Climatic Change (2019) 154:367–386
https://doi.org/10.1007/s10584-019-02411-y

Identifying credible and diverse GCMs for regional climate change studies—case study: Northeastern United States

Ambarish V. Karmalkar[1] · **Jeanne M. Thibeault**[2] · **Alexander M. Bryan**[3] · **Anji Seth**[2]

Received: 16 April 2018 / Accepted: 15 March 2019 / Published online: 13 April 2019
© Springer Nature B.V. 2019

Abstract

Climate data obtained from global climate models (GCMs) form the basis of most studies of regional climate change and its impacts. Using the northeastern U.S. as a test case, we develop a framework to systematically sub-select reliable models for use in climate change studies in the region. Model performance over the historical period is evaluated first for a wide variety of standard and process metrics including large-scale atmospheric circulation features that drive regional climate variability. The inclusion of process-based metrics allows identification of credible models in capturing key processes relevant for the climate of the northeastern U.S. Model performance is then used in conjunction with the assessment of redundancy in model projections, especially in summer precipitation, to eliminate models that have better performing counterparts. Finally, we retain some mixed-performing models to maintain the range of climate model uncertainty, required by the fact that model biases are not strongly related to their respective projections. This framework leads to the retention of 16 of 36 CMIP5 GCMs that (a) have a satisfactory historical performance for a variety of metrics and (b) provide diverse climate projections consistent with uncertainties in the multi-model ensemble (MME). Overall, the models show significant variations in their performance across metrics and seasons with none emerging as the best model in all metrics. The retained set reduces the number of models by more than one half, easing the computational burden of using the entire CMIP5 MME, while still maintaining a wide range of projections for risk assessment. The retention of some mixed-performing models to maintain ensemble uncertainty suggests a potential to narrow the ranges in temperature and precipitation. But any further refinement should be based on a more detailed analysis of models in capturing regional climate variability and extremes to avoid providing overconfident projections.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10584-019-02411-y) contains supplementary material, which is available to authorized users.

✉ Ambarish V. Karmalkar
akarmalkar@geo.umass.edu

Extended author information available on the last page of the article.

1 Introduction

Studies of regional climate change and its impacts often use climate data obtained from global climate model (GCM) simulations, either directly or after downscaling them to regional scales. While the climate data from a large number of GCMs is readily available, the users of that data across disciplines are typically interested in or even constrained to using a handful of climate models due to enormous computational and data storage costs. Selecting a limited number of appropriate climate models that perform well and that represent uncertainty in projections is one of the first challenges faced by the users (Barsugli et al. 2013; Snover et al. 2013; Mote et al. 2011).

Selecting a subset of GCMs is a challenging and contentious topic. It is challenging because GCM credibility varies considerably by metric, spatiotemporal scale, season, and region (Masson and Knutti 2011), and it is contentious because a selected subset can underrepresent the uncertainty in climate change projections (Weigel et al. 2010). The most common approach to sub-select GCMs is to assess models' ability to simulate historical climate and select a few better performing models that capture the full range of uncertainty in future projections (e.g., McSweeney et al. 2015; Monerie et al. 2017). The sub-selection procedure is partly subjective in nature and is intimately linked to choices regarding the metrics and error measures evaluated, the spatial and temporal scales examined, and the level of performance expected. Therefore, eliminating poor-performing models is considered more reasonable than efforts to identify the best models in an ensemble (Knutti et al. 2010; McSweeney et al. 2015). A simple assessment of model performance for a region involves using a set of standard metrics such as biases, mean squared errors (MSEs), and standard deviations (Taylor 2001; Gleckler et al. 2008) to examine how simulated fields deviate from observations. The standard metrics are useful to summarize the overall model performance, but are insufficient to diagnose model errors. For a more complete evaluation, a process-based assessment of model performance is required to identify models that correctly simulate large-scale circulation features that drive changes in key surface variables such as temperature and precipitation. This approach confirms that a good model performance is a result of a physically consistent behavior across variables, and not due to error compensation in models on climate timescales, and thus allows identification of credible models. The process-based evaluation has been used to select GCMs for dynamical downscaling studies over the Arctic (Overland et al. 2011) and Southeast Asia, Europe and Africa (McSweeney et al. 2012; McSweeney et al. 2015).

Combining standard- and process-metric-based evaluation addresses the issue of model credibility; still remaining, however, is how to resolve the contentious issue of appropriately representing the uncertainty in regional projections whilst eliminating models. While the performance-based evaluation helps eliminate poor-performing models, it does not necessarily reduce the uncertainty in projections since model shortcomings are often only weakly related to their projections (Pierce et al. 2009; Knutti et al. 2010; McSweeney et al. 2015). For example, Thibeault and Seth (2015) identified six CMIP5 models from a set of 15 as having a credible representation of the processes responsible for summer precipitation in the northeastern U.S., but these models still had a substantial disagreement in their projections. Similarly, Rowell et al. (2016) found that uncertainty in projections for two regions in Africa cannot be reduced by selecting models with skillful performance over the historical period. A simple magnitude-based classification of projections is commonly used (e.g., (Whetton et al. 2012) to select a small set of models that spans the full range of uncertainty

seen in the MME. But a more sophisticated, quantitative approach would be to eliminate models that are redundant in terms of their projections with other, better performing models, leaving us with the minimum possible subset that spans a wide range of uncertainty. This approach exploits the fact that models in the MME are not fully independent (Sanderson et al. 2015) due to the exchange of concepts, codes, and model components between developers across modeling centers (Knutti et al. 2013). This redundancy can be used to retain a subset of models with diverse climate projections by eliminating poor-performing models that have projections similar to the better performing members of the ensemble.

The assessment of model performance and projections can be impacted by internal climate variability, generated by the chaotic and unpredictable behavior of the coupled atmosphere-ocean system. This internal variability (IV) is one of the dominant sources of uncertainty in climate predictions and projections (Hawkins and Sutton 2009; Wallace et al. 2016) and can often mask the regional climate change signal (a long-term trend) at interannual as well as decadal time scales (Deser et al. 2012). While climate models successfully capture many aspects of IV, simulated variability may not always match the real-world variability. Such a mismatch over the analysis periods can affect the assessment of model performance (e.g., Rupp et al. 2013) as well as projections. The CMIP5 MME is not designed to explore the impact of IV systematically, but some models provide multiple members driven by different initial conditions that can be used to assess how simulated IV affects the model sub-selection framework.

In this paper, we develop a framework to identify a subset of GCMs that are *credible* and *diverse*, as described earlier. The framework is informed by the evaluation of model performance for standard (area-averaged) as well as process metrics, the redundancy in projections across the CMIP5 ensemble, and the assessment of spread in model performance and projections due to internal variability. Note that the attempt here is not to *pick* a small set of models that can be directly used in *all* regional climate change and impacts studies. Instead, our focus is to reduce the size of the MME based on a diverse set of climate metrics, thereby facilitating selection of a handful of models to address downstream regional projections and impacts work. Consequently, further work will be necessary to pick the most relevant models for specific research questions under investigation (e.g., Vano et al. 2015).

As a case study, we apply our framework to the northeastern U.S. (NEUS). This region is one of the fastest warming regions in the contiguous U.S. (Wuebbles D et al. 2017; Karmalkar and Bradley 2017) and has experienced a dramatic rise in extreme precipitation in recent decades alongside a general tendency towards wet conditions since the 1970s (Horton et al. 2014) and in future scenarios (Maloney et al. 2014; Karmalkar and Bradley 2017). The region is home to over 20% of the U.S. population and many vulnerable species and ecosystems (Staudinger et al. 2015). As a result, many states, cities, and conservation agencies in the region have been increasingly relying on climate information to support local conservation and adaptation efforts that can all benefit from a systematic framework for selecting climate model data.

