

Internal variability of a dynamically downscaled climate over North America

Jiali Wang^{1*}, Julie Bessac², Rao Kotamarthi¹, Emil Constantinescu², and Beth Drewniak¹

¹Environmental Science Division, Argonne National Laboratory, Argonne, Illinois, USA

² Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois,
USA

*Corresponding author: J. Wang, Environmental Science Division, Bldg. 240, Rm. 6A22, Argonne National Laboratory, 9700 South Cass Ave., Argonne, IL 60439.

Email address: jialiwang@anl.gov

Telephone number: 630-252-2848

Fax number: 630-252-5880

Acknowledgments

This work is supported under a military interdepartmental purchase request from the Strategic Environmental Research and Development Program, RC-2242, through U.S. Department of Energy (DOE) contract DE-AC02-06CH11357. The CCSM4 data are downloaded from <https://www.earthsystemgrid.org/home.htm>. Computational resources are provided by the DOE-supported National Energy Research Scientific Computing Center.

Abstract

This study investigates the internal variability (IV) of a regional climate model, and considers the impacts of horizontal resolution and spectral nudging on the IV. A 16-member simulation ensemble was conducted using the Weather Research Forecasting model for three model configurations. Ensemble members included simulations at spatial resolutions of 50 km and 12 km without spectral nudging and simulations at a spatial resolution of 12 km with spectral nudging. All the simulations were generated over the same domain, which covered much of North America. The degree of IV was measured as the spread between the individual members of the ensemble during the integration period. The IV of the 12 km simulation with spectral nudging was also compared with a future climate change simulation projected by the same model configuration. The variables investigated focus on precipitation and near-surface air temperature. While the IVs show a clear annual cycle with larger values in summer and smaller values in winter, the seasonal IV is smaller for a 50-km spatial resolution than for a 12-km resolution when nudging is not applied. Applying a nudging technique to the 12-km simulation reduces the IV by a factor of two, and produces smaller IV than the simulation at 50 km without nudging. Applying a nudging technique also changes the geographic distributions of IV in all examined variables. The IV is much smaller than the inter-annual variability at seasonal scales for regionally averaged temperature and precipitation. The IV is also smaller than the projected changes in air-temperature for the mid- and late 21st century. However, the IV is larger than the projected changes in precipitation for the mid- and late 21st century.

Keywords

Internal variability, regional climate model, spectral nudging, high spatial resolution, climate change

1 Introduction

Unlike weather forecasts, which are short-term (<1 week) numerical model integrations, regional climate simulations are performed over longer integration times, ranging from seasonal to decadal time scales. Regional scale climate simulations are initialized only once and then extended for multiple months or multiple years, driven at the lateral boundaries by time-dependent large-scale meteorological data, which is obtained either from reanalysis data or a global climate model (GCM). When regional climate simulations were explored in the early 1990s (Giorgi and Mearns 1991), it was thought that these kind of long-term simulations were almost totally constrained by the lateral boundary forcing and admitted only one solution. However, recent studies have shown that regional climate simulations have certain freedom and significant variabilities despite being controlled at their boundaries by large-scale atmospheric flow (e.g., Giorgi and Bi 2000; Christensen et al. 2001; Caya and Biner 2004; Alexandru et al. 2007; Lucas-Picher et al. 2008; Nikiéma and Laprise 2011; Braun et al. 2012). This suggests that the regional climate model (RCM) can generate different solutions even when the model is driven by the same lateral boundary conditions (LBCs) under slightly different initializations.

The variability of different solutions generated by RCMs is called internal variability (IV). IV comes from the nonlinear physical and dynamical processes that are described by RCM equations and develops under given large-scale conditions. Due to the restrictions on the large-scale atmospheric flow imposed by the LBCs, the level of IV generated by RCMs would be smaller than those generated by GCMs (at least at the large scale); however, it is important to evaluate the IV of an RCM because this variability may modulate, or even mask, physically forced signals in the model (Braun et al. 2012). The regions with different solutions can be identified as those where the RCM develops its own processes uninhibited by the lateral boundary processes. Giorgi and Bi

(2000) were among the first to study the IV using an RCM. They randomly perturbed the initial conditions in a set of seasonal RCM simulations, and showed that the level of IV was insensitive to either the magnitude or the source of the perturbation, but was sensitive to synoptic circulations, different seasons (i.e., IV is stronger in summer than in winter), region, and model configurations. Giorgi and Bi (2000) also noted that perturbations in initial conditions modified the day-to-day solutions, but did not significantly affect the domain-wide average 3-month climatology. Caya and Biner (2004) conducted a three-member ensemble, 1-year RCM simulation, initiated with different atmospheric and surface initial conditions. They detected a clear annual cycle in the IV, with smaller values in winter and larger values in summer. They also found that the IV was mostly caused by different time-evolutions of the atmospheric flow generated by different ensemble members.

With the growth of computational power over the years, the domain size of regional climate simulations have increased considerably. As domains expand, the LBCs' control over RCMs is reduced and RCMs have more freedom to develop their own circulations. Therefore, the IV of the RCM could intensify. Vanitsem and Chomé (2005) showed that domain size influences the RCM's sensitivity to initial conditions, with stronger sensitivity for simulations over a larger domain, and weaker sensitivity over a smaller domain. Alexandru et al. (2007) performed simulations with five different domain sizes over eastern North America (all domains started from the same point in the Southwest), and found a general increase in the IV with enlarged domain sizes. Braun et al. (2012) found the IV was neither affected by the time period (i.e., past versus future) nor by the type of driving data (i.e., reanalysis data versus GCM output). Lucas-Picher et al. (2008) found that, from a 10-member ensemble of 10-year simulations over North America, there was no long-term tendency in the IV, but there were fluctuations of the IV in time, such as in day-to-day solutions.

