and resized to domain d02. Note, the gridded observations over inner domain cover only the neighboring Great Lakes states and not the rectangular domain d02.

In the analysis, we used precipitation and temperature variables in our assessment. For all datasets, daily mean, maximum, and minimum values were calculated. We used decadal averages of daily climatology (based on daily means) and interannual variability (based on the standard deviations of daily means), percentile and dry days as metrics for inter-comparison for precipitation. We also calculated decadal averages of summer and winter seasonal minimum, maximum, and average daily temperature to delineate the temperature trends. Note, gridded observations lack data for some of the regions in domain 2. For statistical analysis, we did not account for those regions in the simulated datasets where observations were missing. Furthermore, note that gridded observations are missing over the Great Lakes themselves, due to insufficient meteorological stations over open water to support a gridded product.

3. Results

3.1. Gridded Observations

Figure 2 shows seasonal climatology and interannual variability of observed precipitation for 1991 to 2000. For comparison, we defined the warm season from April to September and the cold season from October to March. During the warm season, the Midwest U.S. and Great Lakes region gets ample precipitation as rain mostly used for agriculture production. During the cold season, the temperature drops significantly below freezing. Warm season observations showed an average 4 mm/day precipitation (mostly in the form of rain) in the lower Great Lakes States and reduced precipitation (~3 mm/day) for upper Great lakes U.S. region and lower Canadian provinces (Figure 2a). Indiana, Illinois, Ohio, Wisconsin, and parts of Minnesota showed a relatively higher rainfall in comparison with other U.S. states. Observations showed that with higher precipitation, the interannual variability in the warm season was also higher (2–3 mm/day in wet areas vs. 1 mm/day in drier areas) (Figure 2c). In the cold season, most of the precipitation occurs in the form of snow. Observations showed that the eastern part of the U.S. and Canada had relatively large cold season precipitation in comparison to the western parts of the domain. The average amount of precipitation in the cold season was higher than warm season mostly due to the contribution of lake-effect snow in winters on the east side of the Great Lakes (Figure 2b). Overall, the interannual variability of precipitation for the cold season was lower than that for the warm season. However, the highest interannual variability was found in the same locations that are prone to lake-effect snow.
Figure 2. Seasonal climatology of observed gridded precipitation for (a) warm season from April–September and (b) cold season from October to March for a period of 10 years from the year 1991 to 2000. Similarly, seasonal interannual variability of observed precipitation for (c) warm and (d) cold season.

3.2. NARR Datasets

During the warm season, NARR datasets showed a similar climatological trend as observations over the southwest region of the domain (Figure 3a). However, NARR substantially underestimated precipitation over the Great Lakes states. During the cold season, NARR dataset showed a low precipitation bias throughout the domain (Figure 3b). The interannual variability of NARR dataset during the warm season was less than 1 mm/day for northern and around 1.5–2 mm/day in the southern region of the domain (Figure 3c). Note that due to poor resolution (32-km) of NARR data, we also did not observe fine-scale spatial patterns of precipitation. Moreover, the interannual variability of the simulations was also low in the cold season (Figure 3d). Surprisingly, NARR data did not capture lake-effect snow bands observed on the east side of the Lakes in the cold season, resulting in an underprediction of cold-season precipitation in these areas.
3.3. WRF Model Simulations

Simulations with the WRF model were performed at 12- and 4-km resolutions for the years 1991 to 2000. Climatological model outputs for the warm and cold season at both 12- and 4-km showed a similar spatial pattern (Figure 4). However, the amounts of precipitation at 4-km were less than at 12-km resolution. Recall that at 12-km resolution, we parameterize the convective precipitation using Kain–Fritsch scheme, while at 4-km we explicitly resolve the convective precipitation. Warm season climatology comparison of WRF simulations (Figure 4a,c) with observations (Figure 2b) suggested that the increase in resolution for our case diminished the warm season precipitation amounts over Wisconsin and Ohio. Both model resolutions also underestimated warm season precipitation amounts (e.g., for Illinois and Indiana). However, during the cold season, WRF showed a good match with observations (Figure 4b,d). The cold season precipitation reached a maximum of 6 mm/day. Interestingly, simulations were able to capture the higher lake-effect precipitation in the east side of the Great Lakes with fine-scale representation at 4-km in comparison to 12-km. The primary reason for the improved capture of lake-effect precipitation was the explicit inclusion of lake surface temperatures as a driver in our model dynamics. As evident, the model outperformed the relatively poor observations in the cold season, as it is difficult to capture the lake-effect processes via observation stations [15]. Both 12- and 4-km simulations captured the seasonal difference in interannual variability (Figure 5). Interannual variability was higher in the warm season (~2 mm/day) and lower in the cold season (~1 mm/day).
Figure 4. Seasonal climatology of simulated WRF model precipitation for domain at 12-km for (a) warm and (b) cold season for a period of 10 years from the year 1991 to 2000. Panels (c) and (d) are the same as (a) and (b) but for domain at 4-km.