2 Data and metrics

2.1 Climate model and reference datasets

The analysis is based on climate model simulations that contributed to the fifth phase of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012). The CMIP5 MME used here consists of monthly outputs from historical and future scenario simulations from

36 models. The simulations over the historical period are compared to gridded observations: Climate Research Unit Time Series (CRU TS version 3.23) data for temperature (TAS; Harris et al. 2014) and the University of Delaware (UDEL) data for precipitation (PR; Matsuura and Willmott 2012). We also use two reanalysis datasets–ERA Interim (ERAI; Dee et al. 2011) and NCEP1 reanalysis (Kalnay et al. 1996)–to evaluate the models' ability to capture process metrics (described below). All model and reference data were first regridded to a common 2.5° × 2.5° latitude-longitude grid using bilinear interpolation to allow for comparisons to be made at the same resolution. Future projections are calculated for the period 2040–2069 for the high emissions scenario, RCP8.5 (Meinshausen et al. 2011) relative to the 1979–2000 baseline. For the mid-century projections considered here, it is sufficient to use a single RCP since the uncertainty in NEUS temperature and precipitation projections is dominated by modeling uncertainty at these timescales (Karmalkar and Bradley 2017). To understand the impact of simulated internal climate variability on model performance and projections, we analyze multi-member initial conditions (IC) ensembles available for 22 and 11 CMIP5 models over the historical and RCP8.5 analysis periods respectively, in addition to using the first realization (r1i1p1) of all 36 models. This brings the total number of model realizations analyzed to 113 for model performance and 74 for climate change projections.

2.2 Evaluation metrics

We use a combination of standard and process metrics to ensure that the models we select adequately capture the important physical processes of the region. We evaluate model performance for four seasons: winter (DJF), spring (MAM), summer (JJA), fall (SON), using the following standard metrics (hereafter, "sm") for both temperature and precipitation:

1. area-averaged seasonal mean squared errors (MSE) calculated using seasonal mean biases (8 metrics).
2. the climatological annual cycles constructed from monthly mean values (2 metrics), and
3. standard deviations (SD) of seasonal means to assess models' ability to capture interannual variability (8 metrics).

These provide two sets of metrics denoted by "sm-TAS" (set 1) and "sm-PR" (set 2). We anticipate that the assessment of the precipitation annual cycle can better enable model discrimination since a subset of CMIP5 models has been shown to be less skillful at capturing the annual cycle of the NEUS precipitation than that of temperature (Lynch et al. 2016). The SD metric used here is the absolute value of the logarithm of the model to observed SD ratio (following Rowell et al. 2016).

For process metrics (hereafter, "pm"), we focus on large-scale climate features associated with summer and winter precipitation in the NEUS. These metrics are chosen following previous work by Thibeault and Seth (2014, 2015) [hereafter TS14-15] for summer precipitation, and by Ning and Bradley (2015) and Bradbury et al. (2003) for winter precipitation. The metrics include:

1. JJA mean precipitation, mean sea level pressure (MSLP), geopotential heights at 500 hPa (Z500), and zonal and meridional winds at 850 hPa (U-V850) (4 metrics), and
2. DJF mean precipitation, MSLP, Z500 (3 metrics).

These two sets of metrics are denoted by "pm-JJA" (set 3) and "pm-DJF" (set 4). For model analysis of process metrics, we use ERAI data for summer, consistent with TS14-15, and NCEP1 for winter variables. TS14-15 analyzed pm-JJA metrics for five and 15 CMIP5

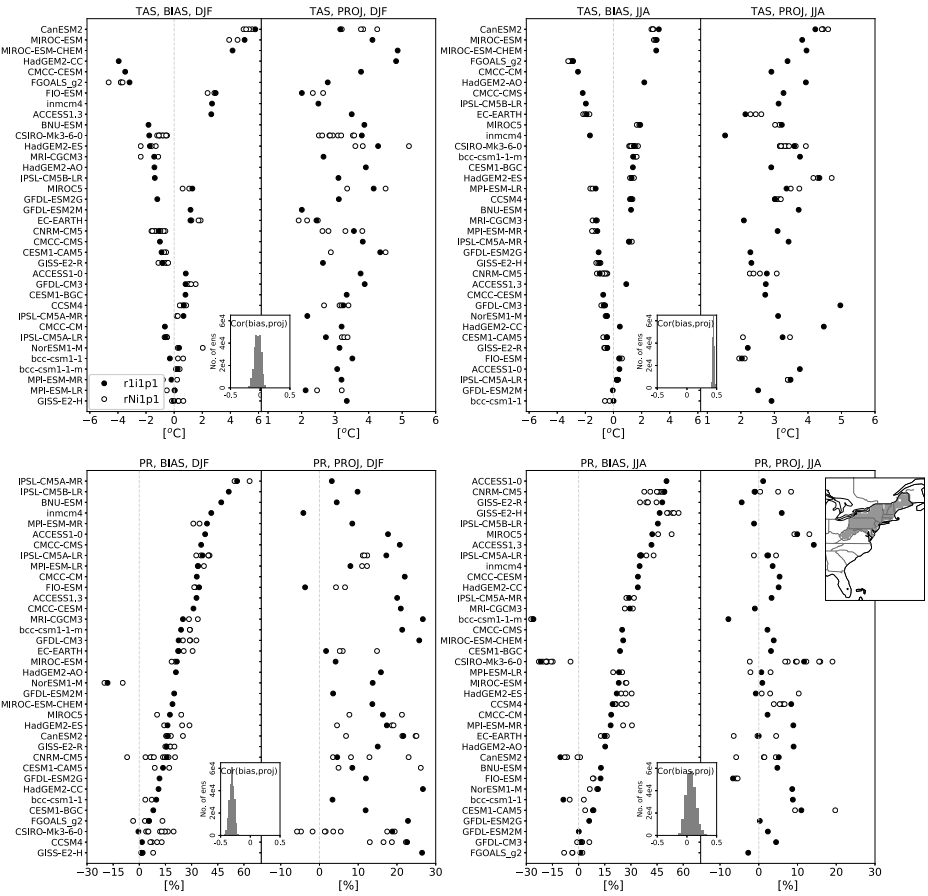

Fig. 1 Winter (DJF) and summer (JJA) temperature biases (BIAS) and projections (PROJ) for 36 CMIP5 GCMs. The black dots show values for the first realization (r1i1p1) and the open circles show values for multiple initial conditions ensemble members whenever available. The models are arranged according to absolute biases in each field. Histograms show correlations between model biases and projections, across 36 models, for all possible combinations (total of 248832) across multiple IC realizations of 11 models over historical and RCP periods. The inset figure shows the NEUS domain used to calculate area-averaged biases and projections

GCMs respectively, and we extend their analysis to 36 models. For pm-JJA, we use 1979–2000 for observations and 1971–1999 for models to evaluate the relationships between precipitation anomalies and atmospheric variables. Because interannual variations in the models do not coincide with observations during any given year, a longer time period for analysis of the models permits more robust statistics. The winter metrics are analyzed using NCEP1 reanalysis data, which allows better characterization of precipitation composites and modes of climate variability using a longer period (1951–2005). The MSLP and Z500 data are used to calculate models' abilities to capture the North Atlantic Oscillation (NAO) and the Pacific-North America (PNA) pattern, both large-scale modes of climate variability important for winter climate in the NEUS. Here, we include well-understood metrics available for DJF and JJA and note that further work will be necessary to include process metrics important in spring and fall seasons (e.g., extratropical storms).

For spatial averages used for sm-TAS and sm-PR, we follow the region definition used in the Third National Climate Assessment (see Fig. 1; Horton et al. 2014) for the NEUS, which includes New York, New Jersey, Pennsylvania, Delaware, Maryland, West Virginia, and the six New England states (see inset in Fig 1). The process metrics (pm-JJA, pm-DJF) use the following domains: 85° W–40° E, 20° N–80° N for NAO, 180° W–65° W, 10° N–80° N for PNA, 83° W–67° W, 36° N–49° N for high minus low precipitation composites. The relevance of using summer and winter process metrics for the NEUS is described in more detail in Section 3.2.

2.3 Model ranks

Models are first ranked for each of the 25 metrics and then individual rankings are added to calculate the total ranks and ranks for each of the four sets described above. To avoid model rankings from being too sensitive to any one performance measure, we have attempted to choose metrics in sets 1 and 2 (sm-TAS, sm-PR) such that they provide independent information on model performance (see Fig. S1). For instance, we include MSEs and not biases in our set of metrics since they capture the same information. We include MSEs for all four seasons despite strong correlations between seasonal errors across the MME because the strong positive relationships arise mostly due to the worst performing models in the ensemble. We emphasize that accounting for redundancy in metrics is less important in our framework since model discrimination does not rely on using a single performance index (e.g., skill score) constructed from all available metrics. Instead, we only eliminate models that have relatively poor performances for all four sets of metrics. For the model to be deemed credible in the case of process metrics (sets 3 and 4), models must perform well (i.e., rank higher) for all the metrics within the set to confirm that they present physically consistent behavior across admittedly dependent metrics. We calculate model rankings based on the first realization of every GCM (r1i1p1) as well as using all 113 realizations for the standard metrics. How model ranks are used to discriminate among models is described in detail in Section 4.

