

Scale-dependent Regional Climate Predictability over North America Inferred from CMIP3 and CMIP5 Ensemble Simulations

Fuqing ZHANG^{*1}, Wei LI^{1,2}, and Michael E. MANN¹

¹*Department of Meteorology and Center for Advanced Data Assimilation and Predictability Techniques, The Pennsylvania State University, University Park, PA 16802, USA*

²*IMSG at NOAA/NWS/NCEP/Environmental Modeling Center, College Park, MD 20740, USA*

(Received 19 January 2016; revised 16 March 2016; accepted 11 April 2016)

ABSTRACT

Through the analysis of ensembles of coupled model simulations and projections collected from CMIP3 and CMIP5, we demonstrate that a fundamental spatial scale limit might exist below which useful additional refinement of climate model predictions and projections may not be possible. That limit varies among climate variables and from region to region. We show that the uncertainty (noise) in surface temperature predictions (represented by the spread among an ensemble of global climate model simulations) generally exceeds the ensemble mean (signal) at horizontal scales below 1000 km throughout North America, implying poor predictability at those scales. More limited skill is shown for the predictability of regional precipitation. The ensemble spread in this case tends to exceed or equal the ensemble mean for scales below 2000 km. These findings highlight the challenges in predicting regionally specific future climate anomalies, especially for hydroclimatic impacts such as drought and wetness.

Key words: regional climate predictability, CMIP5 ensemble, North America climate change

Citation: Zhang, F. Q., W. Li, and M. E. Mann, 2016: Scale-dependent regional climate predictability over North America inferred from CMIP3 and CMIP5 ensemble simulations. *Adv. Atmos. Sci.*, **33**(8), 905–918, doi: 10.1007/s00376-016-6013-2.

1. Introduction

There is widespread scientific consensus that the accumulation of greenhouse concentrations from fossil fuel burning and other human activities is leading to a warming of the globe and other associated changes in large-scale climate (IPCC, 2013). Most assessments indicate that the cost of the resulting damage from climate change will rise to several percent of the global economy in the decades ahead if left unchecked. Yet, our ability to assess the regional impacts of climate change, which are critical both to assessing the damage caused by climate change and the implementation of adaptive strategies, remains hampered by the remaining substantial uncertainties associated with regional climate projections (Murphy et al., 2004; Tebaldi et al., 2005; Hawkins and Sutton, 2009; Deser et al., 2012; Watterson et al., 2014).

Regional climate projections are typically derived by one of two methods: statistical downscaling or dynamical downscaling (IPCC, 2013). In the former case, statistical relationships between coarse and fine scales derived from modern climate data are used to take coarse-scale climate model predictions/projections and estimate the likely impact on climate

statistics at finer spatial and temporal scales. In the latter case, information from coarse climate models is used as boundary constraints on a finer resolution model (a regional climate model) that resolves the smaller spatiotemporal scales of interest. In either case, there is an assumption of a predictable relationship between the large scales captured in the coarse climate model projection and the local scales sought by the downscaling method.

Downscaled climate model projections have increasingly been used as guidance for policymakers and stakeholders at the local, national, and international level in assessing potential impacts and risks associated with human-caused climate change (von Storch et al., 1993; Mearns et al., 1999; Jones et al., 2011). However, the reliability of these projections continues to be debated. There is clearly skill in the largest-scale quantities; for example, the observed increase in global mean temperature (and even continental mean temperatures) can be detected and attributed to anthropogenic climate change (IPCC, 2007). However, confidence in regional-scale projections of surface temperature and precipitation is considerably lower (Whetton et al., 2007; Separovic et al., 2008; Watterson and Whetton, 2011; Deser et al., 2012; Li et al., 2012).

In the current study, we seek to quantify the predictability of regional-scale climate change, with an emphasis on surface temperature and precipitation, across the coterminous United

* Corresponding author: Fuqing ZHANG
Email: fzhang@psu.edu

States and surrounding areas, through analysis of the multimodel ensembles of coupled model simulations collected from both CMIP3 and CMIP5. In section 2 we describe the data and methods used in the study. In section 3 we present an analysis of regional climate predictability based on CMIP5 multimodel historical simulations. In section 4 we provide a complementary analysis of CMIP3 multimodel simulations, analyzing both historical simulations and future projections. Conclusions are presented in section 5.

2. Data and methodology

We analyze surface temperature and precipitation fields across the coterminous United States and neighboring oceanic regions (15°–60°N, 70°–130°W), as derived from both observational data and an ensemble of climate model simulations.

Observational data analyzed include monthly mean 5° latitude × 5° longitude grid-box near-surface temperature anomalies over 1850–2014 from HadCRUT4. We add the reference period climatology over 1961–90 to yield absolute surface temperatures. Precipitation data are taken from the 1° latitude × 1° longitude GPCC dataset over the period 1979–2004.

For the climate model simulations, monthly mean surface (2 m) air temperature and precipitation spanning the period 1979–2004 are available for 38 climate models in the historical late 19th to early 21st century 20C3M CMIP5 simulation archives, and 18 for CMIP3 (Taylor et al., 2012) (Table 1). Where multiple simulations are available for a given model, a single ensemble mean is calculated to ensure that each distinct model is represented equally in the ensuing analysis.

We focus our analysis on the boreal summer (June–August) climatological period during the 1979–2004 period of overlap between observations and model simulations. Both observational and model data are interpolated to a common (T85, ~1.4° latitude × 1.4° longitude) spatial resolution prior to analysis. For the purpose of the ensuing analyses, we define the following terms:

(1) Ensemble mean: the uniform arithmetic mean of all ensemble members;

(2) Ensemble spread: the uniform arithmetic mean of the absolute difference between any two members across all ensemble members;

(3) Ensemble mean error or bias: the (signed) difference between the model ensemble mean and the observational analysis.

In addition to evaluating the ensemble mean, ensemble spread, and error for the observational and model fields themselves, we perform a power spectral analysis (using the Fast Fourier transform) of the fields to evaluate the power spectral density (PSD) characteristics of the various quantities in wavenumber space. In these analyses:

(1) The ensemble mean power spectrum (P) is defined as the PSD of the non-weighted arithmetic mean (V_m) for all ensemble members ($M = 38$):

$$P = \text{PSD}(V_m), V_m = (1/M) \sum_i V_i, M = 38;$$

(2) The PSD of the ensemble spread is defined as:

$$\Delta P = (1/N) \sum_{k=1}^N P_d, P_d = \text{PSD}(V_i - V_j), \text{ where } i \text{ and } j \text{ are any pair of ensemble members and } N = 703.$$

(3) The PSD of the ensemble mean error or bias is defined as the power spectra of the difference between the ensemble mean and the observation, i.e.,

$$P' = \text{PSD}(V_m - V_o), \text{ where } V_m \text{ and } V_o \text{ refer to the ensemble mean and observational variables, respectively.}$$

The ratio of the PSD ensemble mean (signal) and ensemble spread (noise) as a function of wavenumber defines a scale-dependent SNR measure (Bei and Zhang, 2007). A ratio smaller than unity indicates that the noise amplitude is greater than the signal amplitude, and implies that model estimates of mean changes are unreliable at that spatial scale. Wavenumber 1 corresponds to a single sinusoidal fluctuation over the entire circle of latitude, i.e., it is the coarsest possible measure of zonal variability (wavenumber 0 represents the zonal mean). Since the selected U.S. domain is precisely 1/6 of a circle of latitude, the lowest resolvable wavenumber for that domain is global wavenumber 6. The horizontal scale (wavelength) for a given wavenumber is the length of the circle of latitude divided by the wavenumber, e.g., the length scale corresponding to wavenumbers 1, 2, 3, 4, 5, 6 and 9 are ~36 000, 18 000, 12 000, 9000, 7200, 6000 and 4000 km, respectively.