To prevent RCMs from departing strongly from their native model driver (e.g. imposed GCM boundary conditions), nudging techniques are applied in the interior of the RCM domains to keep the large-scale circulation of the RCM close to the driving field (von Storch et al. 2000; Wang and Kotamarthi 2013). In this case, the IV would be smaller than in the simulations that do not apply nudging techniques. For example, Braun et al. (2012) found that the IV generated by deactivating spectral nudging was almost double the IV generated with nudging, and approaches the magnitude of a GCM's IV. It is far from clear, however, how much of the IV within a model is considered as “healthy,” and how much is “to be avoided.” On the other hand, IV is also affected by spatial resolutions of the RCMs. For example, Girard and Bekcic et al. (2005) found that an RCM simulation at a spatial resolution of 50 km generates larger IV than a simulation with exactly the same model setup but at a spatial resolution of 100 km. However, there are very limited studies that compare the impacts of spectral nudging with the impacts of spatial resolution on the IV. While most of the previous studies investigate the IV at a spatial resolution of 50–100 km, this study focuses on the IV of an RCM at a spatial resolution of 12 km, and compares the IV with an RCM at a spatial resolution of 50 km. We also compare the IV between an RCM with and without spectral nudging for 12 km simulations. We answer the following questions: (1) How does IV change with various spatial resolutions? (2) How sensitive is IV to spectral nudging techniques? And (3) How important is IV compared to inter-annual variability and projected future climate changes? This is a follow-on study of Wang and Kotamarthi (2015, WK15 hereafter), which provided future climate projections at a spatial resolution of 12 km over much of North America using the same model configuration as this study. We aim to understand the uncertainty due to IV of the model simulation studied in WK15, and the relative importance of IV compared to inter-annual variability and projected climate changes.

2 Methodology

2.1 Experimental design

This study uses the Weather Research Forecasting (WRF) model version 3.3.1 to dynamically downscale the Community Climate System Model version 4 (CCSM4) to spatial resolutions of 50 km and 12 km with and without spectral nudging. The physics schemes used include the Grell-Devenyi convective parameterization (Grell and Devenyi 2002), the Yonsei University planetary boundary layer scheme (Noh et al. 2003), the Noah land surface model (Chen and Dudhia 2001), the longwave and shortwave radiative schemes of the Rapid Radiation Transfer Model for GCM applications (<http://rtweb.aer.com>) (Iacono et al. 2008), and the Morrison microphysics scheme (Morrison et al. 2009). There are 600 west-east $\times$ 515 south-north grid points for the 12 km simulations, and 150 west-east $\times$ 130 south-north grid points for the 50 km simulation over a large domain covering most of North America (Fig. 1 in Wang and Kotamarthi 2014). This domain contains high topography over the Rocky Mountains and the inflow boundary is mostly located in the Pacific Ocean. Spectral nudging is applied to one of the 12 km simulations above 850 hPa with wavelengths around 1200 km. The nudging coefficient is $3 \times 10^{-5} \text{ s}^{-1}$.

We conducted three sets of simulations. The first set of simulations was at a spatial resolution of 50 km without nudging (hereafter, 50km_no_nudg); the second set of simulations was at a spatial resolution of 12 km without nudging (hereafter, 12km_no_nudg); and the third set of simulations was at a spatial resolution of 12 km with nudging (hereafter, 12km_nudg). Each set of simulations included 16 ensemble members, which shared exactly the same model setup (e.g., nudging technique, domain coverage, and spatial resolution), physics schemes, lateral boundary conditions for the atmospheric fields, prescribed sea surface temperature (SST), and sea ice coverage for the ocean surface, with the exception of the surface and atmospheric initial conditions. The 16

members started with initial conditions 12 hr apart between 0000 UTC on 1 November 1994 and 0000 UTC on 8 November 1994; and end on 0000 UTC 1 January 1996. Therefore, the 16 members overlapped for the entire year of 1995, with a spin-up period varying from 55 to 60 days. The analyses were performed starting 1 January 1995. To investigate the robustness of the number of ensemble members in each set of simulations, we generated 200 sub-samples with 10 members taken with uniform probability and without replacement from the 16-member ensemble for each set of simulations at their original spatial resolution. In addition, to investigate the possible dependence of IV on the integration horizon (i.e., one year versus multiple years), we continued all three sets of simulations for one more year (1996) and generated a 10-member ensemble. Statistical analysis based on these sub-samples is described below. Because all the RCM simulations in this study used the same prescribed SST, the IV was caused by the atmospheric variations rather than the SST variations.