3 Model performance and projections

3.1 Area-averaged metrics

Winter and summer mean biases and projected changes in temperature and precipitation for the NEUS for 36 CMIP5 models and their IC members are shown in Fig. 1. Note that we show biases here which provide more intuitive insight into model performance instead of MSEs which are used in ranking models. Model performance varies considerably across the ensemble with warm and cold as well as dry and wet biases in both seasons, with some of the worst performers in the ensemble having temperature biases that are a factor of four larger and precipitation biases over 20 times larger than the good performers. This indicates a large inter-model range in performances making it feasible to discriminate between models. With the exception of fall (see Fig. S2), the CMIP5 ensemble overestimates precipitation in the NEUS. Most models indicate an increase in winter and spring precipitation in the future (Figs. 1 and S2) with the inter-model range of about 5% decrease to over 30% increase. Projections of summer and fall precipitation, on the other hand, include increases as well as decreases. Overall, the uncertainty ranges in seasonal temperature and precipitation projections of 1.5 °C to 4.5 °C and −10% to +35% respectively (ranges for individual

seasons shown in Figs. 1 and S2) in CMIP5 represent our current understanding of the anticipated changes in the NEUS.

Multiple realizations of a given model that only differ in their initial conditions show a spread in TAS and PR biases and projections (open circles in Fig. 1) highlighting the importance of accounting for internal variability. The temporal averages of over 20 years should ideally minimize the impact of IV on biases and projections but the small domain size used is likely responsible for significant uncertainty in mid-century projections (Deser et al. 2014). Overall, the spread due to IV is higher for precipitation than temperature and for projections than biases. But more importantly, the intra-model ranges due to IV are smaller than the inter-model range in biases, a useful finding for model discrimination. The uncertainty in projections, on the other hand, cannot be simply ascribed to model diversity since the spread due to internal variability in IC ensembles of individual models is comparable to the inter-model spread.

We first calculate correlations between seasonal biases and projections to determine if a relationship between model performance and projections exists for the NEUS. To also include the impact of internal variability, we calculate correlations, across 36 models, for all possible combinations of model biases and projections across multiple IC realizations, presented in the form of histograms in Fig. 1. These correlations demonstrate that there are no strong relationships between model biases and projections, which suggests that the spread in projections cannot be easily constrained by simply eliminating models that have large biases in either field. Even when the correlations are moderately high, as in the case of summer temperatures, discarding models with large biases in these key variables does not necessarily reduce the spread in projections. Also, a lack of any systematic relationship between seasonal temperature and precipitation biases (not shown) suggests that models exist that are successful in capturing seasonal temperatures in the NEUS but have large biases in how well they simulate precipitation climatology and vice versa.

To help explain the spread in area-averaged precipitation projections, we examine the spatial patterns of simulated precipitation over the NEUS (Fig. 2). The region receives

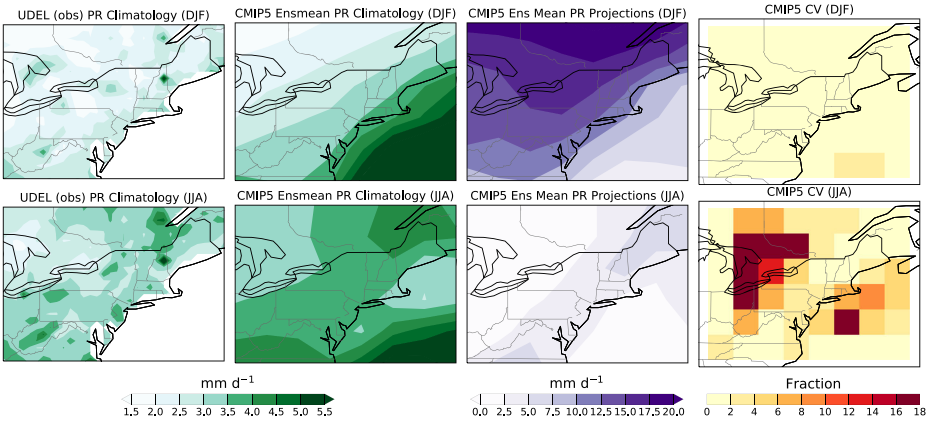

Fig. 2 Seasonal precipitation in observations and CMIP5 ensemble. Observed climatology (column 1), CMIP5 ensemble mean historical climatology (column 2), CMIP5 ensemble mean projections (column 3), and the coefficient of variation (CV; column 4) across the ensemble for winter (DJF; top row) and summer (JJA; bottom row). Projections are for RCP8.5 and for the period 2040–2069 relative to 1979–2000

abundant rainfall throughout the year as a result of a southwesterly flow of moisture from the Gulf of Mexico in the warm season and extratropical storms in the cold season. The comparison between the observed precipitation climatology and the CMIP5 ensemble mean (Fig. 2) suggests that most models overestimate DJF and JJA precipitation in the region, also reflected in the spatially averaged biases shown in Fig. 1. The projected increase in total annual precipitation is robust across models (Seager et al. 2010), mainly due to increases in winter and spring precipitation. Winter precipitation is projected to increase along the entire Eastern Seaboard (Maloney et al. 2014) with higher increases in the NEUS. The ensemble mean pattern and the coefficient of variation (CV; Fig. 2) across the MME together indicate that the inter-model spread in winter precipitation projections is spatially uniform over the NEUS, confirming that the area-averaged metrics are representative of the regional-scale changes in winter precipitation. In contrast, summer precipitation projections indicate a very small increase when averaged across MME members, but the spatial patterns vary across models, as indicated by large CVs across the domain. This highlights that the models exhibit the most diversity in spatial patterns for summer precipitation projections in addition to a wide range in the magnitude in area-averaged projections (Fig. 1), facilitating model discrimination.

3.2 Process-based metrics

We use atmospheric processes identified in previous studies in the evaluation of model performance in addition to using surface variables discussed in Section 3.1. The summer process metrics (pm-JJA) are based on the analysis by TS14-15, who demonstrate that the wet summers in the NEUS are associated with (a) negative Z500 anomalies over the Great Lakes region and positive Z500 anomalies over the central North Atlantic, (b) an increase in MSLP over the North Atlantic potentially related to the westward expansion of the North Atlantic Subtropical High (NASH), and (c) the resultant southwesterly moisture transport and convergence in the NEUS region indicated by low level winds (U-V850).

Following TS14-15, we regress the area-averaged and detrended NEUS summer precipitation anomalies in both observations (ERAI) and models onto the observed and simulated detrended fields of Z500, U and V winds at 500 and 850 hPa, and MSLP (Fig. 3a, U-V850 not shown). These regression maps show changes in the circulation fields for every standard deviation change in NEUS precipitation. The model performance is then quantified by calculating pattern correlations between the observed and simulated regression maps. Naturally, models that capture the NEUS summer precipitation drivers well show strong correlations with the observed regression patterns. Figure 3a shows observed and simulated summer mean precipitation and regression maps for Z500, U500 and V500, and MSLP for observations and for the best and worst performing models for these metrics. The models are ranked for individual metrics within the set first, and then the individual ranks are added to calculate an overall rank for a given set. Models with high ranks for pm-JJA, therefore, suggest that they capture precipitation and the underlying drivers reasonably well. While many models simulate the PR climatology well, the process metrics show significant departures, in their position and magnitude, from observations (Figs. 3a, S3, and S4), enabling model discrimination.