3. CMIP5 multimodel historical experiments

We form estimates of the internal variability in the climate means of summertime precipitation and surface air temperature during the period 1979–2004 using the multimodel ensemble of 38 CMIP5 historical simulations (Fig. S1 in electronic supplementary material). The mean difference between any two models within the 38-model ensemble is defined as the ensemble spread; a measure of uncertainty of any deterministic prediction assuming the truth (as well as any single deterministic prediction) is a random draw out of the multimodel ensemble. The ensemble spread can be regarded as a lower bound on the model uncertainty, since it neither accounts for the potential bias due to deficiencies in model physics that are common among models, nor uncertainties in forcing (both anthropogenic and natural).

Figures 1a and b show the resulting mean and ensemble spread of surface temperature and monthly precipitation over the U.S. domain. The largest uncertainty for the surface temperature field is found over the western U.S., with the ensemble spread as high as 5°C–6°C; followed by the central U.S., with a spread of 3°C–5°C; and the eastern U.S., with a spread of ~1.5°C–3°C (all higher than the surrounding oceans, at ~0.5°C–1.5°C). It is worth noting that the warming trend over the U.S. during the past century is on the order of 1°C (Ji et al., 2014), though it is beyond the scope of this study to examine the scale-, variable- and location-dependent predictability of the climate trend.

The large uncertainty in the mean surface temperature

Table 1. List of the CMIP5 models used in this study (IPCC, 2013).

Model	Sponsor	Atmospheric resolution (lat × lon)
BCC_CSM1.1	Beijing Climate Center, China Meteorological Administration	2.8° × 2.8°
BCC_CSM1.1(m)	Beijing Climate Center, China Meteorological Administration	2.8° × 2.8°
CanCM4	Canadian Centre for Climate Modeling and Analysis	2.8° × 2.8°
CanESM2	Canadian Centre for Climate Modeling and Analysis	2.8° × 2.8°
CCSM4	National Center for Atmospheric Research	0.9° × 1.25°
CESM1-BGC	Community Earth System Model Contributors	0.9° × 1.25°
CESM-CAM5	Community Earth System Model Contributors	0.9° × 1.25°
CESM1-FASTCHEM	Community Earth System Model Contributors	0.9° × 1.25°
CESM1-WACCM	Community Earth System Model Contributors	1.875° × 2.5°
CMCC-CESM	Centro Euro-Mediterraneo per I Cambiamenti Climatici	3.75° × 3.75°
CMCC-CM	Centro Euro-Mediterraneo per I Cambiamenti Climatici	0.75° × 0.75°
CMCC-CMS	Centro Euro-Mediterraneo per I Cambiamenti Climatici	1.875° × 1.875°
CNRM-CM5-2	Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	1.4° × 1.4°
CNRM-CM5	Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	1.4° × 1.4°
CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	1.875° × 1.875°
EC-EARTH	EC-EARTH consortium	1.125° × 1.125°
GFDL CM2.0p1	NOAA Geophysical Fluid Dynamics Laboratory	2° × 2.5°
GFDL CM3	NOAA Geophysical Fluid Dynamics Laboratory	2° × 2.5°
GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	2° × 2.5°
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	2° × 2.5°
GISS-E2-H-CC	NASA Goddard Institute for Space Studies	2° × 2.5°
GISS-E2-H	NASA Goddard Institute for Space Studies	2° × 2.5°
GISS-E2-R-CC	NASA Goddard Institute for Space Studies	2° × 2.5°
HadCM3	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	2.5° × 3.75°
HadGEM2-AO	National Institute of Meteorological Research/Korea Meteorological Administration	1.25° × 1.9°
HadGEM2-CC	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	1.25° × 1.9°
HadGEM2-ES	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	1.25° × 1.9°
INM-CM4.0	Institute for Numerical Mathematics	1.5° × 2°
IPSL-CM5A-LR	L'Institut Pierre-Simon Laplace	1.875° × 3.75°
IPSL-CM5A-MR	L'Institut Pierre-Simon Laplace	1.25° × 2.5°
IPSL-CM5B-LR	L'Institut Pierre-Simon Laplace	1.875° × 3.75°
MIROC5	Atmospheric and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1.4° × 1.4°
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmospheric and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8° × 2.8°
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmospheric and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8° × 2.8°
MRI-CGCM3	Meteorological Research Institute	1.125° × 1.125°
MRI-ESM1	Meteorological Research Institute	1.125° × 1.125°
NorESM1-ME	Norwegian Climate Centre	2.5° × 2.5°
NorESM1-M	Norwegian Climate Centre	1.875° × 2.5°

field (Fig. 1a) can be interpreted through a parallel assessment using the observational (HadCRUT4) surface temperature during the overlapping time period (Fig. 1c). The error/bias can be estimated as the model ensemble mean minus the observations. The domain-averaged ensemble spread ($\sim 2.1^{\circ}\text{C}$) is found to be larger but grossly comparable to the domain-averaged root-mean-square of this estimate of error/bias (1.2°C). Moreover, the spatial pattern of the error/bias is similar to that of the ensemble spread, with the western U.S. displaying the largest mean error, followed by the Great Plains in the central U.S., and finally the eastern U.S. There are, however, some notable differences as well. Of particular interest is the relatively low spread over the North American west and east coasts and neighboring ocean regions (Fig. 1a), which contrasts with the large error/bias estimates over these same regions (Fig. 1c). This suggests the presence of a systematic bias that is common to most or all of the climate models, perhaps associated with deficiencies in the

models' representations of land–sea contrast or continental sea-breeze circulations.

The region of maximum uncertainty (ensemble spread) for precipitation (Fig. 1b) is found over lower latitudes (the south central U.S. and Latin America), with an ensemble spread exceeding 2.5 mm d^{-1} —roughly half the amplitude of the observed mean (signal). A second uncertainty maximum in precipitation is located over the northern Great Plains in the lee of the Rockies, with an ensemble spread exceeding 1.5 mm d^{-1} . The spatial pattern of the ensemble mean error (Fig. 1d) is once again grossly consistent with that for the ensemble spread (uncertainty), in that regions of peak amplitude are similar (e.g., common maxima along the southern edge of the domain and northern Great Plains), though the ensemble mean errors of approximately -2.5 mm d^{-1} are considerably larger than the ensemble spread for the Gulf Coast and Florida Peninsula. In addition, the North American domain-mean absolute ensemble mean error and spread

