3-1-2015

Comparison between Observed and Model-Simulated Atmospheric CirculationPatterns Associated with Extreme Temperature Days over North AmericaUsing CMIP5 Historical Simulations

Paul C. Loikith
Portland State University

Anthony J. Broccoli
Rutgers University

Citation Details

Loikith, P. C., & Broccoli, A. J. (2015). Comparison between Observed and Model-Simulated Atmospheric Circulation Patterns Associated with Extreme Temperature Days over North America Using CMIP5 Historical Simulations. Journal of Climate, 28(5), 2063-2079.

Comparison between Observed and Model-Simulated Atmospheric Circulation Patterns Associated with Extreme Temperature Days over North America Using CMIP5 Historical Simulations

PAUL C. LOIKITH

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

ANTHONY J. BROCCOLI

Department of Environmental Sciences, Rutgers, The State University of New Jersey, New Brunswick, New Jersey

(Manuscript received 5 September 2013, in final form 26 September 2014)

ABSTRACT

Circulation patterns associated with extreme temperature days over North America, as simulated by a suite of climate models, are compared with those obtained from observations. The authors analyze 17 coupled atmosphere–ocean general circulation models contributing to the fifth phase of the Coupled Model Intercomparison Project. Circulation patterns are defined as composites of anomalies in sea level pressure and 500-hPa geopotential height concurrent with days in the tails of temperature distribution. Several metrics used to systematically describe circulation patterns associated with extreme temperature days are applied to both the observed and model-simulated data. Additionally, self-organizing maps are employed as a means of comparing observed and model-simulated circulation patterns across the North American domain. In general, the multimodel ensemble resembles the observed patterns well, especially in areas removed from complex geographic features (e.g., mountains and coastlines). Individual model results vary; however, the majority of models capture the major features observed. The multimodel ensemble captures several key features, including regional variations in the strength and orientation of atmospheric circulation patterns associated with extreme temperatures, both near the surface and aloft, as well as variations with latitude and season. The results from this work suggest that these models can be used to comprehensively examine the role that changes in atmospheric circulation will play in projected changes in temperature extremes because of future anthropogenic climate warming.

1. Introduction

Climate model simulations of future climate project increases in extreme heat events over much of the globe by the end of the twenty-first century because of anthropogenic global warming (Meehl and Tebaldi 2004; Tebaldi et al. 2006; Meehl et al. 2007, 2009). Many recent heat events highlight the dangers of such changes. The European heat wave of 2003 was an example of an event that is likely to become more common in the future, leaving large segments of the world population vulnerable to unprecedented heat (e.g., Beniston 2004; Schär et al. 2004; Stott et al. 2004). Another notable example is the Russian heat wave that occurred in the summer of 2011. Dole et al. (2011) argue that this unusual heat event was largely a result of natural variability while Rahmstorf and Coumou (2011) argue that anthropogenic global warming played a role in the event. Otto et al. (2012) suggest that this particular event was a combination of an unusual natural event and enhanced heat due to global warming.

Recent analyses of extremes indices show a warming of the cold tail of the temperature distribution along with an increase in warm nights globally in recent decades (Frich et al. 2002; Alexander et al. 2006; Griffiths and Bradley 2007; Brown et al. 2010). Much of this warming can be attributed to anthropogenic radiative forcing (Christidis et al. 2005, 2011; Morak et al. 2011; Zwiers et al. 2011). Changes in atmospheric circulation resulting from the changing climate could result in nonlinear changes in the temperature probability distribution, as well as regional variability in how temperature extremes are affected. When applied to general circulation model (GCM)

Corresponding author address: Paul C. Loikith, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., M/S 300-233, Pasadena, CA 91109.
E-mail: paul.c.loikith@jpl.nasa.gov

DOI: 10.1175/JCLI-D-13-00544.1

TABLE 1. List of the CMIP5 models used. The letter assignment used in the results presentation is provided in parentheses to the right of the model name.

Model name	Horizontal resolution (degrees lat × degrees lon)	Modeling group	Z500 available
BNU-ESM (G)	2.81 × 2.81	College of Global Change and Earth System Science (GCESS), China	Yes
CanESM2 (C)	2.81 × 2.81	CCCma, Canada	Yes
CMCC-CM (B)	0.75 × 0.75	CMCC, Italy	Yes
CNRM-CM5 (K)	1.41 × 1.41	CNRM-Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), France	Yes
FGOALS-g2 (J)	3.00 × 2.81	LASG-Center of Earth System Science (CESS), China	Yes
FGOALS-s2 (P)	1.67 × 2.81	LASG-IAP, China	Yes
GFDL-ESM2G (A)	2.00 × 2.50	GFDL, United States	No
GFDL-ESM2M (H)	2.00 × 2.50	GFDL, United States	Yes
HadGEM2-CC (O)	1.25 × 1.88	MOHC, United Kingdom	Yes
INM-CM4 (E)	1.50 × 2.00	INM, Russia	No
IPSL-CM5A-LR (D)	1.88 × 3.75	IPSL, France	Yes
IPSL-CM5A-MR (L)	1.26 × 2.50	IPSL, France	Yes
MIROC5 (N)	1.41 × 1.41	MIROC, Japan	Yes
MIROC-ESM-CHEM (Q)	2.81 × 2.81	MIROC, Japan	Yes
MPI-ESM-LR (I)	1.88 × 1.88	MPI-M, Germany	Yes
MPI-ESM-MR (F)	1.88 × 1.88	MPI-M, Germany	Yes
MRI-CGCM (M)	1.13 × 1.13	MRI, Japan	Yes

simulations, extreme value statistics show a warming trend with the cold tail of the temperature distribution warming more than the warm tail, especially in areas of sea ice and snow cover retreat (Kharin and Zwiers 2000, 2005; Kharin et al. 2007). Donat and Alexander (2012) show that daily temperature distributions have warmed, further confirming the observed increase in warm extremes and decrease in cold extremes globally. Rowe and Derry (2012) show that the frequency of record warm temperatures has been increasing, while the occurrence of record cold temperature has been decreasing across the continental United States in recent years. There is evidence that changes in circulation may result in a dampening of the warming and potential regional increases in extreme cold events (Vavrus et al. 2006; Kodra et al. 2011).