Winter temperature and precipitation variability in the NEUS is influenced by the large-scale modes of climate variability namely the Pacific-North American (PNA) pattern, the North Atlantic Oscillation (NAO), and the ENSO (Ning and Bradley 2015; Bradbury et al. 2003). In particular, when El Niño years coincide with the high phase of PNA, the NEUS typically experiences wet winters characterized by a decreasing coast-to-inland gradient in

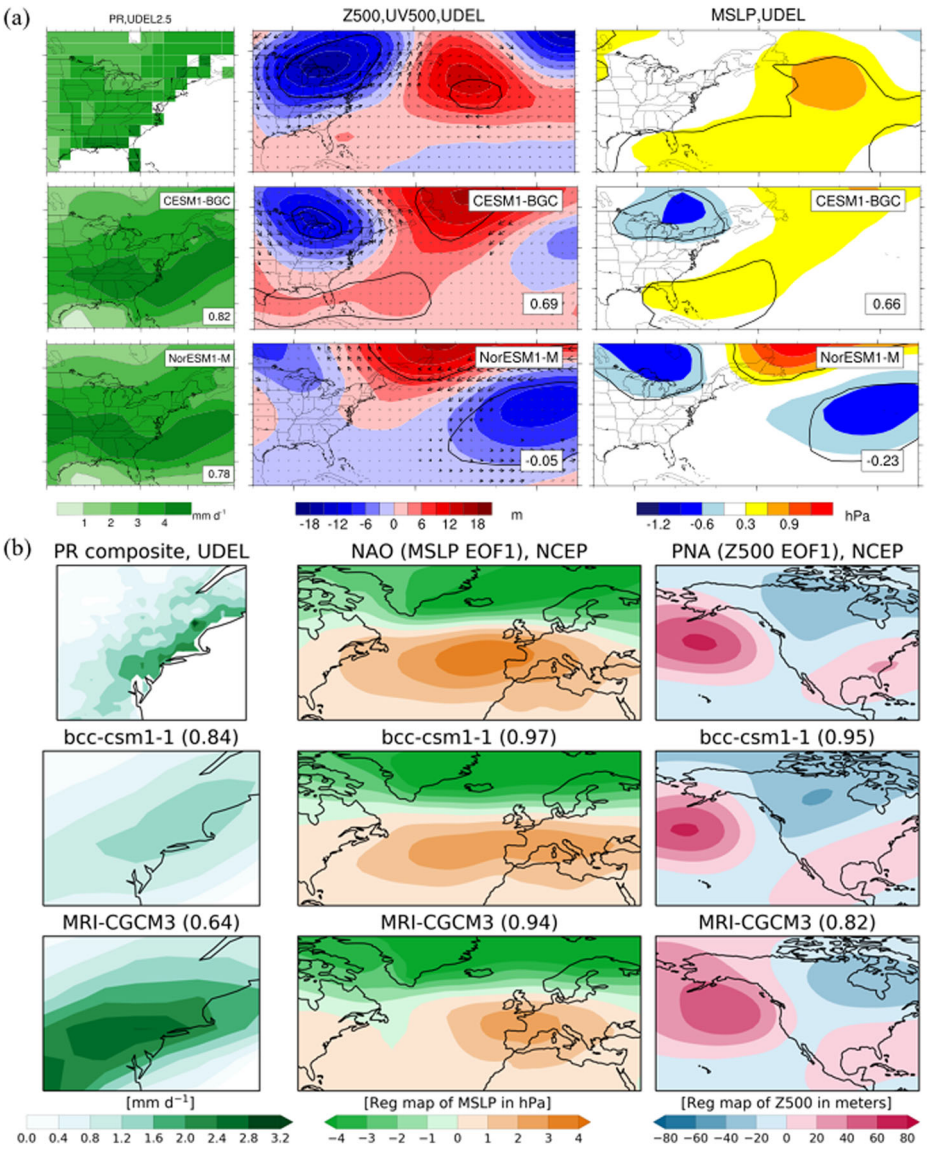

Fig. 3 Observed and simulated process metrics. **a** JJA mean precipitation (column 1) and regression maps between PR and Z500, winds at 500 hPa (column 2), and MSLP (column 3). **b** precipitation composite (column 1) showing the difference in DJF mean precipitation between wet and dry winters and the leading EOFs of MSLP (column 2) and Z500 (column 3) that capture NAO and PNA patterns respectively. The leading EOFs are shown as regression maps. In both (**a**) and (**b**), the top row is for reference data and rows 2 and 3 show the best performing and worst performing models respectively for both summer and winter process metrics. Results for individual models are shown in Figs. S3–S7

precipitation amounts. The gradient results from extreme precipitation events in the form of coastal storms (e.g., Nor'easters; Sheffield et al. 2013). Most CMIP5 models analyzed here cannot simulate storms accurately due to their coarse horizontal resolution, nor do

they capture the phases of different modes of variability exactly. Given the difficulties in examining teleconnections for a large number of models, we limit our analysis to assess the models' ability to capture continental-scale spatial patterns of PNA and NAO in winter (indicated by pm-DJF).

The NAO and PNA modes are defined as the leading modes (EOFs) of MSLP and Z500 variability respectively. We assess the models' ability to capture these modes by comparing the leading EOFs in the models with NCEP1 reanalysis data over a 55-year period (1951–2005) over the domains shown in Fig. 3b. Additionally, spatial variations in winter

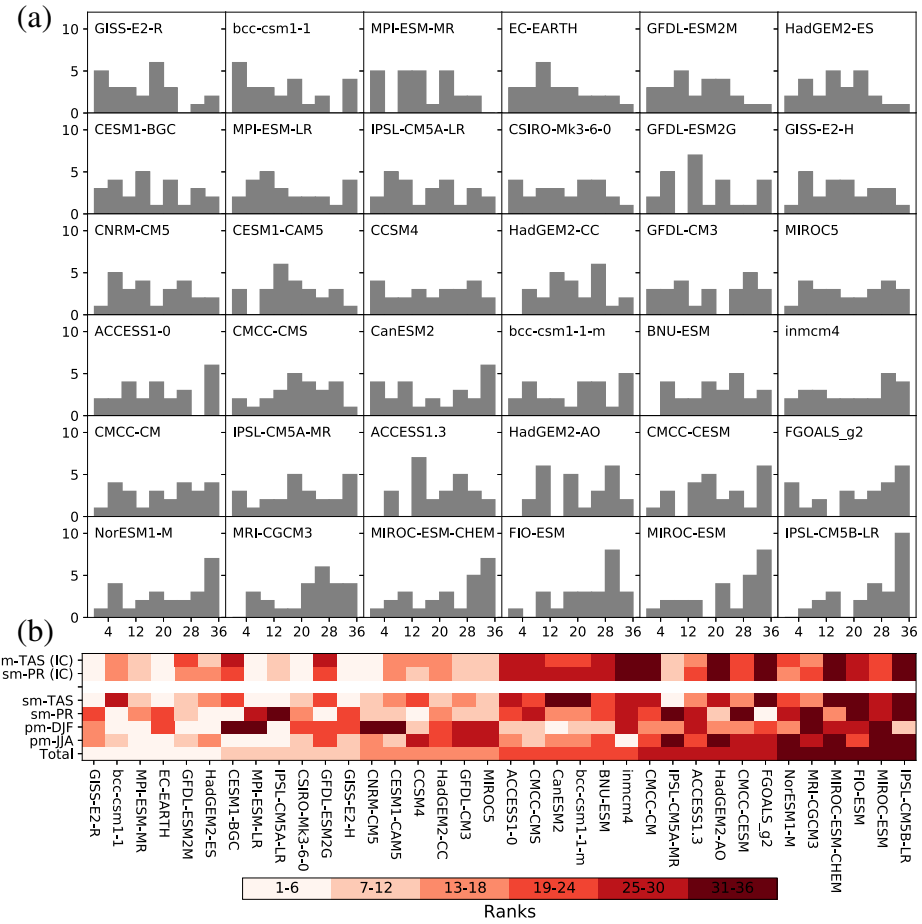

Fig. 4 (a) Histograms showing the distribution of ranks for each model's performance for metrics listed in Section 2. Models are arranged according to the rank totals across all metrics: an overall better performer to the top left, and the poor performer to the bottom right. Models with better (worse) performance for most metrics would have right-skewed (left-skewed) distributions. **b** Model ranks are shown for each of the four sets of metrics (pm-DJF, pm-JJA, sm-TAS, sm-PR). Ranks for standard metrics based on the IC ensemble (sm-TAS (IC) and sm-PR (IC)) are shown for comparison. Models are ordered, shown by "Total," based on their total rank across four sets of metrics, excluding IC. Model ranking for individual metrics in each set are shown in Fig. S8

precipitation in the NEUS are assessed by comparing the observed and simulated differences between composites of high and low winter precipitation years; defined as the years with winter precipitation higher or lower than one standard deviation from the 55-year mean. Correlations between the observed and simulated spatial patterns of high minus low composites quantifies the models' skill in capturing the east-to-west gradient in winter precipitation. These comparisons are shown in Fig. 3b for observations and for the best and worst performing models for metrics in the pm-DJF set. All models capture the NAO and PNA patterns well as evident from high spatial correlations between the observed and simulated patterns (NAO: 0.84–0.98, PNA: 0.63–0.98; see Figs. S5, S6, S7 for individual model results). The models' ability to capture all three winter metrics determines their overall rank for pm-DJF.