2.2 Internal variability calculations

The IV is defined by the spread between the individual members of the ensemble during the integration period. The spread is measured by the standard deviation between the 16 members in the ensemble. First, we calculated the variance of the 16 members:

$$\sigma_{en}^2(i, j, t) = \frac{1}{N} \sum_{n=1}^N [Y_n(i, j, t) - \bar{Y}(i, j, t)]^2 \quad (1)$$

Where $Y_n(i, j, t)$ refers to a variable Y on grid point (i, j) at time t for member n in the ensemble and N is the total number of ensemble members, here $N=16$. $\bar{Y}(i, j, t)$ is the ensemble mean defined as:

$$\bar{Y}(i, j, t) = \frac{1}{N} \sum_{n=1}^N Y_n(i, j, t) \quad (2)$$

A measure of the seasonal average of the IV and its geographical distribution over the model domain is calculated by the square root of the seasonal-average of $\sigma_{en}^2(i, j, t)$ in eq. (1). Domain-averaged IV during the course of the model integration is calculated by the square root of the spatial average of $\sigma_{en}^2(i, j, t)$ in eq. (1). The details of these calculations can be referred in Alexandru et al. (2007) and Lucas-Picher et al. (2008). The meteorological fields focused on were precipitation and near-surface air temperature.

2.3 Statistical analysis

To assess the robustness of the IV of each set of simulations and to investigate the statistical significance of the difference between the three sets of simulations affected by spatial resolution and spectral nudging, we perform statistical analysis on the results of the IV obtained from the 16-member ensemble. The number and the length of the simulations in the ensemble are limited by the computing resources required to run the WRF over such a large domain. Alexandru et al. (2007) suggested 10-members or more should provide a robust estimation of IV. Several earlier studies of IV were based on two to four ensemble members (Giorgi and Bi 2000; Caya and Biner 2004).

First, to investigate the robustness of the number of ensemble members in each set of simulations, we generated 200 sub-samples with 10 members, taken with uniform probability and without replacement, from the 16-member ensemble of each set of simulations at their original spatial resolution. We conducted the same analysis (such as calculating the IVs) for the 200 10-member ensembles, and then computed Δ_i , which is defined as:

$$\Delta_i = \frac{\left| IV_{16,i} - \frac{1}{200} \sum_{r=1}^{200} IV_{10,r,i} \right|}{\sigma_{IV_{10,r,i}}} \quad (3)$$

where i denotes each pixel, $IV_{16,i}$ denotes the IV calculated using the 16-member ensemble at the i^{th} pixel, and $IV_{10,r,i}$ denotes the IV calculated using the 200 10-member ensembles ($r=1, 2, 3, \dots, 200$) at the i^{th} pixel. $\sigma_{IV_{10,r,i}}$ is the standard deviation of the 200 values of IV_{10} . If $\Delta_i > 2$, then, at least approximately, the differences in IV calculated by the 10-member ensemble and by the 16-member ensemble are statistically significant at the 5% level. Otherwise, the differences are not statistically significant. In other words, using Eq. 3 we compared the IV_{16} with the confidence interval: $\left[\mu_{IV_{10,r,i}} \pm 2\sigma_{IV_{10,r,i}} \right]$, where $\mu_{IV_{10,r,i}} = \frac{1}{200} \sum_{r=1}^{200} IV_{10,r,i}$. If IV_{16} falls into the confidence interval, then it indicates the differences in IV calculated by the 10-member ensemble and by the 16-member ensemble are not statistically significant.

Second, to investigate the statistical significance of the differences in IV between the three sets of simulations affected by spatial resolution and spectral nudging, we first aggregated the 12 km model output onto the 50 km grid. Then we generated 200 sub-samples with 10 members taken with uniform probability and without replacement from the 16-member ensemble for each set of simulations at the 50 km spatial resolution. IVs are calculated for the 200 10-member ensembles and for the 16-member ensembles for each set of simulations. Inspired by the previous study Wang et al. (2015), using one set of simulations as a reference data, we calculated θ_i , which is defined as:

$$\theta_i = \frac{\frac{1}{200} \sum_{r=1}^{200} |IV_{10,r,12N,i} - IV_{16,12NN,i}|}{\sigma_{|IV_{10,r,12N,i} - IV_{16,12NN,i}|}} \quad (4)$$

We considered 12km_no_nudg (indicated by $12NN$ in eq. 3) as a reference, and calculated the statistical significance of the difference between 12km_nudg (indicated by $12N$ in eq. 3) and 12km_no_nudg. Similarly, for example, we could consider 12km_nudg or 50km_no_nudg as a reference and calculate the significance of the difference between 12km_no_nudg and the reference. In these cases, $IV_{16,12NN,i}$ would be replaced by $IV_{16,12N,i}$ or $IV_{16,50NN,i}$, and $IV_{10,r,12N,i}$ would be replaced by $IV_{10,r,12NN,i}$. Note that both 12km_no_nudg and 12km_nudg were regridded onto the 50km grid. If $\theta_i > 2$, then, at least approximately, the differences in IV calculated by any two different sets of simulations with various spatial resolution or spectral nudging are statistically significant at the 5% level. Otherwise, the differences between any two sets of the simulations are not statistically significant. In other words, using eq. 4 we compare the average of the distance between two sets of simulations ($\mu_{d_i^r}$) with the standard deviation of the distance ($2\sigma_{d_i^r}$), where $d_i^r = |IV_{10,r,12N,i} - IV_{16,12NN,i}|$. If $|\mu_{d_i^r}| < 2\sigma_{d_i^r}$, then the difference in IV calculated by any two different sets of simulations with various spatial resolution or spectral nudging are not statistically significant.