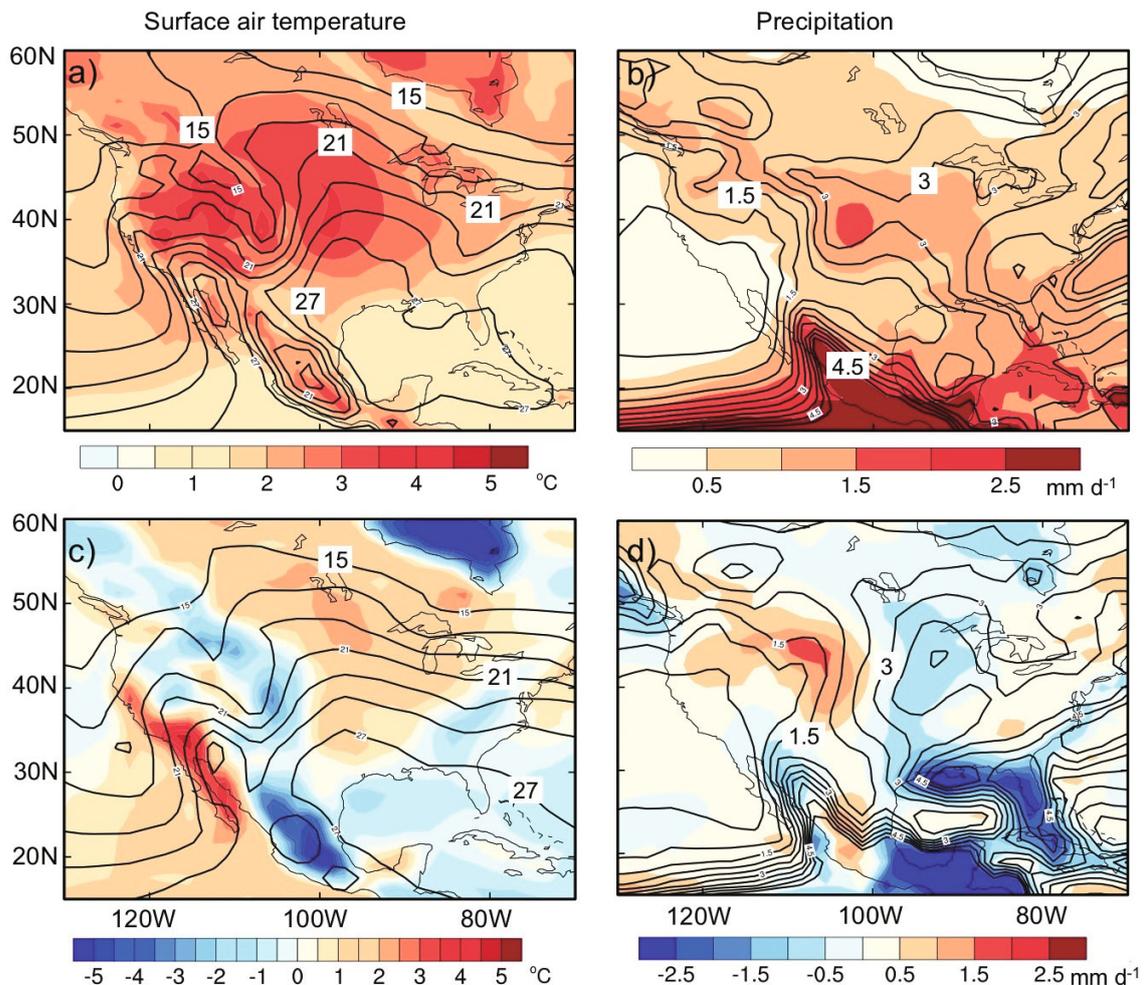


Fig. 1. Ensemble mean and spread/error for CMIP5 historical simulations and observations of summer (June–August) surface air temperature and precipitation over the U.S. during 1979–2004. Top: ensemble mean (contours) and ensemble spread (color-shaded) for (a) surface air temperature and (b) precipitation. Bottom: observational mean (contours) and error (color-shaded) for (c) surface air temperature and (d) precipitation. The corresponding domain-mean spreads are 2.09°C for (a), 0.97 mm d^{-1} for (b), 1.16°C for (c), and 0.58 mm d^{-1} for (d).

are also comparable in magnitude (0.58 mm d^{-1} and 0.97 mm d^{-1} , respectively).

Unlike surface temperature, which is primarily determined by large-scale processes, precipitation is heavily influenced by smaller-scale processes including moist convection, land–sea contrast, and orographic lifting. This distinction is exemplified by the local maxima for both ensemble mean error and spread over the Gulf of Mexico (hot spot of convection) and the mountainous areas of the western U.S. (where orographic effects are important). Comparing Figs. 1b and d suggests that, for the CMIP5 simulations of climatological mean precipitation, the ensemble spread can be used qualitatively to assess the uncertainty in the ensemble mean estimate. To place the U.S. results in a broader perspective, we also compare the ensemble mean, spread, and error for the global domain (not shown). The basic results discussed above appear to apply at this larger scale as well (though a detailed analysis of the global domain is beyond the scope of the current study).

The spatial scale–dependence of the predictability of surface temperature and precipitation is quantified by evaluating the PSD along both global circles of latitude and a latitudinal/longitudinal sub-region containing the coterminous U.S. (15° – 60° N, 70° – 130° W). Figure 2 shows the ensemble mean (left) and ensemble spread (middle) PSD for the CMIP5 surface temperature and precipitation fields, along with the ratio of the ensemble mean to the ensemble spread, i.e., the SNR (right) as a function of global wavenumber. For the global circle of latitude ensemble mean temperature (Fig. 2a), the PSD exhibits a peak at lower wavenumbers (1–3) for the midlatitudes (40° – 60° N); while for the subtropics (20° – 40° N), three distinct spectral peaks are observed (wavenumbers 1, 3 and 5). By contrast, for the ensemble spread, the PSD (Fig. 2b) decreases quite gradually in both the midlatitudes and the subtropics, though greater amplitudes are found across all wavenumbers for the former. The SNR (Fig. 2c) exceeds unity at all latitude and wavenumber ranges, with the exception of (1) wavenumber 4 between 40° – 60° N, (2) wavenumber 6 poleward of 50° N, and (3) wavenumber 9 between 45° – 55° N. Given the SNRs, surface temperature projections can be considered most reliable for wavenumbers 1–2 in the midlatitudes, and wavenumbers 1, 3 and 5 within the subtropics. Therefore, meaningful surface temperature predictions ($\text{SNR} > 1$) appear possible over a somewhat broad range of latitudes and wavenumbers.

For the more limited U.S. sub-region, the ensemble mean (Fig. 2d) and spread (Fig. 2e) are both larger at lower wavenumbers than for their global counterparts, but the SNR (Fig. 2f) falls below unity for global wavenumber 12 (horizontal scale of 30° longitudinal variation, i.e., distances of $\sim 3000 \text{ km}$) over the central latitudes of the U.S. (35° – 45° N), and for nearly all wavenumbers greater than 36 (scales less than 10° in longitudinal distance, i.e., distances less than $\sim 1000 \text{ km}$). This observation implies that state-of-the-art (i.e., CMIP5) climate model projections are likely to exhibit very limited skill in predicting regional variations in surface temperature at scales below 1000 km . It is noteworthy that

wavenumber 18 ($\sim 20^{\circ}$ or $\sim 2000 \text{ km}$ in longitudinal distance) exhibits the maximum SNR at nearly all latitudes for the U.S. domain. We interpret this observation as indicative of the influence of topographical features in the U.S. that induce enhanced predictability at this characteristic spatial scale.

The findings for precipitation (Figs. 2g and h) are quite different from those for surface temperature (Figs. 2a and b). Precipitation exhibits greater spectral amplitude in the subtropics relative to the midlatitudes, especially for lower (1–2) wavenumbers. SNRs at the global scale (Fig. 2i) are generally lower, substantially exceeding unity only for wavenumbers 1–2 between 20° N and 50° N, and wavenumber 4 between 40° N and 60° N. Low predictability ($\text{SNR} < 1$) is observed even at wavenumbers 1–2 poleward of 50° N, implying considerable challenges in predicting regional-scale variations in precipitation at high latitudes. Interestingly, however, for the U.S. regional sub-domain (Figs. 2j–l), there are apparently predictable signals ($\text{SNR} > 1$) for global wavenumbers 6–12 at nearly all latitudes, and for even higher wavenumbers (24–60, i.e., scales as small as 600 km) in the central U.S. latitudes (35° – 45° N). The larger signals over these latitudes in the North American domain may be related to regional-scale terrain effects and land–ocean contrasts, although some models may still have deficiencies in simulating these effects.

To further assess the scale and latitude dependence of surface temperature and precipitation predictability over the U.S. sub-domain, we average the fields over three representative latitude ranges (low latitude, 15° – 30° N; midlatitude, 30° – 45° N; and high latitude, 45° – 60° N; see Fig. 3). Given that the observations represent a single realization drawn from a larger distribution of possible climate histories, if the model ensemble accurately reflects the true climate, the PSD of the observations should be similar to that of individual ensemble members, and the ensemble mean should reflect the approximate mode of the distribution. On the other hand, the PSD of the ensemble spread (representing the uncertainty) should closely resemble that of the difference between the ensemble mean and observations (error/bias) across wavenumbers.