Large-scale drivers of climate variability included in pm-JJA and pm-DJF are likely to be present under future scenarios, albeit with changes in their position and magnitude. Consequently, the models that completely fail to capture these drivers of the NE climate or do not have a satisfactory representation of these features can be deemed unsuitable to provide reliable projections.

4 Model elimination and uncertainty in regional projections

We use the analysis described in Section 3 to rank CMIP5 models based on their performances. They are shown in Fig. 4a in the form of histograms of ranks for the metrics described in Section 2. The bars in each panel signify the number of times, across all 25 metrics, a model is ranked 1 to 36. A model that is ranked high for all the metrics under consideration (a very good model, for instance) would have a right-skewed histogram whereas the worst model would have a left-skewed histogram. Figure 4a clearly demonstrates that the performance is mixed and that it is difficult to identify best or worst models when a large number of metrics are included in the model assessment. Figure 4b shows the models' total ranks based on all the metrics as well as for individual sets. Figure 4 suggests a clear potential to discriminate among models since some models emerge as overall better performers (e.g., bcc-csm1-1, GISS-E2-R, MPI-ESM-MR) while some others are poor performers (e.g., MIROC models, IPSL-CM5B-LR) even though it is difficult to select best models overall. To account for the influence of internal variability on model rankings, we re-calculate model ranks for the standard metrics (Sets 1, 2) by picking the best performing members for each of the 22 models from their respective IC ensembles. These rankings for the standard metrics are shown in Fig. 4b (sm-TAS (IC), sm-PR (IC)) and are used to avoid eliminating models simply based on the performance of one realization (r1i1p1) when the better performing members may be available. CSIRO-Mk3-6-0 is one example where the additional nine IC members all have smaller temperature biases in winter than the first member (see Fig. 1).

Given the diversity in performance shown in Fig. 4, we divide all the CMIP5 models into three categories to help the process of model elimination. Models that rank in the top half (ranked 1–18) for all four sets of performance metrics are categorized as "good" performers. Similarly, models ranked in the bottom half (ranked 19–36) for all four sets of metrics are considered "poor" performers. All the remaining models are considered to have "mixed" performance, which includes models that may perform well for some metrics and poorly for others (two peaks in histograms, e.g., CanESM2). This categorization, shown in Table 1, identifies six good performers (GISS-E2-R, bcc-csm1-1, MPI-ESM-MR, GFDL-ESM2M,

HadGEM2-ES, IPSL-CM5A-LR) that we must retain, and two poor performers (FIO-ESM, MIROC-ESM) that can be eliminated right away. Note that the six "good" performers are not exactly the top six models based on total ranks but are deemed "good" because they rank in the top half for all four sets of metrics.

The remaining 28 models, in the "mixed" set, still have a wide range of performances, which provides an opportunity to eliminate relatively poor-performing models that provide redundant information about projections. First, the "mixed" performers are divided into two categories—essential and optional—based on future projections to help guard against constraining projections without proper justification. Specifically, models with extreme seasonal temperature and precipitation projections (top and bottom 10 percentiles) for at least two of the four seasons are considered "essential" to assess if there are relatively better performing models in the mixed set that, when retained, would maintain the range in seasonal projections. Second, the elimination approach exploits the redundancy in the spatial patterns of summer precipitation projections over the NEUS using hierarchical clustering (Rousseeuw 1987). The winter precipitation projections are uninteresting in this regard since models show very similar spatial patterns (Maloney et al. 2014) despite uncertainty in the magnitude of projected changes (Figs. 1 and 2).

To evaluate the redundancy in projections, correlation coefficients (r) between the spatial patterns of summer PR projections of different models are converted into a distance measure, which implies that models with similar patterns, and therefore strong correlations (positive or negative), will be closer (small "d") to each other. A pictorial representation of model clusters is shown using the *dendrogram* (Fig. 5) and the spatial patterns of individual model projections are shown in Fig. S9. The choice of the number of clusters (7) is based on visual inspection of model projections and is chosen such that the cluster means represent the differences in projections seen across the full ensemble. Note, however, that there can be intra-cluster differences in projection patterns that can lead to further groupings within the cluster. In general, models in clusters 3 and 4 (Fig. S9) indicate overall increases in precipitation but with different spatial patterns whereas models in the other 5 clusters show a variety of patterns. Cluster 6 includes models that indicate drying in the eastern and wetting in the western part of the domain, whereas models in cluster 2 show the opposite pattern. When this analysis is combined with performance-based categories, we find that six overall "good" performers are scattered across six of the seven clusters (Table 1) indicating that credible models are diverse in terms of their summer precipitation projections.

The use of intra-cluster similarities, performance-based rankings, and projections-based categories discussed above allow further model culling (Table 1). We first retain five models (EC-EARTH, CESM1-CAM5, CanESM2, bcc-csm1-1-m, inmcm4) that are "essential" and ranked high for one of the process metrics despite overall mixed performance. We eliminate GFDL-ESM2G, NorESM1-M, and ACCESS1-0 because they have mixed performance and because their spatial patterns of precipitation projections are very similar to GFDL-ESM2M, CESM1-CAM5, and HadGEM2-ES, respectively—three models that are retained. CESM1-BGC from cluster 2 is retained because it is the best model for pm-JJA, which allows us to eliminate two other optional and mixed-performing models (IPSL-CM5A-MR, BNU-ESM) in the same cluster. From clusters 3 and 6, we retain two optional models (CSIRO-Mk3-6-0 and MPI-ESM-LR) that perform well for all metrics except pm-DJF and eliminate other models with mixed performance. Note that two models from cluster 6 (CMCC-CMS, MRI-CGCM3) are eliminated despite being "essential" because of their overall inferior performance compared to the other retained models in the cluster. From cluster 5, we retain two relatively better performing but optional models (GISS-E2-H, CNRM-CM5), and eliminate the other two, one of which is an essential model (HadGEM2-CC). This approach leads to