Loikith and Broccoli (2012, hereinafter LB12) systematically identified and described the primary atmospheric circulation patterns associated with extreme temperature days over North America during the twentieth century. The diagnostic metrics devised in LB12 allow for systematic comparison between observations and model-simulated circulation data. This work follows the framework of LB12 and compares model simulations of atmospheric circulation patterns associated with extreme temperature days over North America from historical climate model simulations of the twentieth century. Section 2 describes the datasets, including observational and model-simulated data, and the methodology used. Section 3 follows with a comparison between observations and models for several metrics developed in LB12, and section 4 uses self-organizing maps as a basis for domain-wide comparison and for individual cases. Section 5 presents concluding remarks and implications for future work and use of these climate models to understand projected changes in temperature extremes.

2. Data and methodology

a. Data

All model output used in this work is from the fifth phase of the Coupled Model Intercomparison Project (CMIP5). This is the latest phase of a coordinated effort by modeling groups worldwide to systematically perform numerous prescribed climate model experiments. The output from the model simulations is archived and available to the scientific community through the Program for Climate Model Diagnosis and Intercomparison (PCMDI). A detailed description of the experimental design is available in Taylor et al. (2012). This work utilizes 17 individual models from 13 different modeling groups (Table 1). All output used comes from the historical simulation of each model, and only one ensemble member from each model is used. The historical simulations are prescribed experiments conducted by each modeling group that simulate the past climate using observed radiative forcing. This work utilized daily maximum and minimum temperature, sea level pressure

(SLP), and 500-hPa geopotential height (Z500) from these simulations. Only models with all three of these variables available for the historical and two global warming experiments (RCP4.5 and RCP8.5) were used. The requirement for future simulation output is imposed in anticipation of future work that will analyze global warming simulations in comparison with historical simulations for each model. Geopotential height was provided directly by each model, with the exception of the INM-CM4.0 and GFDL-ESM2G models, for which the hypsometric equation was vertically integrated to calculate Z500. All models lacking geopotential height on pressure coordinates or the variables needed to vertically integrate the hypsometric equation were not used in this work.

Daily, observed temperature data were obtained from the Hadley Centre Global Historical Climatology Network-Daily (HadGHCND) gridded daily temperature dataset. The dataset is a joint project between the Met Office Hadley Centre for Climate Change (MOHC) and the U.S. National Climatic Data Center (NCDC). The dataset is on a global domain, and the majority of the observations that are applied to the gridding process are from NCDC's Global Historical Climatology Network-Daily (GHCND). The resolution of the grid is 2.5° latitude by 3.75° longitude. The dataset has two products: gridded observed daily maximum and minimum temperatures and their anomalies (both available from 1950 to 2011). We use the anomalies in this work, which are computed by subtracting the 5-day running mean of a 30-yr daily climatology (1961–90) from the actual temperature. A more detailed description of the dataset and the gridding process can be found in Caesar et al. (2006).

The SLP and Z500 fields used to determine circulation patterns associated with observed temperature extremes were obtained from the National Centers for Environmental Prediction's Reanalysis 1 (Kalnay et al. 1996) (subsequent references to the combination of surface air temperature observations and SLP and Z500 from reanalysis will be labeled "observed" to distinguish them from the CMIP5 simulations). To calculate anomalies for observed SLP and Z500, the same method and reference period as Caesar et al. (2006) for the gridded temperature anomalies was used. The same method of calculating anomalies was also applied to simulated SLP and Z500, and the model circulation variables were kept on their respective native grid. Only data for North America are used, defined here as all land north of 17.5°N, bounded on the east and west by the Atlantic and Pacific Oceans respectively. This equates to 315 grid cells in the analysis domain.

b. Methodology

The methods for calculating circulation anomaly patterns using the CMIP5 data follow the technique used by LB12, where a more detailed description can be found. All CMIP5 temperature data were regridded to the resolution of the HadGHCND dataset. Because the resolution for each model is finer than that of the HadGHCND dataset, an area-averaging technique was used in all cases for the regridding process. Temperature extreme days are defined as those days falling below the 5th (Tx5) and above the 95th (Tx95) percentile in the temperature frequency distribution, and extremes were identified for January, April, July, and October. This results in 47 (46) extreme temperature days for 31- (30-) day months, except in the case of a tie for the 47th/46th most extreme temperature day when ties are included and the sample size is larger.

Following the procedure in LB12, composite patterns of anomalies in SLP and Z500 were computed for days concurrent with extreme temperatures at each grid cell for each model. The composite patterns are limited to 4500 km from the grid cell where the extreme temperatures are occurring. This distance is chosen as a compromise between capturing as much large-scale climate variability as possible while limiting the influence from distant features that are not likely to be influential on the occurrence of temperature extremes. All composite patterns were then regridded to a gridcell-relative grid such that the center of the domain is the grid cell where the extreme temperatures are identified (see Fig. 1 in LB12). While patterns for warm and cold daily maximum and minimum temperature extremes were computed and analyzed using both Z500 and SLP anomalies for January, April, July, and October, daily maximum extremes for January and July are the main focuses of this paper. April and October cases tended to resemble a combination of July and January patterns. Similarly, patterns associated with warm daily minimum extremes tended to resemble those associated with warm daily maximum extremes.

3. Comparison between model-simulated and observed composite analysis

This section applies some of the diagnostic metrics developed in LB12 to the CMIP5 models and compares the results with observations. In most of the analyses, the multimodel ensemble is presented as the median metric value of the 17 individual models. Individual models vary in horizontal resolution and physical complexity (Earth system models being the most complex), and the ability of the models to simulate the patterns identified with observations also varies considerably. Model names are not indicated in the text or figures; rather, each model is randomly assigned a letter, which can be referenced against