2.4 Coefficient of variations

To determine the relative importance of IV versus interannual variability, we employed the coefficient of variation (CV). The CV allows for the intercomparison of variables with different means so that results from different variables can be compared (Done et al. 2014). In this study,

we compared the 16-member ensemble for 1995 with a 10-year single member simulation for the period from 1995 to 2004. The CV of the multi-member ensemble is defined as:

$$CV_{16m} = \frac{\sigma_{\bar{X}_n^{i,t}}}{\frac{1}{16} \sum_{n=1}^{16} \bar{X}_n^{i,t}} \quad (5)$$

where X denotes the seasonal and regional average of the 16-member ensemble of a simulated field, such as temperature or precipitation, n is the index of the ensemble member; here $n=1, 2, \dots, 16$ for 1995. Therefore, the CV_{16m} is calculated by dividing the ensemble mean seasonal and regional average by the ensemble standard deviation, indicating the relative importance of IVs to the multi-member ensemble means. Similarly, eq. 5 can be applied to the 200 sub-sample of the 10-member ensemble and for the 10-member ensemble for year 1996, such as:

$$CV_{10m} = \frac{\sigma_{\bar{X}_n^{i,t}}}{\frac{1}{10} \sum_{n=1}^{10} \bar{X}_n^{i,t}} \quad (6)$$

The CV of the 10-year single member simulation is defined by:

$$CV_{10y} = \frac{\sigma_{\bar{X}_l^{i,t}}}{\frac{1}{10} \sum_{l=1}^{10} \bar{X}_l^{i,t}} \quad (7),$$

where l denotes the number of years, here $l=1, 2, \dots, 10$. Therefore, the CV_{10yr} is given by the 10-year standard deviation divided by the 10-year mean seasonal and regional average, indicating the relative importance of interannual variability to multi-year means. A $CV_{16m} < CV_{10yr}$ suggests that the IV is less important than the interannual variability and therefore, IV does not have a significant impact on the annual variation of the meteorological field. We also computed the mean CVs (

$\mu_{CV_{10m}}$) from the 200 10-member ensembles, and the standard deviations of the distance ($\sigma_{d_{10vs16}}$) between the mean CV_{10m} and CV_{16m} , where $d_{10vs16} = |\mu_{CV_{10m}} - CV_{16m}|$. Then we compared the CV_{16m} with the confidence interval: $[\mu_{CV_{10m}} \pm 2\sigma_{d_{10vs16}}]$. If the CV_{16m} at any pixel falls into the confidence interval, then the difference in CV calculated from the 10-member ensemble and the 16-member ensemble are not statistically significant.

3. Results

3.1 A general view

Fig. 1 shows the 500-hPa geopotential height as one example of a randomly chosen five-member (out of 16) ensemble with different initial conditions after six simulated months (0000 UTC 30 June 1995) and 12 simulated months (0000 UTC 30 December 1995). Comparing the three sets of simulations (50km_no_nudg, 12km_no_nudg, and 12km_nudg), in general, indicates the five simulations with different initial conditions and without nudging produced different solutions even with the same set of boundary conditions. In particular, 12km_no_nudg shows the most divergent solutions in the five ensemble members, especially toward the northeastern side of the domain by the sixth month. The 50km_no_nudg simulation also produces slightly different solutions for geopotential height, but the variability is smaller than that produced by 12km_no_nudg. By the 12th month, the geopotential heights are mostly converged for both 12km_no_nudg and 50km_no_nudg, even over the northeastern side of the domain. The five ensemble members of 12km_nudg generate very similar solutions for the geopotential height even by the sixth month of the integration, indicating that nudging reduces the IV of RCMs. While other randomly chosen five-member ensembles do not show exactly the same geospatial pattern as Fig. 1, they agree with the conclusion drawn above that 12km_nudg shows the smallest IV in comparison with 12km_no_nudg and 50km_no_nudg.

3.2 Geographic distribution of IV of precipitation

In the following sections, we focus on near-surface fields, including precipitation and air temperature. Figure 2 shows the IV of the 16-member ensemble of precipitation over four seasons, produced by 50km_no_nudg, 12km_no_nudg, and 12km_nudg. From Eq. 3 in Section 2.3, we computed the mean and the standard deviation of the IV from the 200 10-member ensembles, and then compare with the IV from the 16-member ensemble. If $\Delta_i > 2$, we added cross-hatchings on the i^{th} pixel of the maps in Fig. 2. However, there are no cross-hatchings over the entire domain for precipitation for any simulation. Indeed, the maximum Δ_i over the entire domain for all seasons was smaller than 0.5 for all simulations. This suggests that, the IVs of precipitation generated by the 200 10-member ensembles are not significantly different from those generated by the 16-member ensemble. Alexandru et al. (2007) also found that large ensembles (such as 10, 15, or 20 members) of domain/time-averaged and geospatial distribution of 850-hPa geopotential height IV converged, while small ensembles of two or five members lead to a large spread of IV.