For the global domain, the PSD of the ensemble mean and observational mean are indeed similar for all latitude ranges for both surface temperature and precipitation. An exception is the anomalously low PSD values for surface temperature at wavenumber 2 and those exceeding ~ 50 , the latter of which we attribute to the low spatial density of surface temperature observations over the open ocean. The PSD for the ensemble spread generally exceeds that of the ensemble error/bias at most wavenumbers, and especially at lower wavenumbers (< 10) and for surface temperature. Consistent with our earlier findings (Fig. 2), the PSD for both the ensemble mean and observations (i.e., the signals) exceed those for the ensemble spread and error (noise or uncertainties) for wavenumbers 1–20 for all three latitude ranges for surface temperature (Figs. 3a–c), implying predictability across the associated spatial scales. For precipitation, by contrast, predictability is only evident (Figs. 3g–i) for wavenumbers 1–3 for the low-latitude (15° – 30° N) and midlatitude (30° – 45° N) zone, and

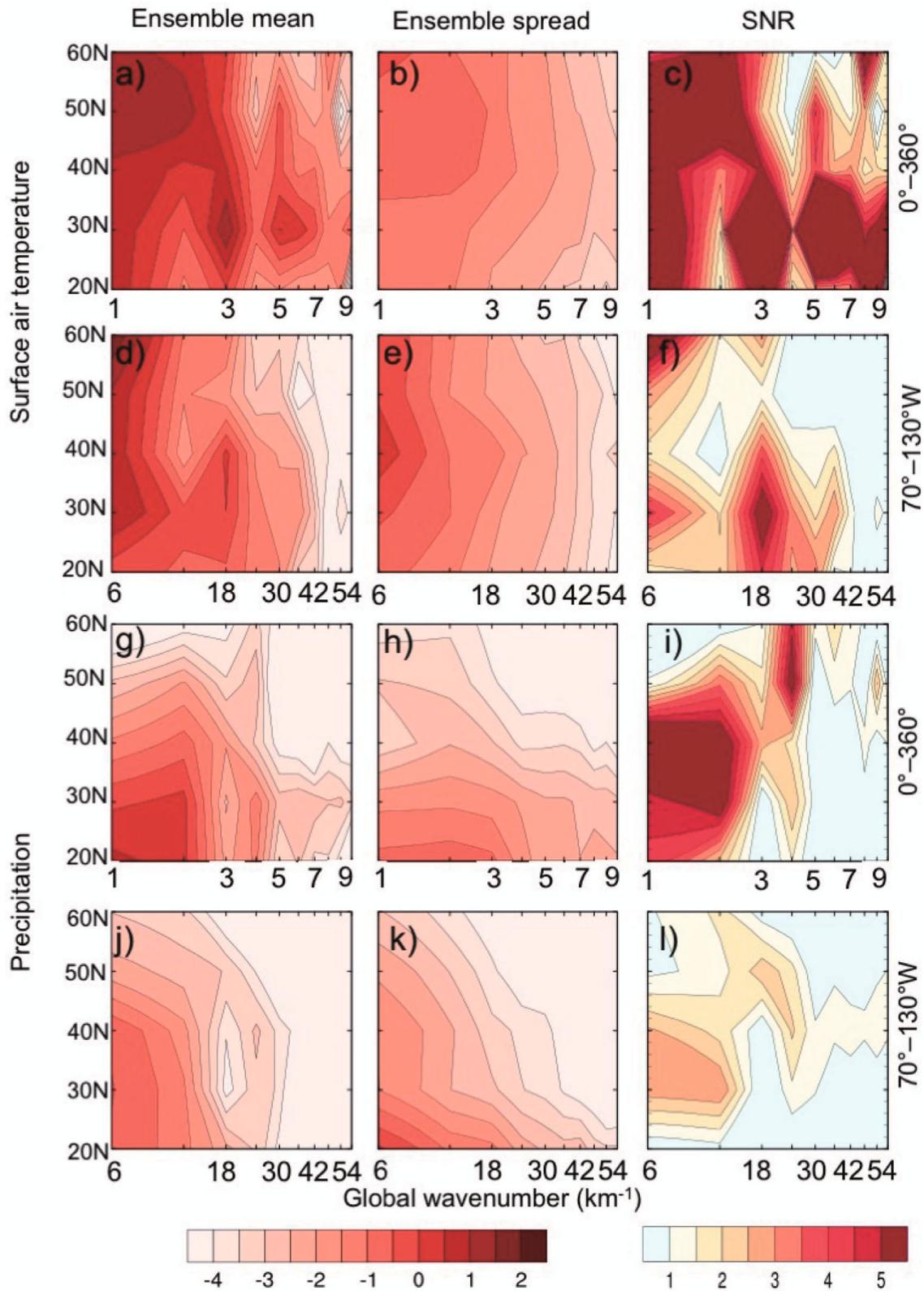


Fig. 2. Wavenumber–latitude distribution of the PSD of (a–f) surface air temperature and (g–l) precipitation over the global (0°–360°) and U.S. regional (70°–130°W) sub-domain. Shown are the PSDs for the ensemble mean, i.e., signal (left); ensemble spread, i.e., noise (middle); and ratio of the former to the latter, i.e., the SNR (right). The PSD amplitude scale is logarithmic.

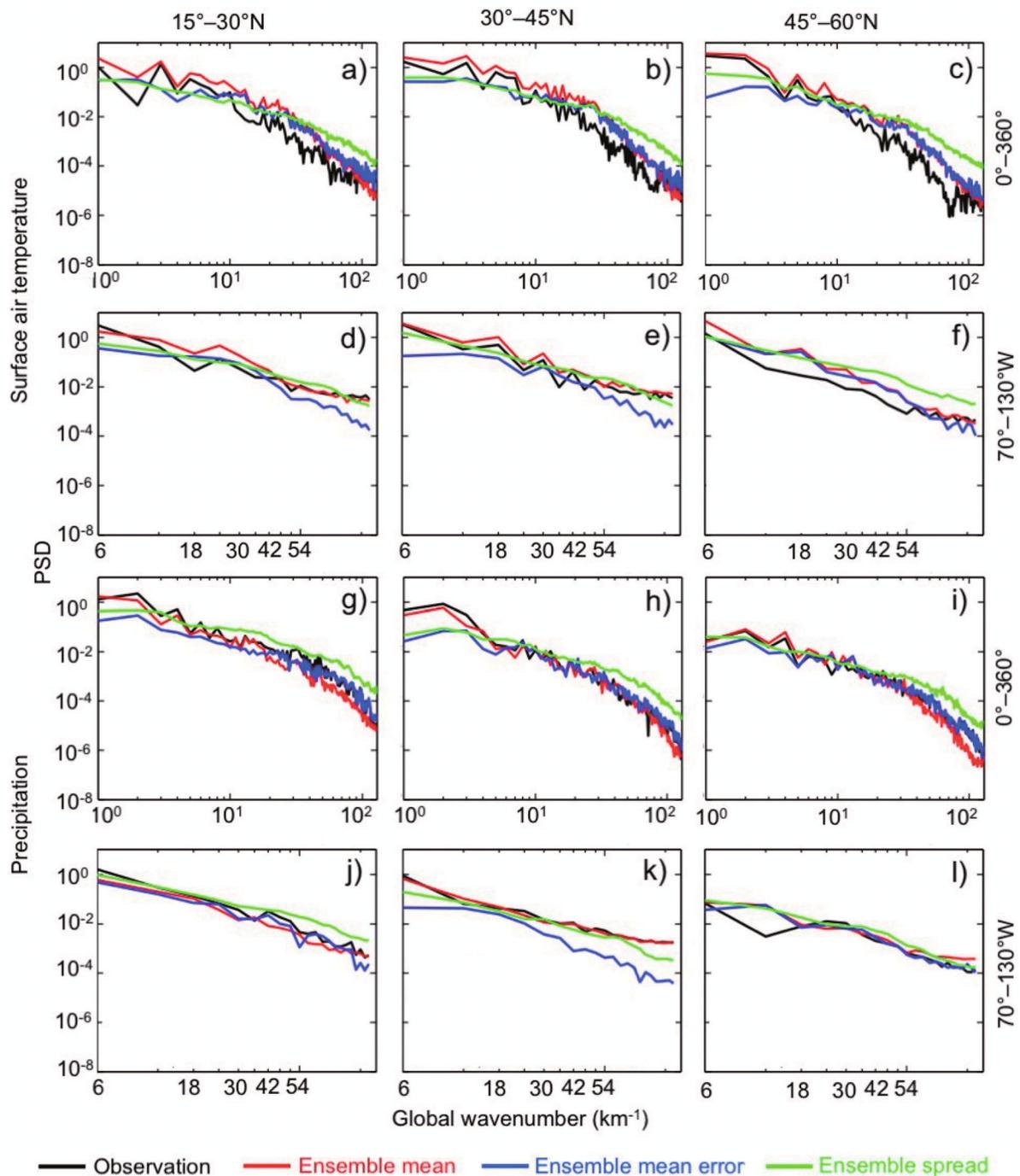


Fig. 3. PSD of the observations (black), ensemble mean (red), ensemble error (blue) and ensemble spread (green) for summer (June–August) (a–f) surface air temperature and (g–l) precipitation over the global (0° – 360° E) and U.S. regional (70° – 130° W) sub-domain averaged over three different latitude bands (left, 15° – 30° N; middle, 30° – 45° N; right, 45° – 60° N). Scales for both axes are logarithmic.

for almost no wavenumbers for the high-latitude (45° – 60° N) zone.