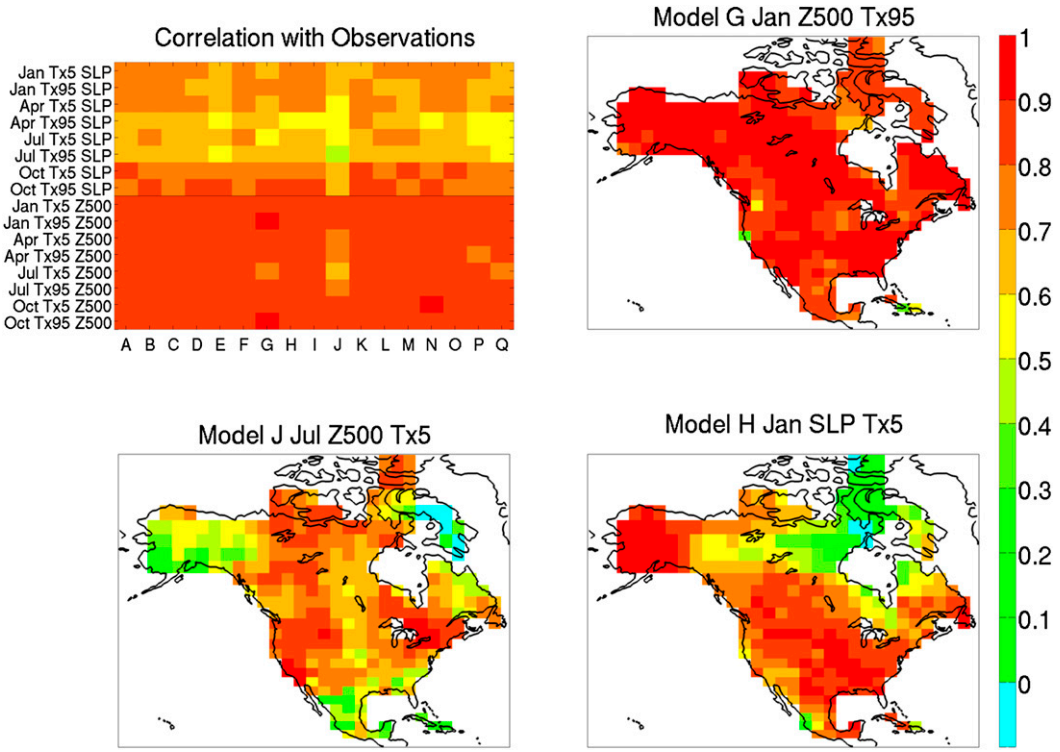

FIG. 1. (top left) Portrait diagram depicting the median value of the 315 correlation coefficients from pattern correlations between the simulated and observed composite patterns at each grid cell. Maps of correlation coefficients for select models and extremes: (top right) January Tx95 Z500 for model G, (bottom left) July Tx5 Z500 for model J, and (bottom right) January Tx5 SLP for model H.

model names in Table 1. The technique used in this section aims to summarize the overall ability of the models to simulate the characteristics of observed temperature extremes given the very large number of dimensions inherent to this analysis.

a. Composite pattern correlations

A portrait diagram (Gleckler et al. 2008) is used in Fig. 1 to summarize the ability of each individual model to simulate the composite pattern for each type of extreme. The random letter assignment for each model is indicated on the *x*-axis. The portrait diagram in Fig. 1 shows the median of the 315 pattern correlation coefficients computed between the observed composite pattern at each grid cell and the corresponding simulated composite pattern for each model. In general, correlation is very high between the model-simulated and observed composite patterns at Z500 in all seasons, with lower values for SLP. Near-surface circulation is more variable than circulation aloft, likely resulting in some of the weaker agreement between the models and observations in SLP patterns. The weakest correlations are for July Tx5 and Tx95 and April Tx95, when SLP circulation is relatively weak. Here, other mechanisms, such as those affecting the surface energy budget (i.e., soil moisture), as well as circulation patterns and processes on smaller scales may be important for extreme temperatures over much of the domain.

The maps in Fig. 1 show examples of individual model results that were used to create the portrait diagram. The map of January Tx95 Z500 for model G is an example of a model that correlates relatively well with observations, while the July Tx5 Z500 map for model J has a relatively low median correlation coefficient. The January Tx5 SLP example is a case that has neither high nor low median correlation in relation to other models. In both of the lower two panels in Fig. 1, large regional variability in correlation values suggests that models are able to simulate patterns in some regions more realistically than other regions, but not always for the same region.

To compare the basic common properties of the composite patterns, all of the 315 different circulation patterns corresponding to each grid cell in the domain were composited together to create a single ''grand composite.'' Figure 2 shows the grand composites for Z500 and SLP patterns for observations and the mean grand composite calculated as the average of the 17 simulated grand

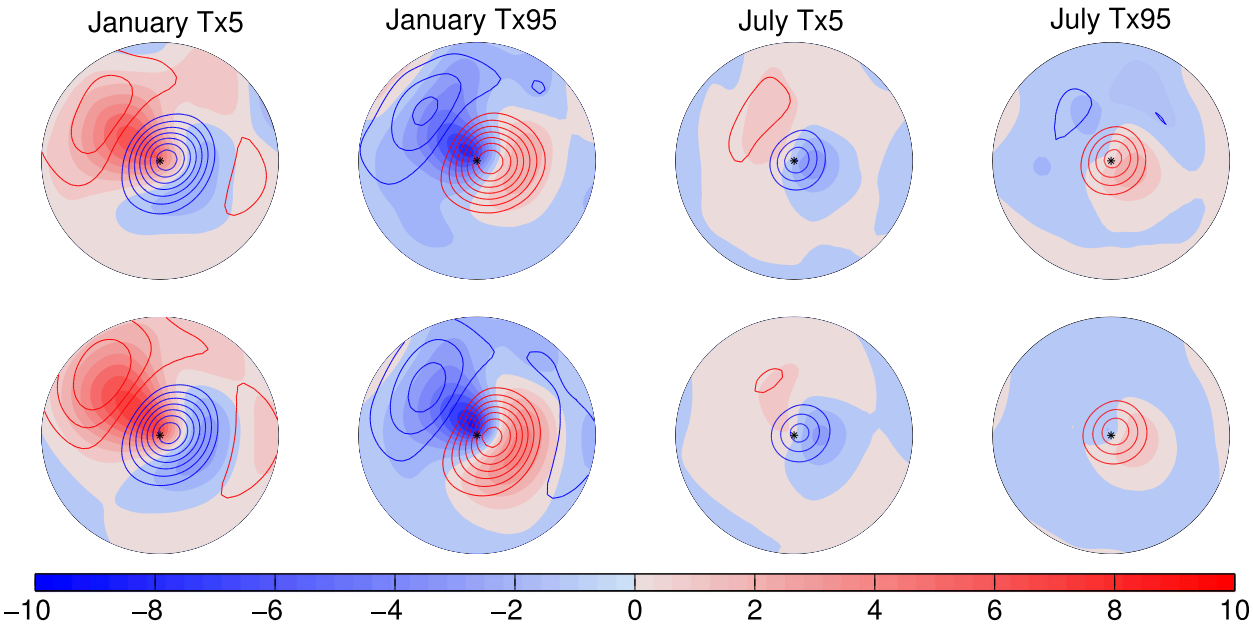

FIG. 2. (top) Observed and (bottom) multimodel ensemble mean grand composites for (from left to right) extreme cold January, extreme warm January, extreme cold July, and extreme warm July maximum temperature days. Color-filled contours (every 2 hPa) are SLP anomalies, and red and blue contours are for positive and negative Z500 anomalies, respectively. The radius of the area depicted is 4500 km with north oriented upward.