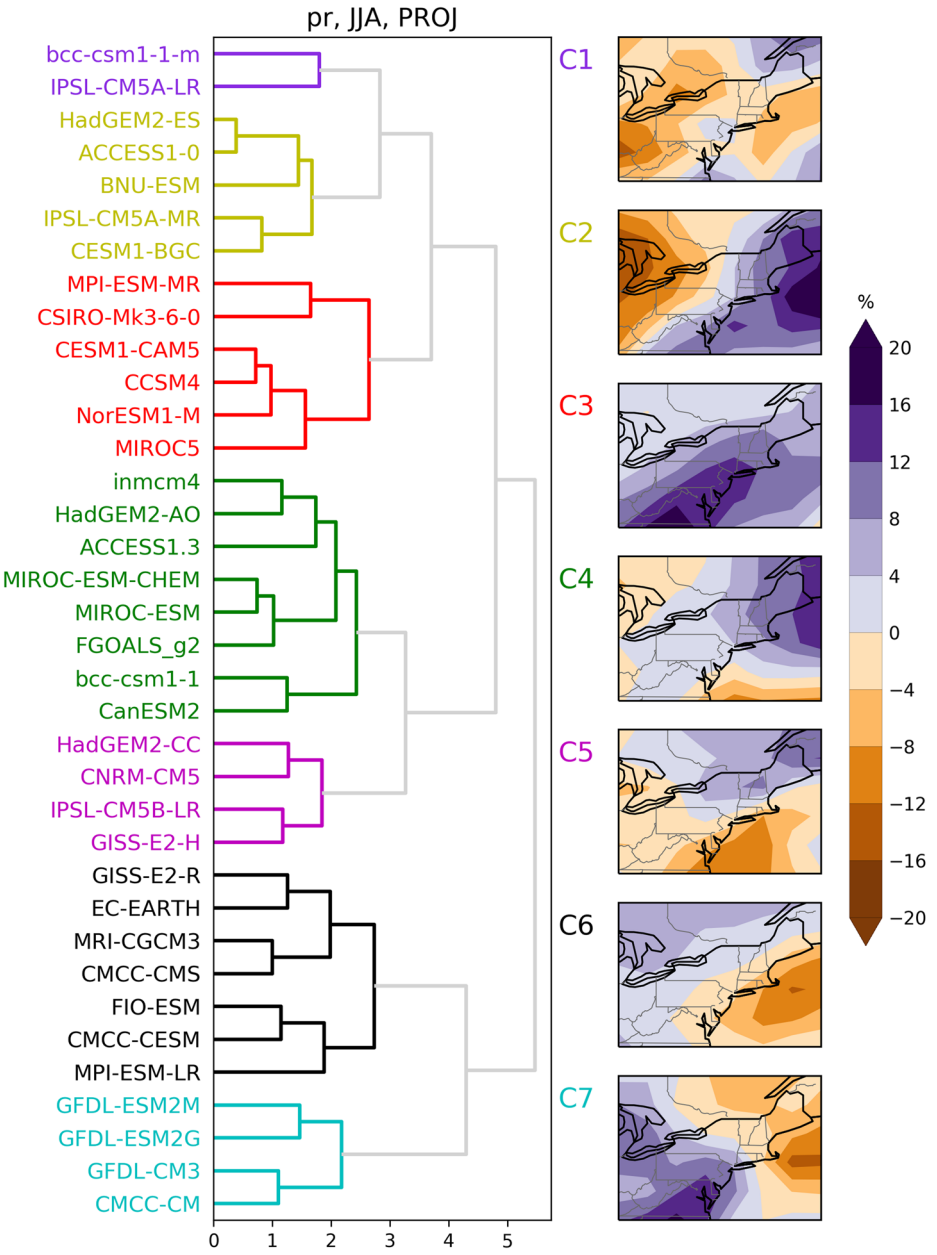

Fig. 5 Hierarchical clustering of CMIP5 models based on their summer precipitation projections (left) and spatial patterns of projections for seven clusters (C1–C7) based on the means across the members of every cluster (right)

the elimination of 20 of 36 models (Table 1). In summary, the retained models are credible since they have an overall good performance across a large set of standard and process metrics. They are also diverse since they span a wide range of projections and include models

Table 1 Table describing the decision framework

Model	Members (HIST, RCP)	sm-TAS	sm-PR	pm-JJA	pm-DJF	Total	Category	Decision
Cluster 1								
IPSL-CM5A-LR	6, 4	4(7)	35(7)	6	4	9	G	y
bcc-csm1-1-m	3, 1	32(19)	18(29)	14	8	22	M (E)	y
Cluster 2								
HadGEM2-ES	4, 4	14(11)	6(15)	16	5	6	G	y
CESM1-BGC	1, 1	20(25)	4(22)	1	32	7	M (O)	y
ACCESS1-0	1, 1	28(28)	20(30)	17	9	19	M (O)	
BNU-ESM	1, 1	22(29)	24(28)	19	7	23	M (O)	
IPSL-CM5A-MR	3, 1	6(10)	33(11)	32	16	26	M (O)	
Cluster 3								
MPI-ESM-MR	3, 1	7(9)	13(10)	3	2	3	G	y
CSIRO-Mk3-6-0	10,10	8(3)	15(6)	9	23	10	M (O)	y
CESM1-CAM5	3, 3	18(13)	9(13)	8	34	14	M (E)	y
CCSM4	5, 5	23(14)	5(12)	25	11	15	M (O)	
MIROC5	3, 3	12(12)	8(9)	29	28	18	M (E)	
NorESM1-M	3, 1	19(21)	25(14)	36	13	31	M (O)	
Cluster 4								
bcc-csm1-1	3, 1	25(13)	2(17)	7	1	2	G	y
CanESM2	5, 5	33(23)	14(26)	15	6	21	M (E)	y
inmcm4	1, 1	27(32)	28(32)	2	27	24	M (E)	y
ACCESS1.3	1, 1	13(17)	26(23)	22	26	27	M (O)	
HadGEM2-AO	1, 1	29(33)	12(35)	31	17	28	M (O)	
FGOALS_g2	5, 1	36(36)	3(36)	28	24	30	M (E)	
MIROC-ESM-CHEM	1, 1	31(34)	16(31)	35	30	33	M (E)	
MIROC-ESM	3, 1	30(24)	29(21)	33	33	35	P	
Cluster 5								
GISS-E2-H	6, 1	5(5)	23(3)	12	21	12	M (O)	y
CNRM-CM5	10, 5	9(2)	7(2)	18	31	13	M (O)	y
HadGEM2-CC	1, 1	11(16)	17(18)	24	20	16	M (E)	
IPSL-CM5B-LR	1, 1	35(35)	36(33)	34	10	36	M (O)	
Cluster 6								
GISS-E2-R	6, 1	1(1)	22(1)	13	14	1	G	y
EC-EARTH	4, 4	3(6)	19(4)	5	19	4	M (E)	y
MPI-ESM-LR	3, 3	2(4)	30(5)	4	35	8	M (O)	y
CMCC-CMS	1, 1	24(30)	27(27)	11	12	20	M (E)	
CMCC-CESM	1, 1	16(22)	31(20)	30	18	29	M (O)	
MRI-CGCM3	3, 1	17(18)	34(19)	26	36	32	M (E)	
FIO-ESM	3, 3	34(26)	32(25)	21	29	34	P	

Table 1 (continued)

Model	Members (HIST, RCP)	sm-TAS	sm-PR	pm-JJA	pm-DJF	Total	Category	Decision
Cluster 7								
GFDL-ESM2M	1, 1	10(20)	11(16)	10	3	5	G	y
GFDL-ESM2G	1, 1	21(27)	1(24)	20	22	11	M (O)	
GFDL-CM3	5, 1	15(8)	10(8)	27	25	17	M (E)	
CMCC-CM	1, 1	26(31)	21(34)	23	15	25	M (O)	

The number of initial conditions realizations included for each model over the historical (HIST) and RCP8.5 periods are shown in the second column. "sm-TAS" and "sm-PR" show model ranks for standard TAS and PR metrics based on one realization of every model whereas the numbers in parentheses indicate model ranks using all available initial conditions ensemble members. "pm-JJA" and "pm-DJF" show model ranks for JJA and DJF process metrics. "Category" shows model categorization based on performance (G: Good, M: Mixed, P: Poor) and as essential (E) or optional (O) for models with mixed performance. Retained models are indicated by "y" under "Decision." Models are listed according to their cluster membership (Fig. 5) and ordered according to the "Total" rank for every cluster (Fig. 4b)

from each of the 7 clusters, thereby preserving the diversity in the spatial pattern of summer precipitation projections.

Figure 6 shows the effect of model elimination on the annual and seasonal mean temperature and precipitation, projections and biases. As anticipated, the 16 models retained span the ranges in temperature projections reasonably well in all four seasons. Apart from HadGEM2-ES that projects one of the highest increases in temperature, most models with high temperature projections are eliminated. While this may suggest a relationship between performance and projections, the top performers in the retained set show very different projections. For instance, two of the top six performers—EC-EARTH and HadGEM2-ES—provide seasonal projections that are different by over 1.8 °C in all seasons (Fig. 6), including the spread across their multiple IC members. In the case of precipitation, models with both increases and decreases are eliminated with no clear consensus in the sign and magnitude of precipitation projections among retained models. The only retained model that projects future drying in annual and winter mean precipitation is inmcm4, which successfully captures summer processes but shows mixed performance for all the other metrics. CanESM2 has one of the largest warm biases in the ensemble, and also, poorly captures summer process metrics leading to a dry bias. Yet, its better performance for winter process metrics makes it relevant to future regional climate change studies. TS14-15 identified six better performing models for summer metrics from a set of 15, and we retain three of those (CanESM2, HadGEM2-ES, GFDL-ESM2M). The inclusion of a large number of metrics and the diversity criterion results in the elimination of the other three models (CCSM4, MRI-CGCM3, FGOALS-g2). For example, FGOALS-g2 performs well for summer precipitation but not for temperature and winter circulation metrics and is eliminated since its summer precipitation projections are very similar to one of the retained models (HadGEM2-ES). Model biases in Fig. 6 and ranks in Table 1 also indicate that at least three retained models (CanESM2, bcc-csm1-1-m, inmcm4) have questionable performances over the NEUS. But they are not eliminated because their relatively skillful performance for at least one set of metrics and diverse spatial patterns for JJA PR projections warrant more

investigation. Overall, the diversity in the magnitude of projections is well represented by the retained set as indicated by very similar ensemble mean projections based on the full CMIP5 MME and the 16 retained models (red symbols in Fig. 6).