Figure 5. Same as Figure 4 but for interannual variability of simulated WRF model precipitation.
To further explore the differences between the 12-km domain with convective parameterization and the 4-km domain that resolves convection explicitly, we investigated the climatology into monthly parts (Figure 6). In this test, we assessed the monthly climatology for April, May, and June. Both simulations at 12- and 4-km showed almost identical precipitation patterns in April. However, the differences began with large-scale convection occurring in May (Figure 6b,e) and were even more evident in June (Figure 6c,f). With a higher convective activity, 4-km simulations without explicitly resolving convection did not capture the higher precipitation amounts as evident in observations (Figure 2). This is counter-intuitive to the findings in Sharma and Huang [59] where the grid-scale (i.e., with no parameterized convection) precipitation amounts increased with an increase in resolution for complex mountainous terrain. Thus, an increase in the resolution better captures mechanical/topographic lifting processes in complex terrain but seems to be a disadvantage in capturing convective processes in the relatively flat Midwestern landscape. It is not immediately clear whether these effects are related simply to the need for additional model tuning and calibration at 4-km resolution, or whether physical processes imperfectly captured at 4-km are likely reasons for the low bias in convective precipitation.

![Figure 6](image-url)

**Figure 6.** Top panel shows the monthly climatological precipitation for WRF model domain at 12-km with explicitly resolve convective precipitation for (a) April, (b) May, and (June). Bottom panel (d-f) are same as top panel (a-c) respectively but for domain at 4-km without explicit resolved convection.

### 3.4. Inter-Comparison among all Datasets

For a detailed comparison between gridded observations, 32-km NARR, and 12- and 4-km WRF model outputs, we plotted summer (June–July–Aug) and winter (Dec–Jan–Feb) average precipitation (Figure 7). During summers, average NARR precipitation compared well with observations for the regions south of the Great Lakes. North of the Great Lakes, observations and simulations matched well for both 12- and 4-km WRF resolutions. However, 4-km as noted earlier showed relatively poor performance in comparison to observations. Interestingly during winters, NARR overall showed poor average precipitation amounts throughout the domain, and both WRF simulations showed comparable results, especially on the eastern side of the Great Lakes (Figure 7b,d,f,h). This suggests that all three models are inheriting these errors primarily from the large-scale forcing. In particular the 4-km likely “inherits” large-scale winter storms from the 12-km simulations. That the large-scale forcing should play an important role makes sense given that winter storms are often large-scale cyclonic events.
forcing should play an important role makes sense given that winter storms are often large-scale cyclonic events.

Figure 7. Left panel shows summer (June–July–Aug) average precipitation for a period of 10 years from the year 1991 to 2000 from (a) observations, (c) NARR dataset (e) WRF model output at 12-km, and (g) WRF model output at 4-km. The right panel (b,d,f,h) is the same as left, but for winter (Dec–Jan–Feb).

Analysis of 99th percentile extreme daily precipitation (PCP) showed important findings for summer and winter (Figure 8). During summers, NARR showed poor results for PCP, even though it performed well in capturing the mean statistics. This means that extreme storms with high intensity were muted in NARR. However, WRF simulations performed well in comparison with observations for PCP in summers. Interestingly, 4-km WRF simulations that did not show a good match for the mean statistics performed well for the extreme statistics for PCP as shown in Figure 8g. Likewise, WRF PCP at both resolutions also performed well in winters. However, NARR results in PCP underperformed in comparison to observations. The low bias in NARR PCP is to be expected given the scale mismatch between model and observations. That is, event statistics averaged over larger grid scales should show lower values than those observed or simulated at a higher resolution. It is also worth noting that stations often miss extreme storms in summer, so the model underprediction for NARR may be even worse than that shown here.
Similarly, we calculated the number of average dry days per year for the 1991–2000 period (Figure 9). During summers, the 4-km WRF simulations performed the best among all products. For example, observations and 4-km output showed a better comparison than other models for Illinois (~65 dry days), Ohio (~50 dry days) and Wisconsin (~40 dry days). Surprisingly, the 4-km simulations did not perform well in reproducing mean climatology statistics for these states. NARR data showed a lesser number of dry days than observations. However, NARR predicted a similar spatial pattern for dry days in the region south of the Great Lakes in winters. NARR again performed poorly for the region north of the Great Lakes showing its weakness in calculating precipitation extremes in winters. WRF models at 12- and 4-km resolutions also underperformed in calculating dry days throughout the domain in winters.