composites. The multimodel average grand composite patterns resemble the observed grand composites well in both January and July for both Z500 and SLP. The models realistically represent the local, strong anomaly at Z500 near the grid cell where the extreme temperatures are occurring as well as the weaker upstream and downstream anomalies. SLP patterns are also well represented by the model mean in both seasons. July Z500 patterns are slightly weaker in the multimodel ensemble grand composite, possibly resulting from more intracomposite variability inherent in the larger sample size used to create the simulated grand composite.

b. *Pattern symmetry and linearity*

Pattern symmetry is defined as the coefficient resulting from a pattern correlation between the circulation pattern associated with extreme warm days and the pattern associated with extreme cold days for a given location. A location with strong symmetry would have a pattern associated with cold days that is similar to but opposite in sign of the pattern associated with warm days. This would result in a correlation coefficient close to −1, which would indicate perfect symmetry. Here, all correlation coefficients have been multiplied by −1 so that a highly symmetrical pattern has a high, positive correlation coefficient, and asymmetry is characterized by a low or negative correlation coefficient. Pattern linearity describes how well the patterns scale with temperature. A linear pattern would have a composite pattern associated with extreme temperature days that was a linearly scaled version of the patterns associated with the rest of the temperature distribution. The linearity metric is defined as the RMS difference between the composite pattern and the regression pattern. The regression pattern is computed individually for each grid cell by regressing the circulation anomalies for each grid cell within 4500 km on the entire time series of temperature anomalies for the grid cell for which the pattern is being computed. The regression coefficients are then multiplied by the mean temperature anomaly for corresponding extreme temperature days so that the units of the regression patterns and the composite patterns are the same. Finally, this pattern is normalized by the standard deviation of the composite pattern. Both of these metrics were first developed and presented using observed temperatures and associated circulation patterns in LB12.

Because symmetry is calculated using patterns for both warm and cold extremes, results are only provided for maximum and minimum temperatures for each month and circulation variable. Figure 3 shows pattern symmetry for daily maximum temperature extremes, with the left column showing observed and the right column showing model results. In this case, the multimodel results are the median value of the 17 different

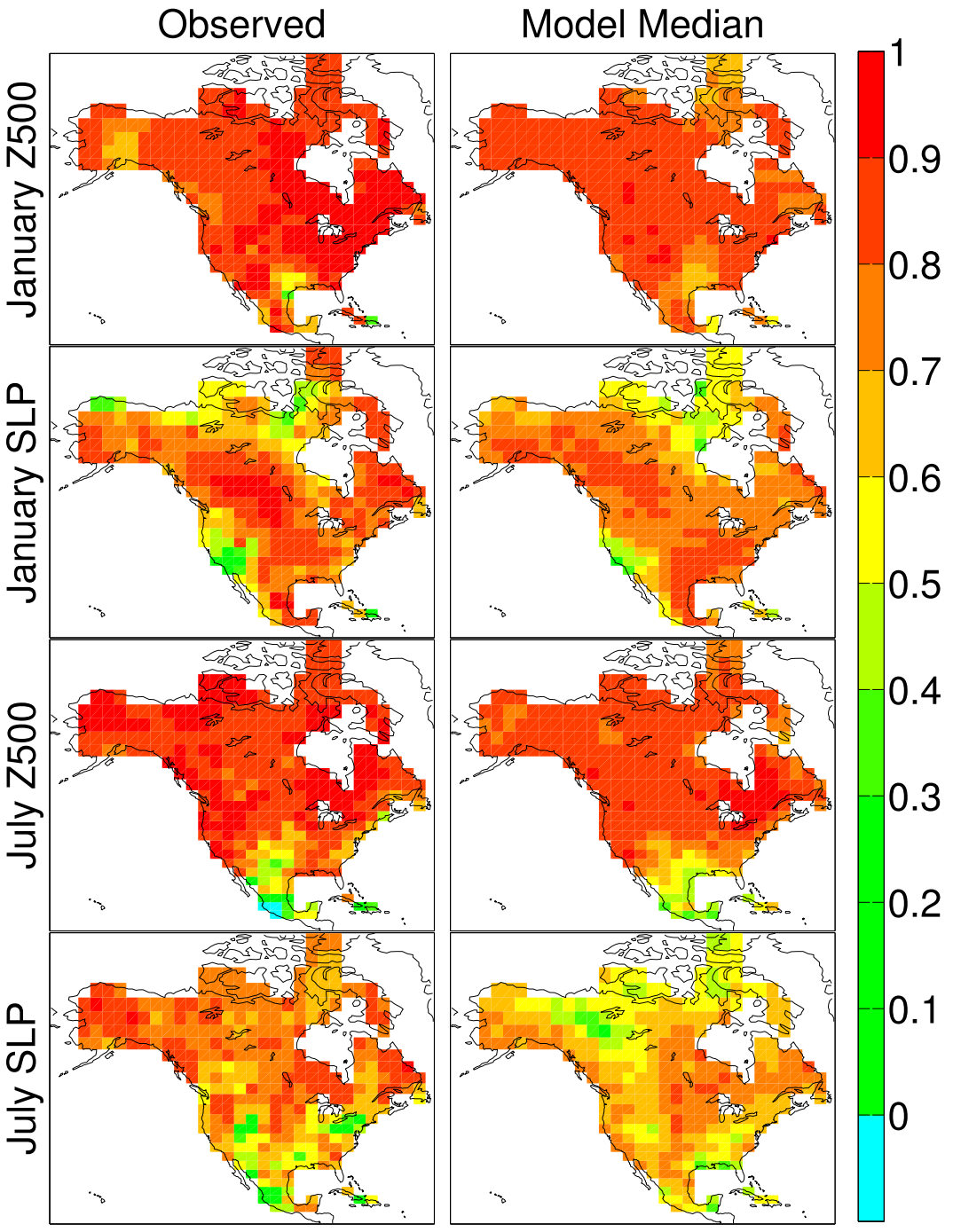

FIG. 3. Maps of symmetry of daily maximum temperature extremes from (left) observations and (right) models. The model results depict the median value across all 17 models at each location.

symmetry values corresponding to each model. Here, the median value was used as a summary of model results to more accurately represent the results of a typical model. In all four cases, the median symmetry maps depict the majority of the features found in observations, with the poorest performance being the July SLP case. The relatively weak resemblance here is consistent with the highly variable and relatively weak near-surface circulation patterns found during the summer.