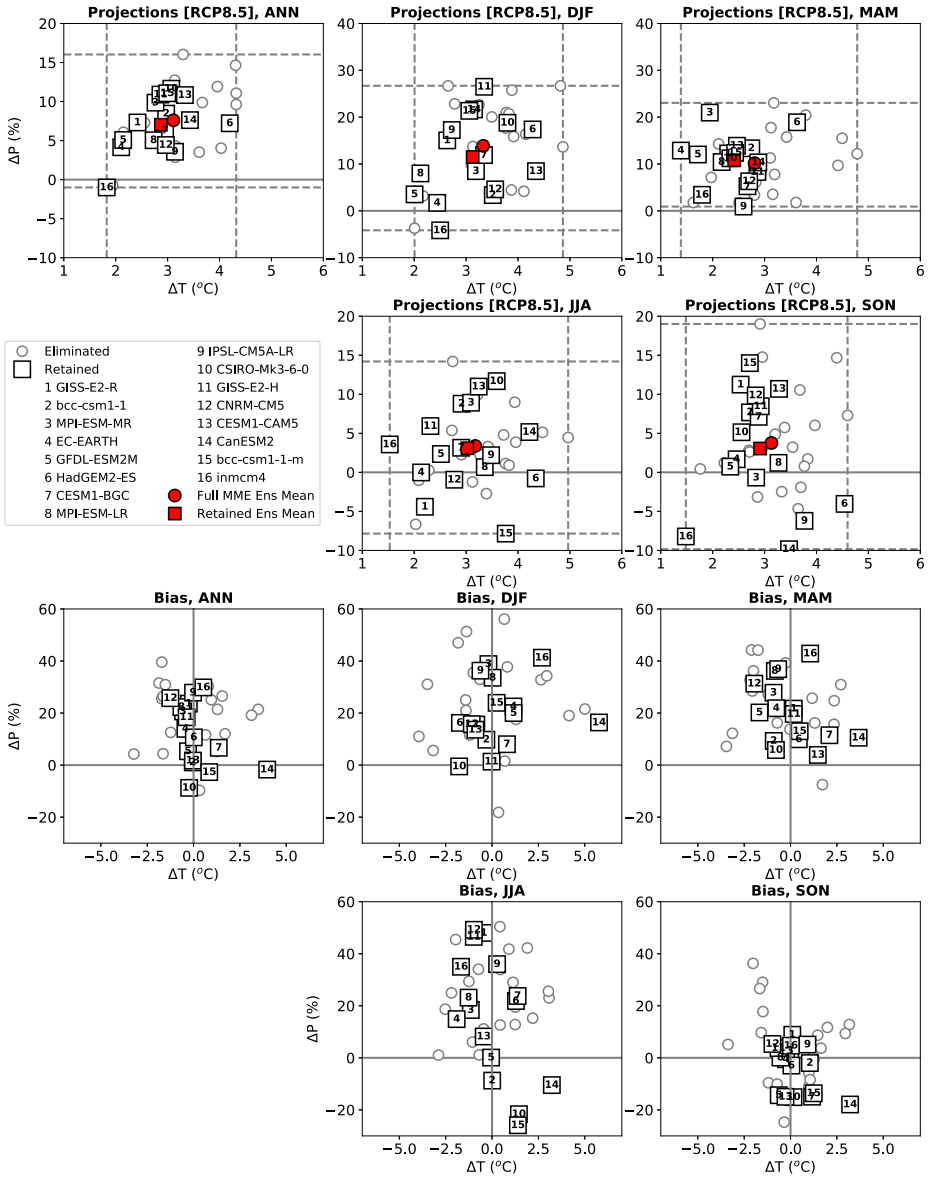

Fig. 6 Annual and seasonal mean projections (top) and biases (bottom) for NEUS temperature and precipitation for eliminated (gray circles) and retained (squares) CMIP5 models. The dotted lines show extreme projections based on the full CMIP5 ensemble and the red symbols show ensemble mean projections based on the full CMIP5 MME (red circles) and the retained models (red squares)

5 Conclusions and discussion

We analyze the performance and projections of 36 CMIP5 climate models over the Northeast U.S. to reduce the size of the ensemble for easier application in regional climate change and impacts studies. We eliminate 20 of 36 models, retaining 16 that are (a) credible in terms of their historical performance across a variety of metrics and (b) diverse in that they provide a wide range of temperature and precipitation projections consistent with uncertainties explored in the multi-model ensemble. We discriminate among models based on their ability to capture seasonal mean climatology, annual cycles, interannual variability, and large-scale atmospheric processes important on seasonal scales—aspects that GCMs are designed to capture reasonably well. This assessment, however, is not sufficient to establish the reliability of model projections as we find no consistent relationship between model biases and future projections for seasonal temperature and precipitation in the Northeast. This makes criterion (b) a critical component of our model sub-selection framework.

There is no one best performing model for the NEUS. This finding is not surprising given that we use a multitude of assessment metrics. Many studies focusing on global (Gleckler et al. 2008) and regional (Sheffield et al. 2013; Rupp et al. 2013) assessments demonstrate that identifying best climate models is difficult since their performance varies significantly across regions, seasons, and variables. Nevertheless, we show that it is possible to reduce the ensemble size substantially without overly constraining the full range of MME projections. In fact, we find that better performers in the retained set show diverse patterns of summer precipitation projections that are consistent with uncertainties in the full ensemble.

Past research has shown that internal climate variability is an important source of uncertainty at regional scales (Hawkins and Sutton 2009; Deser et al. 2014). The impact of IV on model selection is evaluated by analyzing all available multi-member initial conditions ensembles for a subset of CMIP5 models. The IC ensembles show a substantial spread in biases and projected changes in NEUS seasonal temperature and precipitation for many individual models. The inter-model spread in seasonal biases, however, is generally significantly higher than that for the multi-member IC ensemble of any given model. This is helpful for model discrimination even though the actual rankings may be different when the impact of internal variability is suitably explored. While the final model ranks are constructed using only the first realization (r1i1p1) of every model, we use the assessment of multiple IC realizations to inform performance-based model categorization (i.e., good, mixed, poor). This helps us avoid eliminating models when their alternate IC realizations are more skillful (for standard metrics) than the first realization (e.g., CSIRO-Mk3-6-0). Nonetheless, further evaluation of the impact of IV on process metrics will be necessary to help find realizations from multiple IC members that produce credible overall performances. The impact of IV on projections, while substantial, is only highlighted in this study but not taken into account in the decision framework because multiple IC members for RCP8.5 are available only for a small fraction of CMIP5 models (11 of 36). These results indicate the need to analyze systematically designed large ensembles (e.g., Kay et al. 2015; Mote et al. 2016) to explore the impact of internal variability on the assessment of model performance and projections.

The idea of either constraining or producing probabilistic projections, while desirable, faces several methodological and interpretive challenges (Tebaldi and Knutti 2007; Weigel et al. 2010). In this study, we simply use binary weights (eliminate or select) to discriminate among models and consider projections from the 16 retained models equally likely. The full CMIP5 ensemble as well as models within our retained set, however, exhibit a wide range in

their historical performance for any individual metric with a few models performing better overall. Selecting say only the top six better performers (see "Total" row in Fig. 4b) would eliminate most models with high-end temperature and precipitation projections (see Fig. 6). This suggests a scope to refine the projections by developing a weighting scheme that takes into account model performance as well as interdependence (e.g., Knutti et al. 2017). Such an approach, however, requires model weights to be derived using a detailed physical understanding of model behavior, mainly of simulated climate variability and extremes in addition to the metrics considered here, to avoid providing overconfident projections. Analysis of extremes, which is the focus of ongoing research, is also crucial to establish model reliability in capturing variables relevant for impacts assessment and could lend further credibility to the selected models.

Acknowledgements We acknowledge the WCRP Working Group on Coupled Modelling, and thank the climate modeling centers for producing and making available model output. We thank the three anonymous reviewers for their careful review and insightful comments that have helped improve the manuscript substantially.