In Fig. 2, there is a clear seasonal cycle for the IV for each set of simulations, with the largest IV in summer (June, July and August) and the smallest in winter (December, January and February). Notably, the geographic distribution of the IV of precipitation generated by nudged (12km_nudg) and non-nudged runs (50km_no_nudg and 12km_no_nudg) are different. Non-nudged runs show relatively large IVs over eastern North America during all four seasons, especially over the southeastern continental United States (CONUS) during summer months when large convective precipitation occurs. In contrast, non-nudged runs show relatively small IVs over western North America, especially the northwestern part of the domain. Alexandru et al. (2007) also found large precipitation IVs in the southeastern CONUS during summer using a regional climate model without spectral nudging. They suggested a link exists between 850-hPa geopotential height in the

northeast portion of the domain and precipitation in the southeast portion. Precipitation may act to trigger IV in geopotential height, which develops along the storm track and reaches a maximum in the northeast portion of the domain. The 500-hPa geopotential height in this study shows the same geographic distribution, with the largest IVs over the northeast portion of the domain (Fig. 3). Comparing the three sets of simulations, 12km_no_nudg has the largest IV, followed by the 50km_no_nudg, while 12km_nudg has the smallest IV. These differences can be seen in all four seasons and for both precipitation and geopotential height. Applying spectral nudging not only reduces the magnitude but also modifies the geographic pattern of IV.

Figure 4 shows the difference in IV of precipitation between each of the simulations with hatchings to indicate the statistical significance of the differences, using Eq. 4. In addition to the figures shown here, we also calculated the differences when using 12km_nudg as a reference, compared with 12km_no_nudg and 50km_no_nudg, as well as using 50km_no_nudg as a reference, compared with 12km_no_nudg. The maps are similar to Fig. 4 and hence not shown here. Overall, the largest differences in IV were between the nudged and non-nudged simulations, while the differences in IV generated by different spatial resolutions without nudging were much smaller. It should be noted that the differences in IV are statistically significant even between the 12km_no_nudg and 50km_no_nudg. The differences in IV between different model simulations were largest in summer, and smallest in winter and spring. Compared with simulations without nudging, the 12km_nudg generated smaller IV over most of North America in all four seasons except over the Midwest in summer. It is likely that in the Midwest, the 12km_nudg experienced more locally generated precipitation uninhibited by the lateral boundary processes, leading to larger IV. The 12km_nudg generates much smaller IVs over the southeast CONUS and

northeastern North America in all four seasons. The difference in IV generated by 50km_no_nudg and 12km_no_nudg, while small, was statistically significant over much of the model domain.

3.3 Geographic distribution of IV of temperature

While spectral nudging reduces the IV of precipitation, there is still a clear seasonal cycle of IV (Fig. 2). However, the IV of temperature was reduced by spectral nudging sufficiently enough to damp the IV seasonal cycle for temperature. Figure 5 shows the IV of the 16-member ensemble of temperature for four seasons, as produced by 50km_no_nudg, 12km_no_nudg, and 12km_nudg. From Eq. 3 in Section 2.3, we computed the mean and the standard deviation of the IV from the 200 10-member ensembles, and then compared with the IV from the 16-member ensemble. If $\Delta_i > 2$ we added cross-hatchings on the maps in Fig. 5. However, similar to precipitation, we did not have any hatchings over the entire domain for IV of temperature from the three sets of simulations. Indeed, the maximum Δ_i over the entire domain for all seasons was smaller than 0.4 for all three sets of simulations. This suggests that, the IV of temperature generated by the 200 10-member ensembles are not significantly different from those generated by the 16-member ensemble.

In Fig. 5, overall, the IV magnitudes produced by 12km_no_nudg and 50km_no_nudg were similar in summer and fall, but the IV generated by 12km_no_nudg was larger than that generated by 50km_no_nudg in spring and winter. The IVs of temperature produced by 12km_no_nudg and 50km_no_nudg also showed clear seasonal cycles. There was large IV over the Great Plains in summer, which moved to the Midwest in fall. The IV was reduced significantly in winter and spring over the CONUS.

Figure 6 shows the differences in IV of temperature between each of the simulations with hatchings indicating statistical significance, calculated using Eq. 4. In addition to the figures shown here, we

also calculated the differences when using 12km_nudg as a reference, compared with 12km_no_nudg and 50km_no_nudg, as well as using 50km_no_nudg as a reference, compared with 12km_no_nudg. These maps are similar to what we show here and hence not discussed further. Overall, the largest difference in IV was between the nudged simulation and non-nudged simulation. Compared with simulations without nudging, the 12km_nudg generated smaller IV over the entire domain and in all four seasons. The difference in IV generated by 50km_no_nudg and 12km_no_nudg were very small and not statistically significant in summer and fall.

3.7 IV versus interannual variability

Using Eqs. 5-7, we calculated CV_{16m} and CV_{10yr} from daily model output. The 10-year simulations and the 16-member ensemble of 1-year simulations used the same boundary condition and model configurations; see WK15 for more details about the model setup. The simulations compared here are at a spatial resolution of 12 km. Table 1 compares CV_{16m} for 12km_nudg and 12km_no_nudg and CV_{10yr} for the 10-year simulation of temperature. The values of CV are based on seasonal averages over seven subregions that are applied by the National Climate Assessment (NCA). For example, the CV_{16m} for summer temperature over the Midwest is 0.012 for simulations without nudging, and is 0.0007 for simulations with nudging, indicating that summer temperature simulated by different initial conditions varies by 1.2% (without nudging) and 0.07% (with nudging) of the ensemble mean. The CV_{10yr} of summer temperature over the Midwest is 0.039, indicating that the summer temperature in different years from 1995 to 2004 varies by 3.9% of the 10-year mean. In winter, the CV_{16m} for temperature over the Midwest is 0.078 for simulations without nudging, and is -0.007 for simulations with nudging, indicating that winter temperature simulated by different initial conditions varies by 7.8% (without nudging) and 0.7% (with nudging) of the ensemble mean. The CV_{10yr} of winter temperature over the Midwest is 0.778, indicating that

the winter temperature in different years from 1995 to 2004 varies by 77.8% of the 10-year mean. A comparison between the CV_{16m} and the CV_{10yr} suggests that, the IV generated by the 16-member simulations is not sufficiently large to affect the interannual variability of temperature. This is also true for the other six subregions and all seasons for temperature.