Some notable regional variations in the degree of symmetry are well captured by the median model values. For example, the area of weak symmetry over

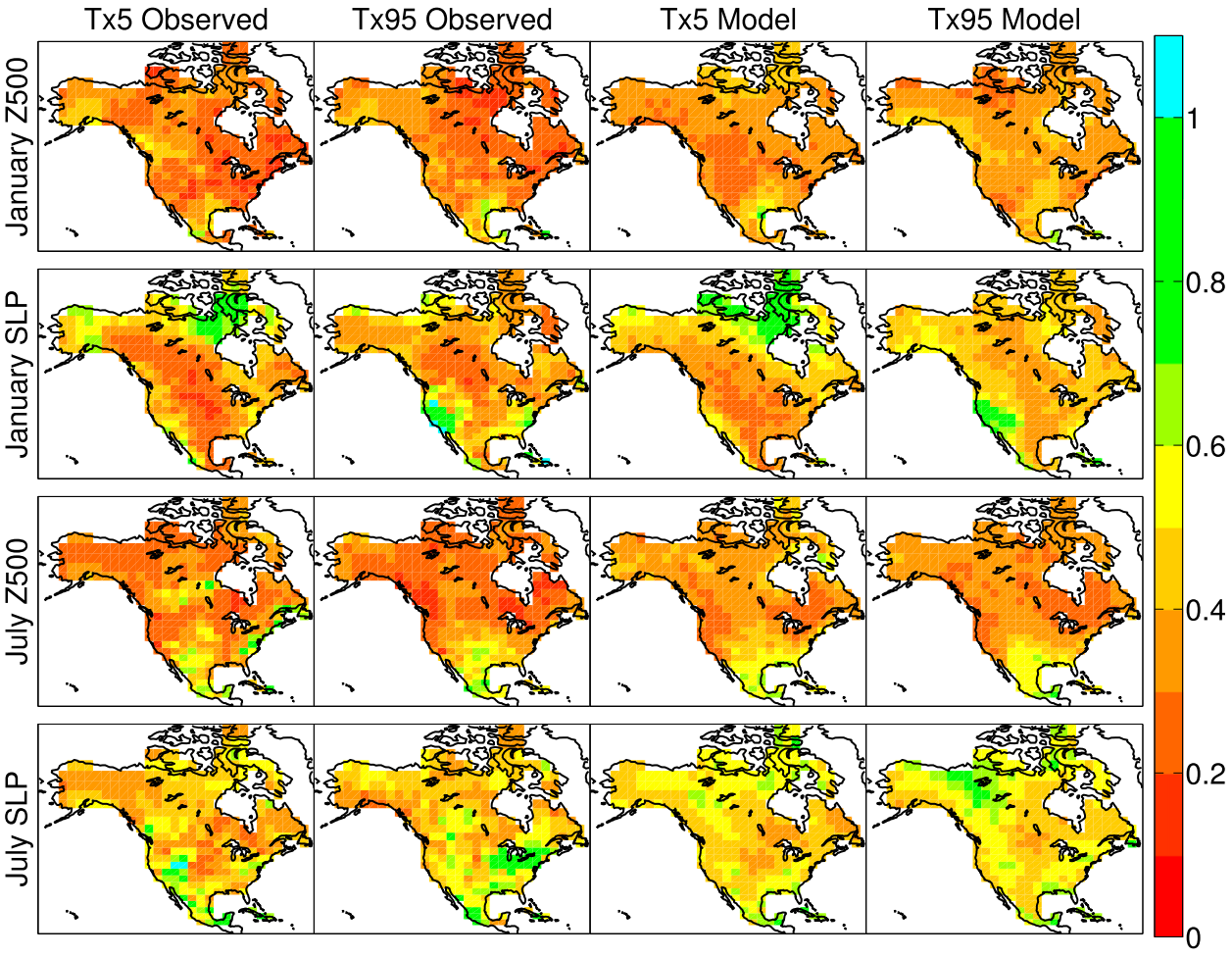

FIG. 4. Maps of pattern linearity for cold and warm daily maximum temperature extremes from (left),(middle left) observations and (middle right),(right) models: (top),(top middle) for January and (bottom middle),(bottom) for July. The model results depict the median value across all 17 models at each location.

the southwestern portion of the continent for the January maximum SLP example is depicted in the multimodel ensemble, but with slightly higher symmetry values. The band of weak symmetry in northern Canada is also represented in the model ensemble for January maximum SLP. The area of weak symmetry in the southern portions of the continent for July maximum Z500 is present in the model ensemble, while July SLP is overall less symmetrical in the model ensemble compared with observations.

Linearity values are presented in Fig. 4, with the observed values in the left two columns and model median values in the right two columns. For reference, a linearity value of zero would indicate perfect linearity. Because linearity is normalized by the standard deviation of the composite pattern, a value of one would indicate that the difference between the composite and regression-derived circulation patterns is as large as the spatial variability of the composite pattern itself and would indicate strong nonlinearity. Linearity shares some commonalities with symmetry, as they both describe how the patterns associated with days in the tails relate to each other, and the model ensemble captures these similarities in many cases. Unlike symmetry, linearity is calculated for each tail separately, so results are presented for both cold and warm extremes. In general, the model median linearity maps capture the broad features shown in the observational analysis. The area of weak linearity for the January Tx95 SLP case in the southwestern portion of the continent (corresponding to weak January maximum SLP symmetry) is present in the model mean, as is the area of weak linearity in northern Canada. Similar to symmetry, the poorest representation by the models is for July SLP, where several regional differences exist. In the July Tx5 SLP case, the area of the Rocky Mountains that has negative values for linearity is not present in the