For the U.S. regional sub-domain, the PSD for the ensemble mean is generally consistent with that for the observations for both surface temperature and precipitation, and low and intermediate wavenumbers. However, for the high-latitude

zone (45° – 60° N) the ensemble-mean PSD considerably exceeds that of the observations for higher (> 24 for surface temperature and > 12 for precipitation) wavenumbers. The discrepancy between the ensemble spread and the ensemble mean error/bias is considerably greater for the U.S. regional domain than for the global domain as well.

The inferred predictability of surface temperature and precipitation for the U.S. regional domain varies considerably between the two variables and three latitude ranges (Figs. 3d–f, j–l). For example, the SNR for surface temperature exceeds unity for all wavenumbers lower than 36 (spatial scales as small as ~ 1000 km) for the low-latitude (15° – 30° N) zone, but the SNR is close to the “no predictability” value of unity for nearly all wavenumbers for the midlatitude (30° – 45° N) and high-latitude (45° – 60° N) zones. For precipitation, only for the midlatitude zone (30° – 45° N) is there evidence of predictability, and at fairly low (6–12) global wavenumbers (i.e., spatial scales no less than ~ 6000 km). These examples highlight the challenge for regional-scale climate predictability in North America with existing state-of-the-art global climate models.

4. CMIP3 historical experiments and future projections

To further investigate the robustness of our findings based on the CMIP5 historical simulations (Figs. 1–3) we perform

parallel analyses using the CMIP3 (Meehl et al., 2007; Watterson et al., 2014) (Table 2) multimodel ensemble simulations using both (1) the same historical period (1979–2004) (Zhou and Yu, 2006; Timm and Diaz, 2009) and (2) the CMIP3 (“A2” scenario) 21st century climate change projections.

Our conclusions regarding the predictability of regional-scale climate over North America with the CMIP5 historical simulations (Figs. 1–3) are similar to those obtained with the CMIP3 historical multimodel simulations (Figs. 4–6), with only one minor discrepancy: slightly lower SNR values are found for both the global surface temperature and precipitation fields over the midlatitudes of North America for global wave numbers 6–12. The fact that little-to-no improvement in regional predictability is found to result from the substantial model development and improvement reflected by the 5-year period between CMIP3 and CMIP5 suggests that, even with increasingly refined and detailed climate models, our conclusions regarding an apparent scale limit for regional-scale climate predictability are likely to remain true.

Similar conclusions are also obtained for the 21st century projections using the “A2” emissions scenario for the pe-

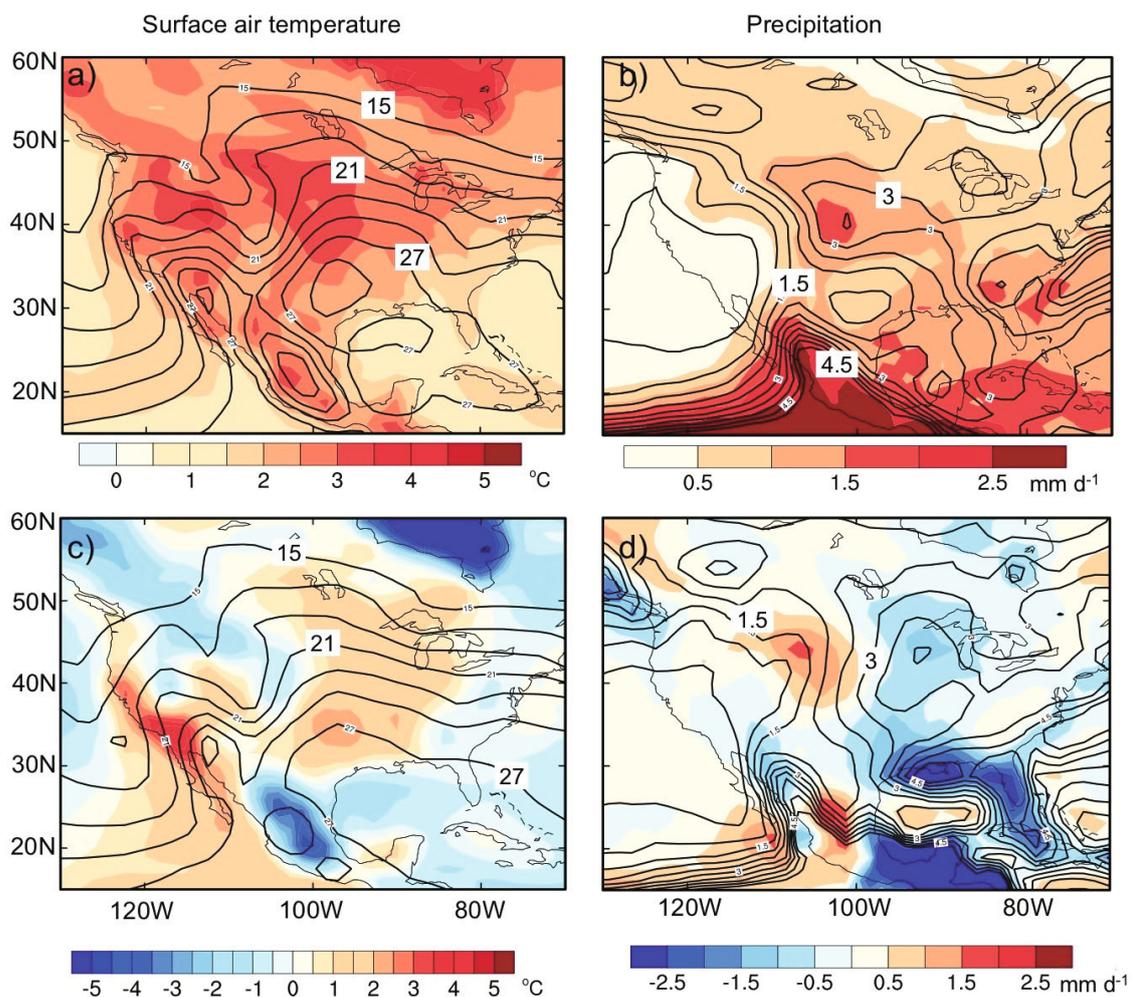


Fig. 4. As in Fig. 1 but using CMIP3 historical simulations.