Table 2 compares the CV_{16m} for 12km_nudg and 12km_no_nudg and the CV_{10yr} for the 10-year simulation of precipitation. Compared with Table 1, both CV_{16m} and CV_{10yr} for precipitation are larger than those for temperature. For example, the CV_{16m} for summer precipitation averaged over the Midwest is 0.083 for the simulation without nudging, and 0.008 for the simulation with nudging, indicating that summer precipitation varies by 8.3% (without nudging) and 0.8% (with nudging) of the ensemble mean. The CV_{10yr} for summer precipitation over the Midwest is 0.15, indicating that summer precipitation in different years from 1995 to 2004 varies by 15% of the 10-year mean, which is much larger than CV_{16m} . In winter, the CV_{16m} for winter precipitation averaged over the Midwest is 0.03 for the simulation without nudging, and 0.0018 for the simulation with nudging, indicating that winter precipitation varies by 3.0% (without nudging) and 0.18% (with nudging) of the ensemble mean. The CV_{10yr} for winter precipitation over the Midwest is 0.365, indicating that winter precipitation in different years from 1995 to 2004 varies by 36.5% of the 10-year mean, which is much larger than CV_{16m} . This suggests that, similar to temperature, the IV generated by the 16-member ensembles is too small to contribute to the interannual variability of precipitation. This is also the case for the other six subregions and seasons for precipitation.

To assess the robustness of a 10-member ensemble, we calculated the CV_{10m} of the 200 10-member ensembles, as shown in Figs.7-10, with the median and percentiles shown in a boxplot. The CV_{16m} (red crosses) mostly overlap with the median CV_{10m} (bold black lines in the boxes). The difference between the CV_{16m} and the median CV_{10m} is smaller than the spread of the CV_{10m} from the 200

sub-samples. Indeed, all the CVs calculated from the 16-member ensemble over the seven regions and during all four seasons fall within the confidence intervals as calculated by $[\mu_{CV_{10m}} \pm 2\sigma_{d_{10vs16}}]$, described in Section 2.4 (see Tables S5-S6 in supporting information). This indicates that the IV estimates provided by a 10-member ensemble and by a 16-member ensemble are not significantly different. However, we do notice that the dispersion of the CV_{10m} varies across regions and seasons.

3.8 IV versus climate change Tables 3 and 4 compare the relative importance of IV versus projected climate change (PCC) of near-surface temperature and precipitation in the mid- and late-21st century under representative concentration pathway (RCP) 8.5. IV is one of the sources of uncertainty for the climate change signal (Hawkins and Sutton 2009). Both IVs and PCC in different seasons were calculated using monthly data. The PCC was computed as the difference between the corresponding future (2045–2054 or 2085–2094 under RCP8.5) and present (1995–2004) 10-year time periods. The values of IV and PCC were first calculated over each grid point, and then the regional averages were calculated. IVs in both present and future simulations contributed to uncertainties when estimating the climate change. Therefore, we expected the PCCs to be at least larger than the IV by a factor of two, to ensure the confidence of the PCC. As shown in Table 3, the IVs of modeled temperature are much smaller than the PCCs in all four seasons over all seven subregions in the mid- and late 21st century, indicating that the IV does not significantly affect the PCCs of regional and seasonal averaged temperature by the mid-century and beyond.

Unlike temperature, the IVs of simulated precipitation are much larger and show a clear seasonal cycle over the seven NCA subregions (Table 4), consistent with Fig. 2. The PCCs, however, are mostly smaller than the IVs. Therefore, we should expect large uncertainties in projected

precipitation change due to the IV of the climate system. For example, using WRF as a regional climate model, Gao et al. (2014), found a spring-drying signal in the Southwest in the late 21st century, which resulted from a decrease in precipitation and an increase in evapotranspiration. We find the dry signal in spring over the Southwest as well, and also in summer and winter over the Southern Great Plains. However, the IVs over these regions and during these seasons were larger than the PCCs, especially in summer. This demonstrates that precipitation is a difficult variable to simulate accurately, and may be poorly estimated from a single simulation. It is important to mention that, in both Table 3 and Table 4, the IVs are calculated using the 12km_nudg data, since our goal was to compare with the PCCs from the previous study, WK15. When the IV was calculated with the non-nudged simulation (not shown), the IVs increased for most subregions and during most seasons, except for the precipitation IV during summer over the Midwest, which was reduced, as shown in Fig. 4. Figures 11 and 12 show the IV calculated by the 200 10-member ensembles as well as the 16-member ensemble, indicating that the 10-member ensemble provided a fairly robust estimate of the model IVs for both temperature and precipitation.