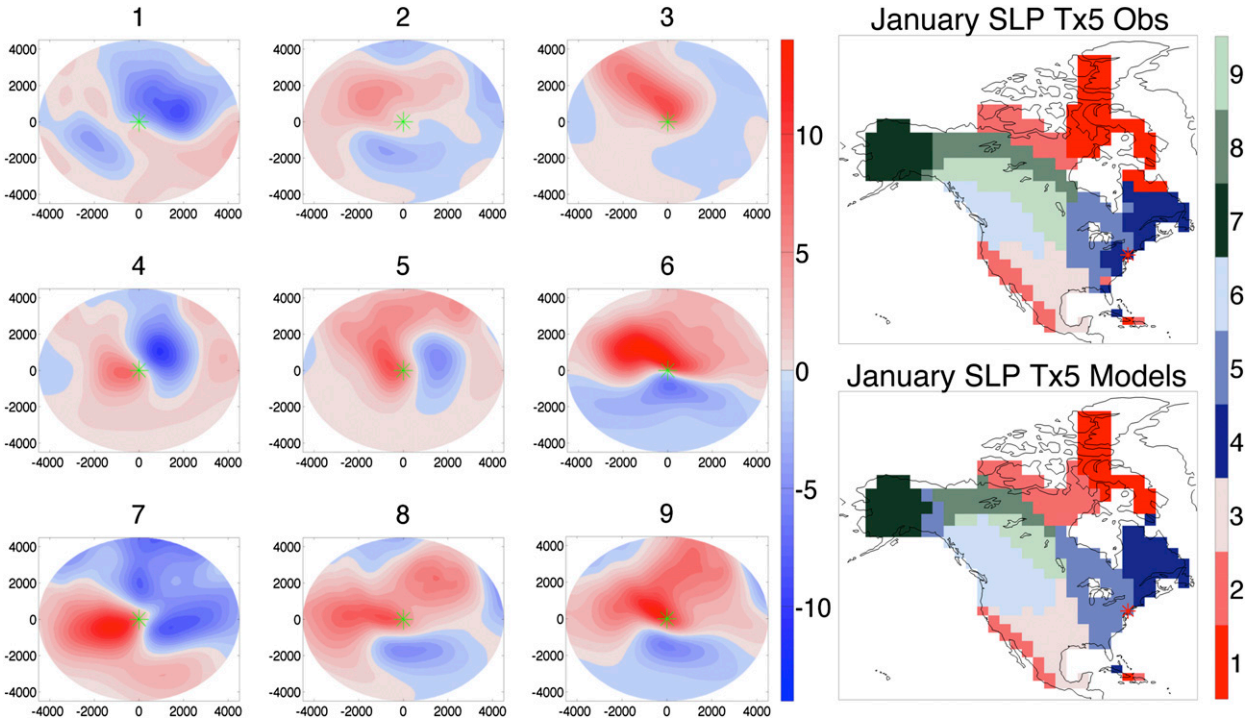

FIG. 5. (left) Self-organizing maps for composite patterns of SLP anomalies in hPa for January Tx5 days. Nodes are identified by the numbers above each panel. (right) Pointwise node assignments for (top) observed patterns and (bottom) multimodel ensemble mean patterns. The number indicated with shading in the maps indicates the node to which the local composite pattern is assigned. The asterisk in the maps indicates the grid cell used for the example in the following figure.

models. The area of weak linearity in the vicinity of the Great Lakes in the July Tx95 SLP case is also not present in the model ensemble. These are both areas where complex terrain and proximity to large water bodies have strong local influences on surface temperature, and because both the observations and the models are generally of coarse horizontal resolution, capturing the effects of features like the Great Lakes on circulation patterns may be challenging.

4. Synoptic analysis using self-organizing maps

In this section, self-organizing maps (SOMs) (Sheridan and Lee 2011) are employed as a tool to provide additional information about the composite patterns at all grid cells in observations and the CMIP5 models. SOMs are commonly used for studying and understanding synoptic climatology, often for a specific region of interest (e.g., Cavazos 2000; Hewitson and Crane 2002; Cassano et al. 2006). Similar to cluster analysis, SOM analysis uses a neural network algorithm to group patterns of 2-dimensional data into a user-defined number of reference clusters or ''nodes.'' Conventionally, SOM analysis uses a time series of a synoptic field (such as SLP anomalies) as input, and each time step is assigned to the reference node that is the most similar to the field for that time step. Here, SOMs are employed in an approach that differs slightly from this convention, in that the input is a vector in space rather than time. In other words, the 315 composite patterns corresponding to the 315 grid cells in the domain are the input, and the output is a node assignment for each grid cell based on the reference pattern that is most similar to its composite pattern. The choice of how many nodes to use is subjective, and a 3×3 matrix was chosen as the best representation of the spatial variability of the composite patterns for this application. SOMs were computed using the Matlab SOM Toolbox (Vesanto et al. 2000), with an initial radius of three, a final radius of 1, and using the ''ep'' neighborhood function, as these settings provided the most easily interpretable results. A detailed description of these parameters can be found in Liu et al. (2006).

To compare models with observations using SOMs, first the analysis was applied to only the observed patterns for a given extreme and circulation variable. These SOMs assignments and patterns are referred to as ''reference.'' Then, for each grid cell, the root-mean-square difference (RMSD) between the multimodel ensemble mean composite pattern and each of the nine reference patterns was