Figure 8. Same as Figure 7, but for 99th percentile extreme daily precipitation (PCP).
simulations matched observations better than precipitation in part because the temperature is a prognostic variable rather than a diagnostic variable like precipitation.

Figure 9. Same as Figure 7, but for the average number of dry days per year.

We also compared summer and winter 2-m average daily maximum, mean, and minimum temperatures (refer to Supplementary material). All datasets showed a decreasing trend in temperature with an increasing latitude, as expected. Both NARR and WRF simulations at 12- and 4-km showed a good match with observations for daily maximum and mean temperatures for summers. While WRF simulations well represented the decadal daily average minimum temperatures for summers, the NARR datasets produced 3 deg C higher temperatures for regions south of the Great Lakes (see Figure S1). The wintertime, average daily maximum, mean, and minimum temperatures compared well for all three datasets (see Figure S2). The temperature simulations matched observations better than precipitation in part because the temperature is a prognostic variable rather than a diagnostic variable like precipitation.

4. Discussion and Conclusions

This article investigated the dependence of simulated precipitation and temperature validation on the model type and resolution over a ten-year period from 1991 to 2000. Specifically, NARR data at 32-km resolution and, 12-km and 4-km WRF simulations were compared to gridded
observations for the Midwest U.S. and the Great Lakes region for both the warm and cold seasons, and summer and winter seasons. We designed this study with an overarching hypothesis that model performance would improve with an increase in resolution from NARR (32-km) to WRF at 12- and 4-km. Temperature, which is a prognostic variable, and was well captured by all three models. However, we found different strengths and weaknesses related to precipitation for different model formulations.

A detailed evaluation of NARR data for the Great Lakes region showed that it is a useful precipitation product for mean warm season and summer climatological studies, but is extremely poor for winter and cold seasons. One plausible reason for the poor performance for NARR in cold conditions is an inaccurate representation of snow processes and/or data assimilation from observed station datasets that contain an inherent cold-season low bias (snow undercatch). A similar problem with negative precipitation bias was found in observed station datasets, and thus gridded observed stations in this study were bias corrected for precipitation using surface winds (please refer Byun and Hamlet [13] for more details). NARR products, in general, also do not perform well for extreme precipitations, i.e., for both dry days and 99th percentile PCPs. For flood assessment and developing Intensity-Duration-Frequency (IDF) curves, NARR data must not be used for the Great Lakes region. However, NARR data can be used to further downscale for regional and local climate modeling for quiescent summer periods with no convective activity (e.g., a few studies that have successfully used NARR data with such constraints (e.g., [20,24])).

In winter, the NARR showed a strong underprediction of precipitation over most of the inner domain, suggesting too much simulated flow of cold, dry arctic air from the north in winter. These effects could be caused by missing Great Lakes effects in the NARR simulations [15]. Furthermore, it is likely that explicit coupling of a hydrodynamic lake model with an atmospheric model such as WRF can improve the model performance significantly in future projections [15]. The need for such an integrated land-lake-atmosphere modeling system is further exemplified in this study, especially for future climate simulations that lack the lake surface temperatures to drive the regional climate models [82].

In this study, we hypothesized that increasing the model resolution from 12- to 4-km using the WRF model would improve the ability to capture spatial patterns and the intensity of the precipitation. However, the 12-km resolution average seasonal summertime precipitation performed better than the 4-km model. Figure 7 showed that even though the 12-km model outputs drove the 4-km simulations, the 12-km results were superior in reproducing summer precipitation averages in the inner domain, especially during summers. In particular, the 4-km simulations in the southern parts of the domain in Illinois, Indiana, and Ohio showed substantially lower precipitation than observations in summer, and also a substantially larger number of dry days (Figure 9) whereas the 12-km simulations were a closer match with observations. The same underlying issues were present for the 99th percentile extremes shown in Figure 8, however, in this case, the 12-km and 4-km simulations were comparable and showed a similar low bias in the southern parts of the domain. These results demonstrate that the 12- and 4-km simulations are a part of a “grey zone” for which the convective phenomenon is only partly resolved. Both these resolutions are equally capable of capturing strong convective storms in summer that produce extreme precipitation. The results also suggest that the initiation of convective storms and precipitation recycling is weak in the 4-km simulations, resulting in the elevated number of dry days (Figure 9), and low summer precipitation bias for the 4-km simulations in the southern parts of the domain (Figure 7). In summer, however, the results are strongly related to internal dynamics associated with small-scale convective storms, precipitation recycling, interactions with the Great Lakes, etc. Thus, we hypothesize that the poor agreement with observations in summer for the 4-km model is likely due to the problems with convection and weak precipitation recycling. This also explains the inability of our high-resolution WRF model (4-km) without convective parameterization to capture summertime convective precipitation (Figure 6) in comparison with 12-km simulations invoking convective parameterization. Our high-resolution 4-km WRF model simulations captured both extremes, viz., dry days and 99th percentile PCPs. Overall, the difficulty was observed in capturing
seasonal precipitation totals and variability. The very close agreement between WRF winter simulations at 12- and 4-km shows the dominant contribution of the large scale forcing in this season.