Funding information This research was supported by the U.S. DOI's Northeast Climate Adaptation Science Center by Grant or Cooperative Agreement No. G12AC00001 from the United States Geological Survey (USGS) and in part by NSF CAREER Award No. 1056216. Its contents are solely the responsibility of the authors.

References

Barsugli JJ, Guentchev G, Horton RM, Wood A, Mearns LO, Liang XZ, Winkler JA, Dixon K, Hayhoe K, Rood RB et al (2013) The practitioner's dilemma: how to assess the credibility of downscaled climate projections. Eos Trans Amer Geophys Union 94(46):424–425

Bradbury JA, Keim BD, Wake CP (2003) The influence of regional storm tracking and teleconnections on winter precipitation in the northeastern United States. Ann Assoc Am Geogr 93(3):544–556

Dee DP, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P et al (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597

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

Deser C, Phillips AS, Alexander MA, Smoliak BV (2014) Projecting North American climate over the next 50 years: uncertainty due to internal variability*. J Clim 27(6):2271–2296

Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. J Geophys Res Atmos 113(D06104). https://doi.org/10.1029/2007JD008972

Harris I, Jones P, Osborn T, Lister D (2014) Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 dataset. Int J Climatol 34(3):623–642

Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. Bull Am Meteorol Soc 90(8):1095–1107

Horton R, Yohe G, Easterling W, Kates R, Ruth M, Sussman E, Whelchel A, Wolfe D, Lipschultz F (2014) Ch. 16: Northeast. Climate Change Impacts in the United States: The Third National Climate Assessment, pp 371–395

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J et al (1996) The NCEP/NCAR 40-year reanalysis project. Bullet Amer Meteorol Soc 77(3):437–472

Karmalkar AV, Bradley RS (2017) Consequences of global warming of 1.5 °C and 2 °C for regional temperature and precipitation changes in the contiguous United States. PloS one 12(1):e0168697

Kay J, Deser C, Phillips A, Mai A, Hannay C, Strand G, Arblaster J, Bates S, Danabasoglu G, Edwards J et al (2015) The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. Bull Am Meteorol Soc 96(8):1333–1349

Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl GA (2010) Challenges in combining projections from multiple climate models. J Clim 23(10):2739–2758

Knutti R, Masson D, Gettelman A (2013) Climate model genealogy: Generation CMIP5 and how we got there. Geophys Res Lett 40(6):1194–1199

Knutti R, Sedláček J, Sanderson BM, Lorenz R, Fischer EM, Eyring V (2017) A climate model projection weighting scheme accounting for performance and interdependence. Geophys Res Lett 44(4):1909–1918

Lynch C, Seth A, Thibeault J (2016) Recent and projected annual cycles of temperature and precipitation in the northeast United States from CMIP5. J Clim 29(1):347–365

Maloney ED, Camargo SJ, Chang E, Colle B, Fu R, Geil KL, Hu Q, Jiang X, Johnson N, Karnauskas KB et al (2014) North American climate in CMIP5 experiments: Part III: Assessment of twenty-first-century projections. J Clim 27(6):2230–2270

Masson D, Knutti R (2011) Climate model genealogy. Geophys Res Lett 38(L08703). https://doi.org/10.1029/2011GL046864

Matsuura K, Willmott CJ (2012) Terrestrial precipitation: 1900-2010 gridded monthly time series (v. 3.01). Center for Climatic Research, Department of Geography, University of Delaware Newark, DE, USA

McSweeney CF, Jones RG, Booth BB (2012) Selecting ensemble members to provide regional climate change information. J Clim 25(20):7100–7121

McSweeney C, Jones R, Lee R, Rowell D (2015) Selecting CMIP5 GCMs for downscaling over multiple regions. Clim Dyn 44(11-12):3237–3260

Meinshausen M, Smith SJ, Calvin K, Daniel JS, Kainuma M, Lamarque JF, Matsumoto K, Montzka S, Raper S, Riahi K et al (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Clim Change 109(1-2):213

Monerie PA, Sanchez-Gomez E, Boé J (2017) On the range of future Sahel precipitation projections and the selection of a sub-sample of CMIP5 models for impact studies. Clim Dyn 48(7-8):2751–2770

Mote P, Brekke L, Duffy PB, Maurer E (2011) Guidelines for constructing climate scenarios. Eos Trans Amer Geophys Union 92(31):257–258

Mote PW, Allen MR, Jones RG, Li S, Mera R, Rupp DE, Salahuddin A, Vickers D (2016) Superensemble regional climate modeling for the western United States. Bull Am Meteorol Soc 97(2):203–215

Ning L, Bradley RS (2015) Winter climate extremes over the northeastern United States and southeastern Canada and teleconnections with large-scale modes of climate variability. J Clim 28(6):2475–2493

Overland JE, Wang M, Bond NA, Walsh JE, Kattsov VM, Chapman WL (2011) Considerations in the selection of global climate models for regional climate projections: The Arctic as a case study. J Clim 24(6):1583–1597

Pierce DW, Barnett TP, Santer BD, Gleckler PJ (2009) Selecting global climate models for regional climate change studies. Proc Natl Acad Sci 106(21):8441–8446

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65

Rowell DP, Senior CA, Vellinga M, Graham RJ (2016) Can climate projection uncertainty be constrained over Africa using metrics of contemporary performance? Clim Chang 134(4):621–633

Rupp DE, Abatzoglou JT, Hegewisch KC, Mote PW (2013) Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. J Geophys Res: Atmos 118(19):10–884

Sanderson BM, Knutti R, Caldwell P (2015) A representative democracy to reduce interdependency in a multimodel ensemble. J Clim 28(13):5171–5194

Seager R, Naik N, Vecchi GA (2010) Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming. J Clim 23(17):4651–4668

Sheffield J, Barrett AP, Colle B, Nelun Fernando D, Fu R, Geil KL, Hu Q, Kinter J, Kumar S, Langenbrunner B et al (2013) North American climate in CMIP5 experiments. Part I: evaluation of historical simulations of continental and regional climatology. J Clim 26(23):9209–9245

Snover AK, Mantua NJ, Littell JS, Alexander MA, Mcclure MM, Nye J (2013) Choosing and using climate-change scenarios for ecological-impact assessments and conservation decisions. Conserv Biol 27(6):1147–1157

Staudinger MD, Morelli TL, Bryan AM (2015) Integrating climate change into northeast and midwest state wildlife action plans. DOI Northeast Climate Science Center Report, Amherst

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res: Atmos 106(D7):7183–7192

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93(4):485–498

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society of London A: Mathematical. Phys Eng Sci 365(1857):2053–2075

Thibeault JM, Seth A (2014) A framework for evaluating model credibility for warm-season precipitation in northeastern North America: a case study of CMIP5 simulations and projections. J Clim 27(2):493–510

Thibeault JM, Seth A (2015) Toward the credibility of Northeast United States summer precipitation projections in CMIP5 and NARCCAP simulations. J Geophys Res Atmos 120(19):10050–10073. https://doi.org/10.1002/2015JD023177

Vano JA, Kim JB, Rupp DE, Mote PW (2015) Selecting climate change scenarios using impact-relevant sensitivities. Geophys Res Lett 42(13):5516–5525

Wallace JM, Deser C, Smoliak BV, Phillips AS (2016) Attribution of climate change in the presence of internal variability. In: Climate change: Multidecadal and Beyond, World Scientific, pp 1–29

Weigel AP, Knutti R, Liniger MA, Appenzeller C (2010) Risks of model weighting in multimodel climate projections. J Clim 23(15):4175–4191

Whetton P, Hennessy K, Clarke J, McInnes K, Kent D (2012) Use of representative climate futures in impact and adaptation assessment. Clim Change 115(3-4):433–442

Wuebbles D, Fahey D, Hibbard K, Dokken B, Stewart B, Maycock T (2017) Climate science special report: Fourth national climate assessment, Volume I. In: Washington, DC, pp 470

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Ambarish V. Karmalkar[1] · Jeanne M. Thibeault[2] · Alexander M. Bryan[3] · Anji Seth[2]

1 Northeast Climate Adaptation Science Center and Department of Geosciences, University of Massachusetts Amherst, Massachusetts, Amherst, MA, USA

2 Department of Geography, University of Connecticut, Storrs, Connecticut, USA

3 Northeast Climate Adaptation Science Center, U.S. Geological Survey, University of Massachusetts Amherst, Massachusetts, Amherst, MA, USA

Reproduced with permission of copyright owner.
Further reproduction prohibited without permission.