4 Summary and discussion

The impacts of spatial resolution and spectral nudging on the IV of temperature and precipitation were investigated with a 16-member ensemble performed with the WRF model at spatial resolution of 12 km and 50 km. The 16-member ensemble simulation for each set of runs differed only by an offset of 12 hours from the starting date of the integrations to trigger IV. The output of the ensembles were analyzed over North America and seven subregions applied by the NCA over the CONUS. The experiment reinforced the notion from earlier studies that IV is sensitive to spatial resolution and spectral nudging. We add that the spectral nudging has a much greater effect on IV than spatial resolution. For all variables examined, smaller IVs were found in simulations with

coarser spatial resolution. Applying nudging reduced the magnitude of IV by a factor of two, and also modified the geographic distribution of IV in comparison with that generated by simulations without nudging. The IV does not affect the seasonal interannual variability at a regional scale, nor does it affect the climate change signal for temperature. However, the IV is an important source of uncertainty for climate change signals of precipitation. All of the above analyses were also conducted using 200 10-member ensembles taken as a sub-sample of the 16-member ensemble with uniform probability and without replacement, which enabled us to assess the uncertainty of the computed IV due to the size of the ensemble members. Our results showed that the IV generated by a 10-member ensemble can be considered similar to one generated using a 16-member ensemble and is able to provide a fairly robust estimate of IV, similar to that suggested by Alexandru et al. 2007.

As the climate changes, it is expected that certain types of weather/climate extremes will also change. Therefore, understanding the IV in the tails of the distribution of certain variables is also of major interest. However, due to computing constraints, we have only a 2-year simulation without nudging and a 3-year simulation with nudging, and therefore, the dataset is not sufficiently large enough to represent particularly rare events. Braun et al. (2012) carried out two 30-year simulations with the Canadian regional climate model, and analyzed the 10% and 90% quantiles (Q10 and Q90, hereafter) means. They found the IV of the Q10 and Q90 was roughly double the IV of the mean for both nudged and non-nudged runs over North America. This indicates that the use of spectral nudging does not affect the relationship between the IV of the mean or extreme values. However, the domain size does impact the relationship between the IV of extremes and the IV of the mean. For example, Braun et al. (2012) found that the IV in Q10 and Q90 was triple the IV of the mean over a smaller domain covering eastern Canada.

In order to investigate the possible dependence of IV on the integration horizon (i.e., one year versus multiple years), we continued a 10-member ensemble for all three sets of simulations for one more year, and conducted the same analysis as for 1995. While the amplitudes of the IV for all the examined fields were not exactly the same as 1995, there were no significant changes from one year to the other. The geospatial distribution and the seasonal pattern of the IVs in 1996 were very similar to those in 1995, as shown in the supporting information (Figures S1-S2, and Tables S1-S4). Indeed, Lucas-Picher also found different sizes of ensemble members generated different IVs (especially if the ensemble size was smaller than 10), but the amplitude of the IV in different years did not change significantly. In other words, they did not find any long-term tendency in the IV over the ten-year period. Therefore, conducting 10-member ensemble simulations for 1996 does not change the conclusions drawn for this study, for example, spectral nudging not only reduces the magnitude of IV but also modifies the geographic distribution of IV. The relationship between IV and interannual variability as well as the relationship between IV and climate change for both temperature and precipitation also stay the same.

This study aims to understand the IV at a spatial resolution of 12 km over North America with the same physics and model configurations (e.g., spectral nudging) as applied over future projections, as produced by WK15. We studied the IV by comparing with simulations at a spatial resolution of 50 km and at a spatial resolution of 12 km but without spectral nudging. We did not consider other impacts, such as physics parameterizations and domain size of the simulation on the IV. Thus, the detected IV is likely to be a lower bound of IV. We looked at the subregion-wide seasonal climatology and did not find that IV is more important than the interannual variability for temperature and precipitation. However, we expect the IV to be doubled or even tripled for smaller

time scales or for extremes (Braun et al. 2012). This conclusion is also true when comparing IV and climate change signals.

References

- Alexandru A, de Elía R, Laprise R (2007) Internal variability in regional climate downscaling at the seasonal time scale. *Mon Weather Rev* 135:3221–3238
- Braun M, Caya D, Frigon A, Slivitzky M (2012) Internal variability of the Canadian RCM's hydrological variables at the basin scale in Quebec and Labrador. *J Hydrometeor* 13:443–462. doi:[10.1175/JHM-D-11-051](https://doi.org/10.1175/JHM-D-11-051).
- Caya D, Biner S (2004) Internal variability of RCM simulations over an annual cycle. *Clim Dyn* 22:33–46
- Chen F, Dudhia J (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
- Christensen OB, Gaertner MA, Prego JA, Polcher J (2001) Internal variability of regional climate models. *Clim Dyn* 17:875–887
- Deser C, Knutti R, Solomon S, Phillips AS (2012) Communication of the role of natural variability in future North American climate. *Nat Clim Chang* 2(11):775–779. doi:[10.1038/nclimate1562](https://doi.org/10.1038/nclimate1562)
- Done, J. M., C. L. Bruyère, M. Ge, and A. Jaye (2014), Internal variability of North Atlantic tropical cyclones, *J. Geophys. Res. Atmos.*, 119, 6506–6519, doi:[10.1002/2014JD021542](https://doi.org/10.1002/2014JD021542).
- Gao Y, Leung LR, Lu J, Liu Y, Huang M, Qian Y (2014) Robust spring drying in the southwestern U.S. and seasonal migration of wet/dry patterns in a warmer climate. *Geophys Res Lett* 41: doi:[10.1002/2014GL059562](https://doi.org/10.1002/2014GL059562)
- Giorgi F, Bi X (2000) A study of internal variability of regional climate model. *J Geophys Res* 105:29503–29521
- Giorgi F, Mearns LO (1991) Approaches to the simulation of regional climate change: A review, *Rev. Geophys.*, 29(2): 191–216, doi:[10.1029/90RG02636](https://doi.org/10.1029/90RG02636).