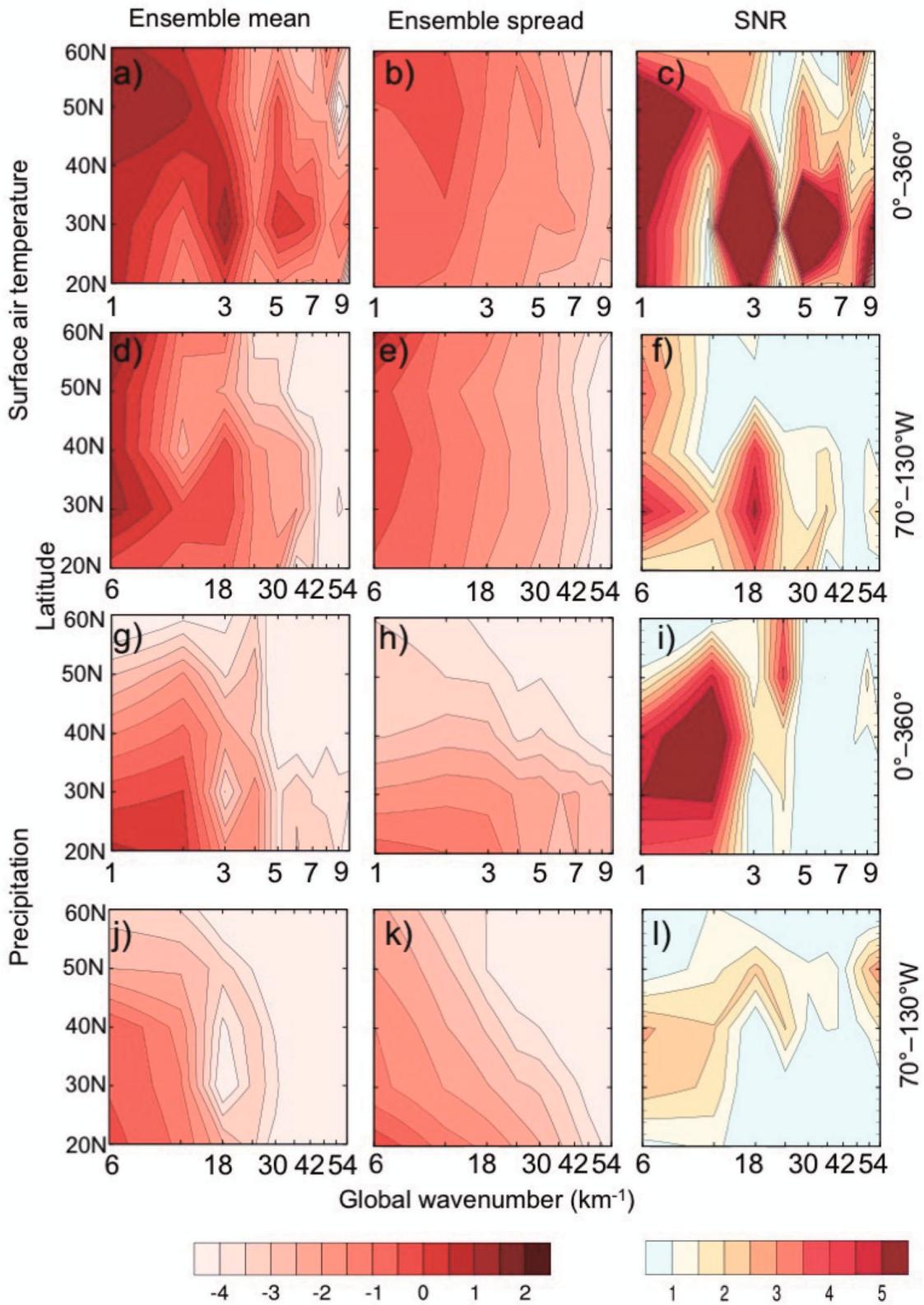


Fig. 5. As in Fig. 2 but using CMIP3 historical simulations.

riod 2074–99 using the CMIP3 multimodel ensemble simulations (Figs. 7–9). Even though obviously we do not have observations to verify these simulations, our conclusions again

remain mostly unchanged: there is an apparent scale limit by which the uncertainty in the prediction (noise) becomes greater than the ensemble mean prediction (signal). This fur-

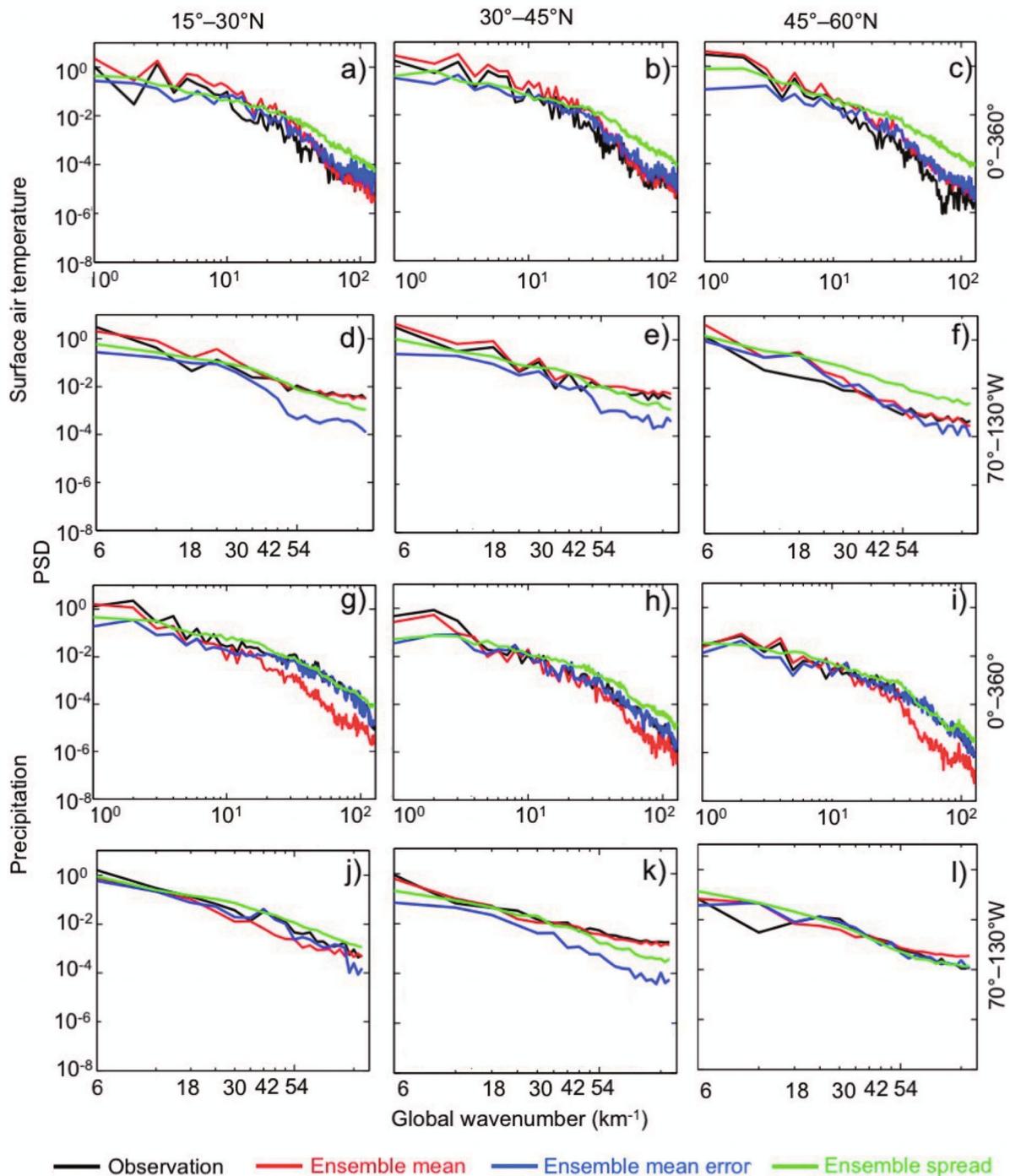


Fig. 6. As in Fig. 3 but using CMIP3 historical simulations.

ther highlights the limited predictability of climate models, especially at regional scales for different climate scenarios.