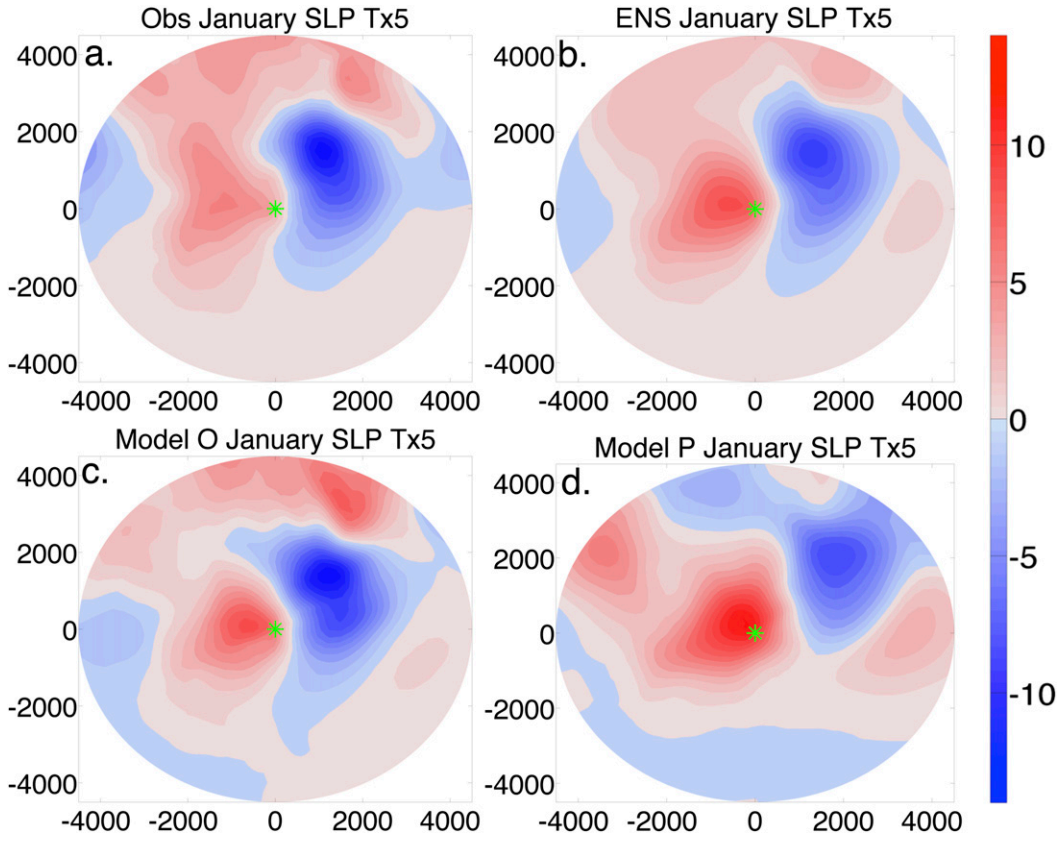

FIG. 6. Examples of SLP anomaly patterns in hPa for January Tx5 days at the mid-Atlantic grid cell, highlighted with a red asterisk in Fig. 5. Patterns are from (a) observations, (b) the multimodel ensemble mean, (c) model O, and (d) model P. The green asterisk is the point where the extreme temperatures are occurring.

computed. The model grid cell was assigned to the node where the RMSD for the simulated pattern was smallest. In the following sections, the results of the SOMs analysis are presented for Tx5 January and Tx95 July SLP and Z500 patterns. For each of the four analyses, an example is provided for a grid cell for the observed, multimodel ensemble mean, a model with strong resemblance to observations, and a model with weak resemblance to observations. To the knowledge of the authors, there is little guidance in the scientific literature regarding the assessment of statistical significant for SOMs, providing a limitation to this type of analysis. Therefore, the following results lean toward the qualitative; however, it is recognized that additional quantitative statistical rigor could benefit future application of SOMs in model comparison.

a. January Tx5

Figure 5 shows the SOM matrix on the left and the corresponding pointwise node assignments on the right for January Tx5 SLP composite patterns. Each node is referenced by the number above the panels. While some adjacent patterns exhibit similar characteristics, such as nodes 4 and 5, there is substantial variability across the matrix. In general, no individual node has a pattern that closely matches the grand composite in Fig. 2. Here, SOMs provide additional information on what the patterns look like at places that differ from the grand composite. The reference node assignments are generally spatially cohesive, although in some cases, regions that are climatologically unrelated are assigned to the same node. For example, node 2 is the closest match for both north-central Canada and the southwestern coast of North America. While the climate of these regions is quite different, the coldest air available for surface advection lies to the northeast in both cases. It follows that a pattern like the one for node 2 would be associated with cold extremes in both cases, as the SLP anomaly gradient would promote surface wind anomalies with a northeasterly trajectory. This feature is reasonably well captured by the CMIP5 models.

The northwestern portion of the domain is assigned to node 1, which is notably different than the other eight patterns. Here, the coldest air is to the west, rather than

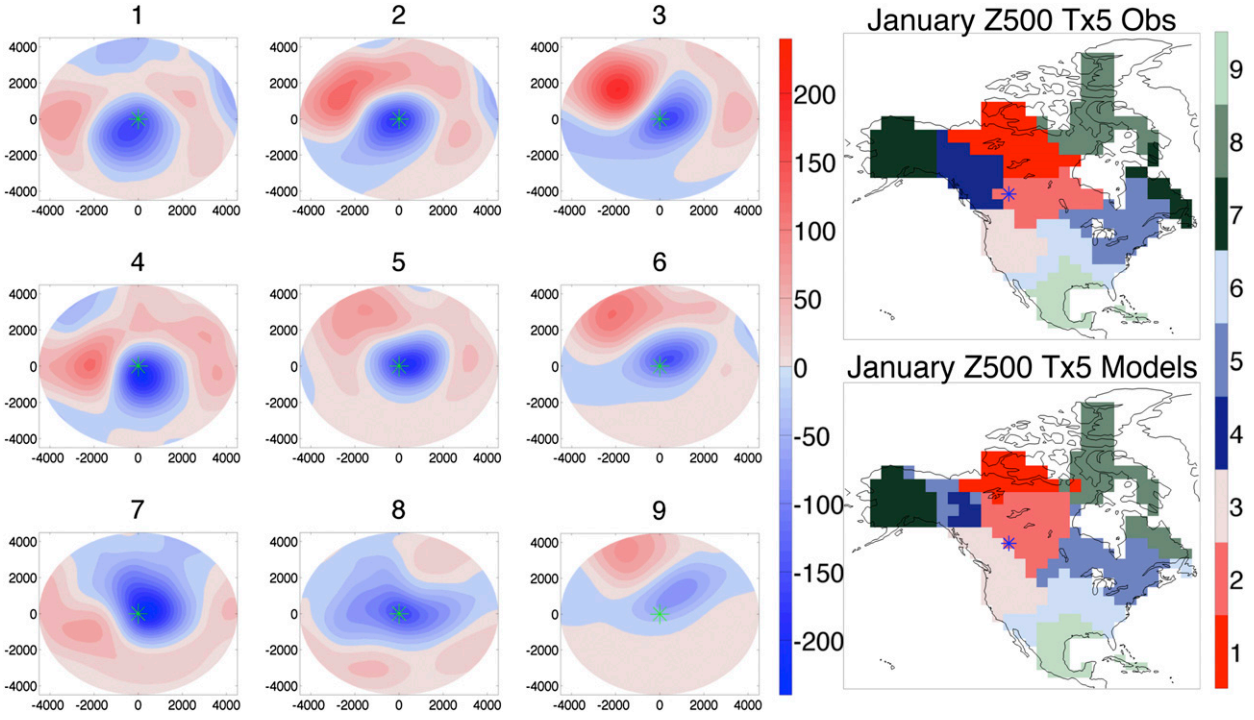

FIG. 7. As in Fig. 5, but for Z500 anomalies (m).