The 12- and 4-km WRF simulations performed much better in winters, albeit with a substantial wet bias over the western part of the inner domain. In winter, although biased with respect to observations, 12-km and 4-km WRF simulations showed very close agreement for seasonal totals, extremes, and dry days. This was expected, since the storms are typically large-scale events in winter, and the large-scale forcing is in a close agreement between the 12- and 4-km and 12-km runs.

Thus, there may be no need to push for higher resolution for studying convective rainfall. This exercise implies that the parameterization schemes for precipitation are not resolution-independent; thus, a refinement of resolution is no guarantee of a better result. Both WRF model resolutions were able to capture the lake-effect precipitation successfully which was muted/absent in previous studies (e.g., [83,84]). This was due to the inclusion of SST/LST in the WRF model as more accurate lower-boundary surface conditions driving the WRF model that captured the lake effect. At the same time, there need to be considerable efforts to improve numerical model sensitivities to convection, microphysics and the water temperature in simulating winter lake-effect snow [73,85–87]. It is beyond the scope of this manuscript to study these aspects.

In conclusion, our study cautions against hastily pushing for increasingly higher resolution in climate studies especially within the “grey zone” around 4-km, and also highlights the need for the careful selection and validation of large-scale boundary forcing data. Alternatively, to avoid the “grey zone,” future studies should use either 10–12 km grid resolution with convective parameterization, or use a very high resolution, say 1-km and explicitly represent the cumulus cloud systems. As mentioned in our experimental design, the choice of the resolution is also dictated by the tradeoffs between computational requirements for large-scale and long-term climate simulations. At the same time, improper validation of input forcing data will also likely produce systematic errors and forced bias correction for one set of simulations may not work for a similar study for a different period or region.

We also recognize that the WRF simulations may improve with some form of land data assimilation or longer spin up time (e.g., [20]). We hope to test this hypothesis in future studies. Given our findings, we encourage readers to carefully validate the observed, assimilated, and simulated datasets by validating climatology and extremes before using them for further downscaling and/or impact assessment studies. We hope that lessons from this study will help direct the efforts of researchers to design better-informed climate experiments with improved impact assessments.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4433/10/5/266/s1, Figure S1: Left panel shows observed summer 2-m daily maximum average from (a) observations, (d) NARR dataset (g) WRF model output at 12-km, and (j) WRF model output at 4-km for a period of 10 years from year 1991 to 2000. Middle and right panels are same as left, but for mean and minimum temperatures respectively, Figure S2: Same as Figure S1, but for the winter season.


Funding: This research was supported by the Notre Dame Environmental Change Initiative (ND-ECI) and computing grant from the Great Lakes Consortium for Petascale Computation (GLCPC) for the Blue Waters supercomputers.

Acknowledgments: The research work was supported by the Notre Dame Environmental Change Initiative (ND-ECI). Simulations were performed with NCAR Cheyenne and the National Center for Supercomputing Applications (NCSA) Blue Waters GLCPC computing grants for supercomputing facilities.

Conflicts of Interest: The authors declare no conflict of interest.

References


32. Goswami, B.B.; Mukhopadhyay, P.; Khairoutdinov, M.; Goswami, B.N. Simulation of Indian summer monsoon intraseasonal oscillations in a superparameterized coupled climate model: need to improve the embedded cloud resolving model. *Clim. Dyn.* 2013, 41, 1497–1507. [CrossRef]


72. Xiao, C.; Lofgren, B.M.; Wang, J. Improving the lake scheme within a coupled WRF-lake model in the Laurentian Great Lakes. *JAMES* 2016, 11. [CrossRef]


75. McMillen, J.D.; Steenburgh, W.J. Capabilities and Limitations of Convection-Permitting WRF Simulations of Lake-Effect Systems over the Great Salt Lake. *Weather Forecast.* 2015, 30, 1711–1731. [CrossRef]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).