5. Concluding remarks

In summary, through an analysis of surface temperature and precipitation variability in the CMIP5 historical simu-

lations and comparisons with observational data during the overlapping (1979–2014) interval of the late 20th/early 21st century, we have found that there appears to be a fundamental scale limit below which refinement of climate model predictions may not be possible. While the predictability limit depends on the variables and regions analyzed, the averaging period and/or season, a seemingly robust result is that, for North America, the uncertainty due to intrinsic noise ap-

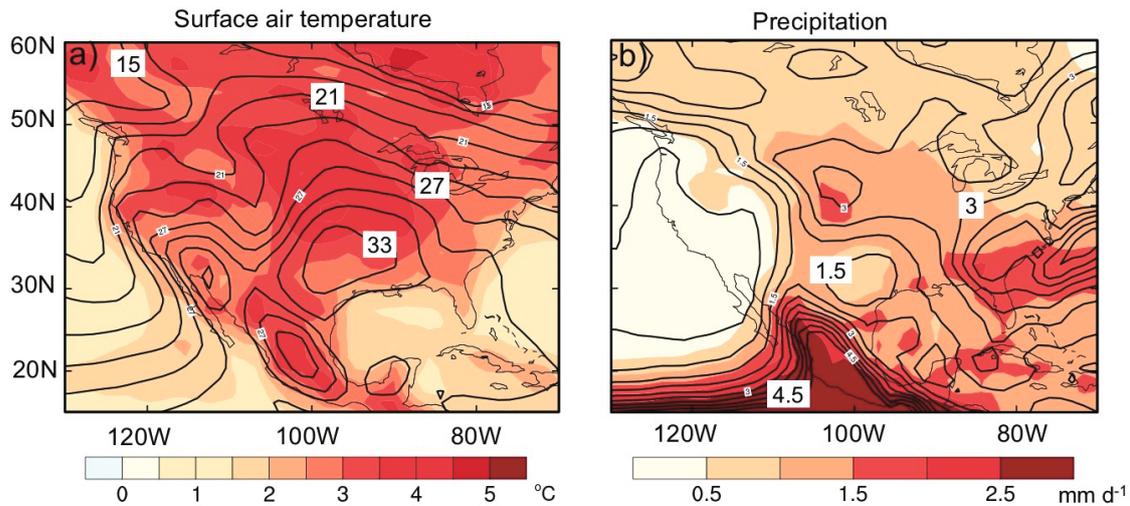


Fig. 7. Ensemble mean (contours) and ensemble spread (color-shaded) for CMIP3 “A2” scenario future projections (AD 2074–99) for (a) averaged summer (June–August) surface air temperature and (b) precipitation over the U.S.

Table 2. List of CMIP3 models used in this study (IPCC, 2007).

Model	Sponsor	Atmospheric resolution (lat × lon)
BCCR-BCM2.0	Bjerknes Centre for Climate Research, Norway	2.8° × 2.8°
CCCMA_CGCM3_1.5	Canadian Centre for Climate Modelling & Analysis	2.8° × 2.8°
CNRM-CM3.1	Météo-France/Centre National de Recherches Météorologiques	2.8° × 2.8°
CSIRO MK3.0.2	CSIRO Atmospheric Research	1.875° × 1.875°
CSIRO MK3.5.3	CSIRO Atmospheric Research	1.875° × 1.875°
GFDL CM2.0_0.3	US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory	2.5° × 2°
GFDL CM2.0_1.3	US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory	2.5° × 2°
GISS_MODEL_E_R.9	NASA/Goddard Institute for Space Studies	5° × 4°
INGV_ECHAM4.1	Instituto Nazionale di Geofisica e Vulcanologia	1.125° × 1.125°
INM CM3.0.1	Institute for Numerical Mathematics	5° × 4°
IPSL-CM4.1	L’Institut Pierre Simon Laplace	3.75° × 2.5°
MIROC3.2 (medres)	Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC)	2.8° × 2.8°
MPIECHAM5.4	Max Planck Institute for Meteorology	1.875° × 1.875°
MRI-CGCM2_3_2A.5	Meteorological Research Institute	2.8° × 2.8°
NCAR_CCSM3_0.8	National Center for Atmospheric Research	1.4° × 1.4°
NCAR_PCM1.4	National Center for Atmospheric Research	2.8° × 2.8°
UKMO_HADCM3.2	Hadley Centre for Climate Prediction and Research/Met Office	3.75° × 2.75°
UKMO_HADGEM1.2	Hadley Centre for Climate Prediction and Research/Met Office	3.75° × 2.75°

proaches in magnitude the amplitude of the climate change signal at horizontal scales below about 1000 km for surface temperature, and 2000 km for precipitation.

Our findings generalize beyond the specifics of the CMIP5 historical simulation ensemble. Parallel analyses of both (i) the earlier generation CMIP3 historical simula-

tion ensemble and (ii) 21st century climate projections based on the CMIP3 “A2” emissions scenario, yield qualitatively very similar conclusions. Given that downscaling methods (whether based on statistical or dynamical approaches) require information from large-scale climate model simulations as boundary conditions and/or reference states, the lack of

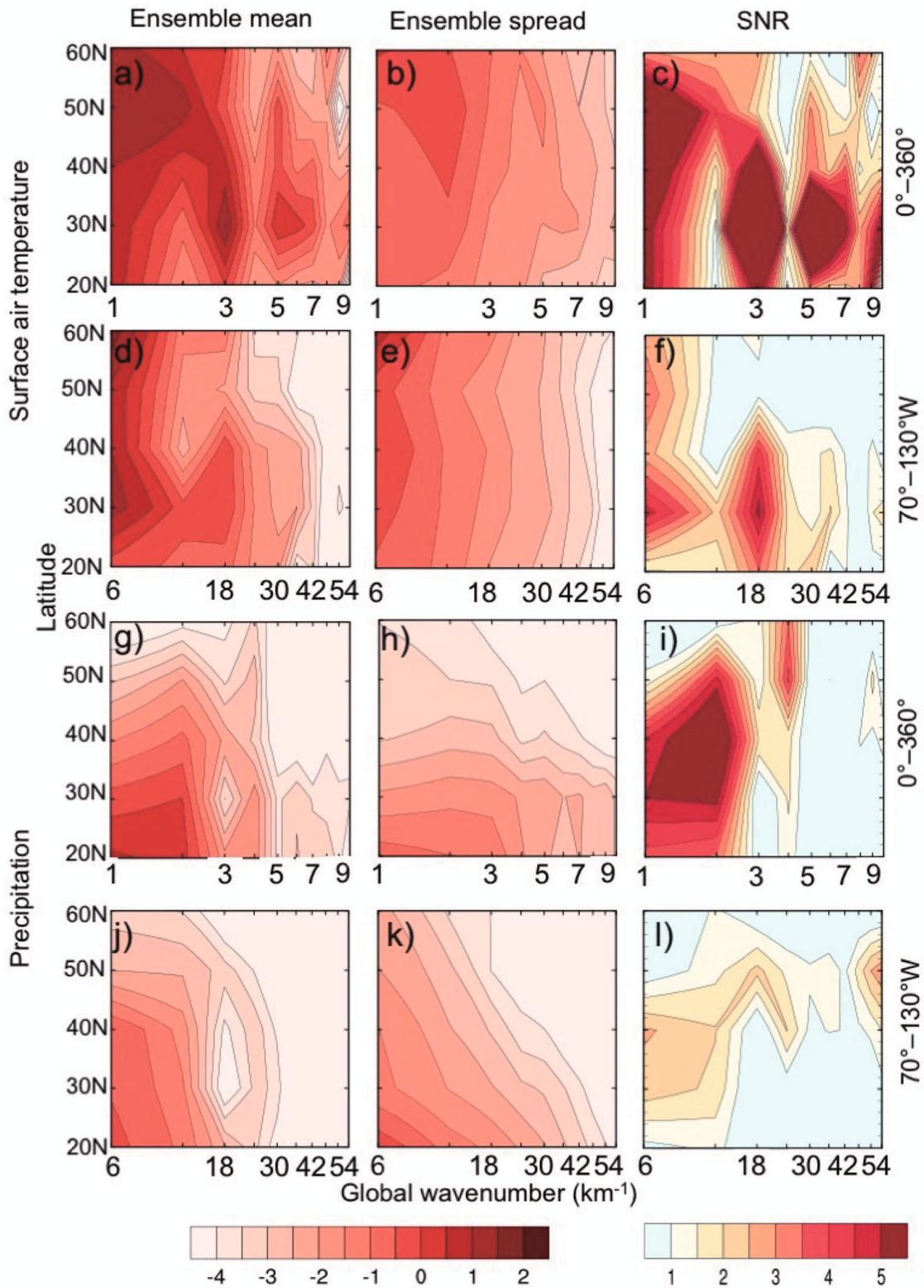


Fig. 8. As in Fig. 2 but using CMIP3 “A2” scenario future projections (AD 2074–99).

predictability at these larger scales likely translates to a lack of predictability at local scales. One apparent exception, based on our findings, are cases where smaller-scale orographic forcing or land–sea contrasts provide additional pre-

dictability at smaller scales.