- Girard E, Bekcic B (2005) Sensitivity of an arctic regional climate model to the horizontal resolution during winter: implications for aerosol simulation. *Int J Climatol* 25:1455–1471
- Grell GA, Devenyi D (2002) A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys Res Lett* 29:1693–1697
- Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. *Bull Amer Meteor Soc* 90: 1095–1107. doi:10.1175/2009BAMS2607.1
- Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008) Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J Geophys Res* 113: D13103. doi:[10.1029/2008JD009944](https://doi.org/10.1029/2008JD009944)
- Lucas-Picher P, Caya D, de Elía R, Laprise R (2008) Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain. *Clim Dyn* 31:927–940
- Morrison H, Thompson G, Tatarskii V (2009) Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon Wea Rev* 137: 991–1007. doi:<http://dx.doi.org/10.1175/2008MWR2556.1>
- Nikiéma O, Laprise R (2011) Budget study of the internal variability in ensemble simulations of the Canadian RCM at the seasonal scale. *J Geophys Res Atmos* 116:D16112. doi:[10.1029/2011JD015841](https://doi.org/10.1029/2011JD015841)
- Noh Y, Cheon WG, Hong SY, Raasch S (2003) Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound Layer Meteor* 107:401–427

- Vanitsem S, Chomé F (2005) One-way nested regional climate simulations and domain size. *J Clim* 18:229–233
- von Storch H, Langenberg H, Feser F (2000) A spectral nudging technique for dynamical downscaling purposes. *Mon Weather Rev* 128:3664–3673.
- Wang J, Kotamarthi VR (2013) Assessment of dynamical downscaling in near-surface fields with different spectral nudging approaches using the Nested Regional Climate Model (NRCM). *J Appl Meteor Climatol* 52:1576–1591
- Wang J, Kotamarthi VR (2014) Downscaling with a nested regional climate model in near-surface fields over the contiguous United States. *J Geophys Res Atmos* 119: 8778–8797.
doi:[10.1002/2014JD021696](https://doi.org/10.1002/2014JD021696)
- Wang J, Kotamarthi VR (2015) High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over North America. *Earth's Future* 3.
doi:[10.1002/2015EF000304](https://doi.org/10.1002/2015EF000304)

Figure and table captions

Fig. 1 500 hPa geopotential height of five of the ensemble members with a delay of 12 hr from initial conditions, valid at 0000 UTC, 30 June 1995 (6 simulated months after the integration starts; left panel) and 0000 UTC, 30 December 1995 (12 simulated months after the integration start; right panel) for 12km_no_nudg, 50km_no_nudg, and 12km_nudg

Fig. 2 Geographic distributions of IV of precipitation amount (mm/day), calculated from the 16-member ensemble, for four seasons during 1995. Left to right: 50km_no_nudg, 12km_no_nudg, and 12km_nudg.

Fig. 3 The same as **Fig. 2**, but for geopotential height (dam) at 500 hPa.

Fig. 4 Differences in IV of precipitation (unit: mm/day) between each two sets of simulations. Left to right: 12km_no_nudg versus 12km_nudg; 50km_no_nudg versus 12km_nudg; and 12km_no_nudg versus 50km_no_nudg. The hatchings indicate the differences in IV between simulations are statistically significant at the 5% level.

Fig. 5 The same as **Fig. 2**, but for daily averages of near-surface air temperature (°C).

Fig. 6 The same as **Fig. 4**, but for temperature (°C).

Fig. 7 Boxplot of the coefficient of variations (CV) of temperature using the 200 10-member ensembles over the seven subregions for four seasons. The simulations are at a spatial resolution of 12 km without spectral nudging. The red cross is the CV calculated using the 16-member ensemble, as presented in Table 1. Boxes indicate the 25th and 75th quantiles, with a horizontal line indicating the median and whiskers showing the extreme range of CVs calculated from the 200 10-member ensembles.

Fig. 8 The same as Fig. 7 but for the simulation at 12km with spectral nudging.

Fig.9 The same as Fig. 7, but for precipitation.

Fig.10 The same as Fig. 8, but for precipitation.

Fig. 11 Boxplot of IVs for temperature using the 200 10-member ensembles over the seven subregions for four seasons. The simulations are at a spatial resolution of 12 km with spectral nudging. The red cross is the IV calculated using the 16-member ensemble, as presented in Table 3. Boxes indicate the 25th and 75th quantiles, with a horizontal line indicating the median and whiskers showing the extreme range of IVs calculated from the 200 10-member ensembles.

Fig. 12 The same as Fig. 11, but for precipitation

Table 1 Coefficient of variations (CV) of near-surface air temperature for a 16-member ensemble generated by simulations with and without nudging at a spatial resolution of 12km in 1995 and a 10-year simulation (1995–2004).

Table 2 The same as **Table 1**, but for precipitation.

Table 3 IV versus projected climate change for temperature during 2045–2054 and 2085–2094 under RCP 8.5. Spectral nudging is applied to all the simulations listed in the table. Units: Celsius.

Table 4 The same as **Table 3**, but for precipitation. Units: mm/day