Given the importance of future projections of surface temperature and precipitation for assessing climate change impacts such as heat stress, flooding potential and drought

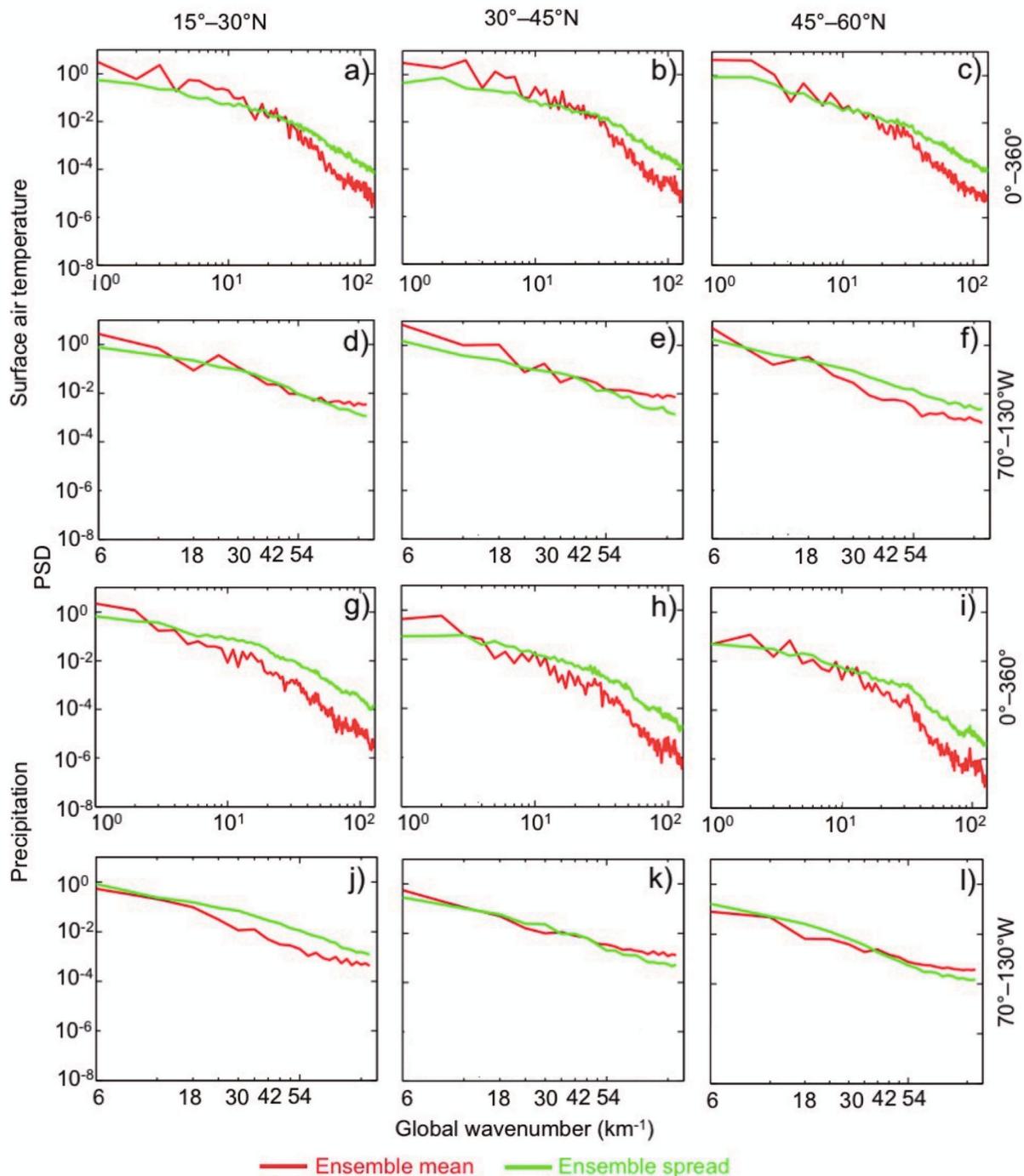


Fig. 9. As in Fig. 3 but using CMIP3 “A2” scenario future projections (AD 2074–2099).

magnitude, duration and extent, our findings suggest great challenges in assessing climate change risk and damage at regional scales most important to stakeholders and policymakers. One potential implication of our findings is that regional adaptation efforts might, in some circumstances, be better focused on reducing vulnerability to climate change in general, rather than planned adaptation to specific projected climate changes.

Acknowledgements. We thank the Hadley Centre and Climate Research Unit for the HadCRUT4 data and NOAA for the GPCC precipitation data. We thank CMIP and the climate modeling groups for the CMIP3/5 model output. The CMIP3/5 data used can be freely obtained from http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html. The comments from two anonymous reviewers were beneficial. This study was partially supported by the NSF (Grant No. AGS-1305798) and the ONR (Grant No. N000140910526).

Electronic supplementary material: Supplementary material is available in the online version of this article at <http://dx.doi.org/10.1007/s00376-016-6013-2>.

REFERENCES

- Bei, N., and F. Q. Zhang, 2007: Impacts of initial condition errors on mesoscale predictability of heavy precipitation along the Mei-Yu front of China. *Quart. J. Roy. Meteor. Soc.*, **133**, 83–99.
- Deser, C., R. Knutti, S. Solomon, and A. S. Phillips, 2012: Communication of the role of natural variability in future North American climate. *Nature Clim. Change*, **2**, 775–779.
- Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**, 1095–1107.
- IPCC, 2007: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
- Ji, F., Z. H. Wu, J. P. Huang, and E. P. Chassignet, 2014: Evolution of land surface air temperature trend. *Nature Clim. Change*, **4**(6), 462–466.
- Jones, C., F. Giorgi, and G. Assar, 2011: The coordinated regional downscaling experiment: CORDEX, an international downscaling link to CMIP5. *CLIVAR Exchanges*, **16**, 34–39.
- Li, W., C. E. Forest, and J. Barsugli, 2012: Comparing two methods to estimate the sensitivity of regional climate simulations to tropical SST anomalies. *J. Geophys. Res.*, **117**, D20103.
- Mearns, L. O., I. Bogardi, F. Giorgi, I. Matyasovszky, and M. Palecki, 1999: Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling. *J. Geophys. Res.*, **104**, 6603–6621.
- Meehl, G. A., and Coauthors, 2007: The WCRP CMIP3 multi-model dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772.
- Separovic, L., R. de Elía, and R. Laprise, 2008: Reproducible and irreproducible components in ensemble simulations with a Regional Climate Model. *Mon. Wea. Rev.* **136**, 4942–4961.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498.
- Tebaldi, C., R. L. Smith, D. Nychka, and L. O. Mearns, 2005: Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles. *J. Climate*, **18**, 1524–1540.
- Timm, O., and H. F. Diaz, 2009: Synoptic-statistical approach to regional downscaling of IPCC twenty-first-century climate projections: Seasonal rainfall over the Hawaiian Islands. *J. Climate*, **22**, 4261–4280.
- von Storch, H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Climate*, **6**, 1161–1171.
- Watterson, I. G., and P. H. Whetton, 2011: Distributions of decadal means of temperature and precipitation change under global warming. *J. Geophys. Res.*, **116**, D07101.
- Watterson, I. G., J. Bathols, and C. Heady, 2014: What influences the skill of climate models over the continents? *Bull. Amer. Meteor. Soc.*, **95**, 689–700.
- Whetton, P., I. Macadam, J. Bathols, and J. O’Grady, 2007: Assessment of the use of current climate patterns to evaluate regional enhanced greenhouse response patterns of climate models. *Geophys. Res. Lett.*, **34**, L14701.
- Zhou, T.-J., and R. C. Yu, 2006: Twentieth-century surface air temperature over china and the globe simulated by coupled climate models. *J. Climate*, **19**(22), 5843–5858.