to the north or east, as in the rest of the domain, and the negative SLP anomaly to the north of the grid cell in the node 1 pattern promotes advection of this cold air. Nodes 6, 7, and 8 all show conditions where cold extremes are associated with easterly flow at the surface, and these correspond to places in the western half of the domain, including Alaska, where the coldest air available for surface advection lies to the east. This is captured well in the models with the exception of eastern Alaska and western Canada, where some grid cells that are assigned to node 7 in the reference are assigned to node 5 in the models. Nodes 4 and 5 cover much of the eastern half of the domain in the reference and models, where cold extremes are associated with north and northwesterly winds behind a cyclone and ahead of an anticyclone.

Composite patterns for the mid-Atlantic grid cell highlighted in Fig. 5 are presented in Fig. 6 for reference and the CMIP5 ensemble mean (ENS) (Fig. 6a,b), as well as for a model that closely resembles reference and one that does not (Fig. 6b,c). This grid cell is assigned to node 4 in the reference and node 5 in ENS. Nodes 4 and 5 are similar, in that they both show positive SLP anomalies to the west and negative SLP anomalies to the east, but node 4 has stronger negative anomalies than node 5, and node 5 has positive SLP anomalies to the north. The reference pattern resembles node 4 well, with a pattern correlation of 0.84, suggesting that the SOM assignments are realistic. The ENS pattern also resembles the reference pattern [pattern correlation (PC) of 0.88], with the main difference being that the positive SLP anomaly is stronger and the negative anomaly weaker in ENS compared with reference. This is similar to node 5. Model O strongly resembles the reference pattern (PC of 0.92), while model P differs considerably (PC of 0.50). In model P, the strongest positive SLP anomalies are nearly overhead, while the negative anomalies are much further downstream than in reference. This scenario would promote maximum radiational cooling at night but does not seem realistic for extremely cold maximum temperatures.

Figure 7 shows the results for January Tx5 Z500 patterns in the same format as Fig. 5. All nodes in Fig. 7 show cold extremes associated with negative Z500 anomalies in the vicinity of the grid cell, with positive anomalies upstream and downstream, as in the grand composite. Differences across nodes are primarily in the strength of the anomaly with node 9 showing the weakest central negative anomaly and nodes 4 and 5 the strongest. Another differentiating factor is the orientation of the Z500 wave train. The western US is mostly assigned to node 3, which shows a highly amplified wave train pattern, oriented mostly east–west. Nodes 7 and 8 differ from this pattern with negative anomalies collocated with the anomalously cold air and weaker area of elongated positive anomalies to the south. While the CMIP5 models capture many of

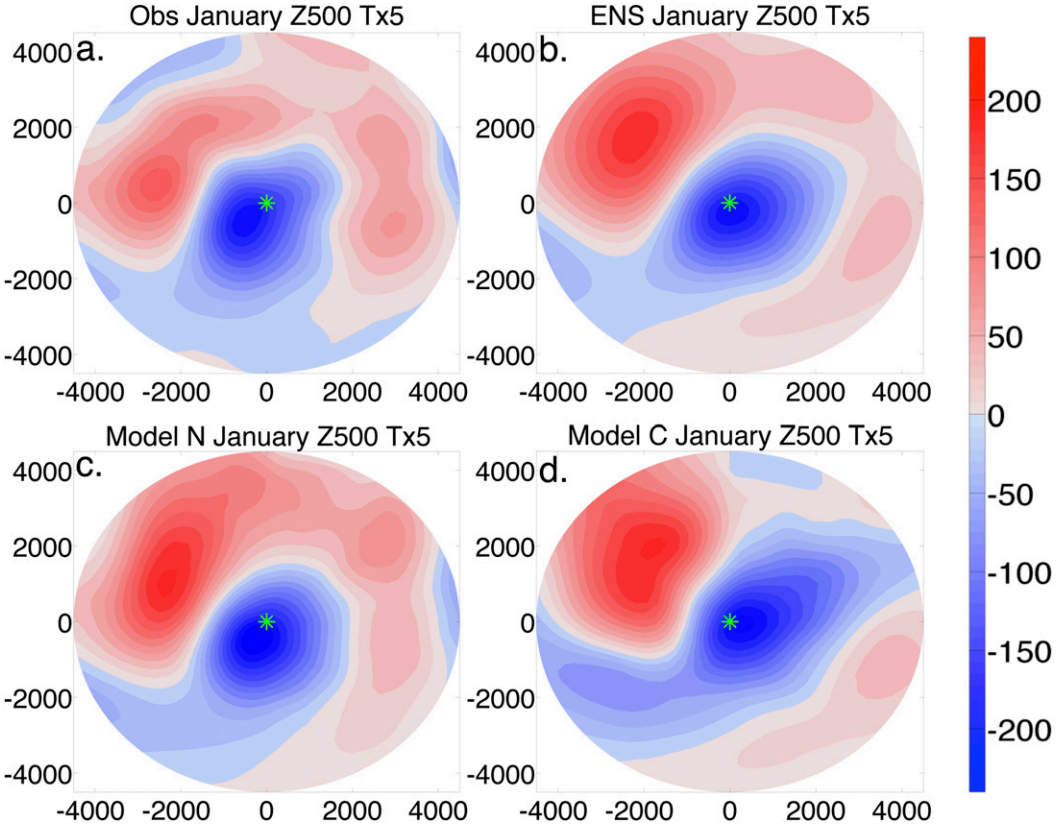

FIG. 8. Examples of Z500 anomaly patterns (m) for January Tx5 days at the western Canada grid cell, highlighted with a blue asterisk in Fig 7. Patterns are from (a) observations, (b) the multimodel ensemble mean, (c) model N, and (d) model C. The green asterisk is the point where the extreme temperatures are occurring.

these features, some notable differences exist. Node 3 is present further north in the models, reducing the area of node 4 assignments over northwest Canada. This discrepancy could be a consequence of the smoother topography in the models, which may require a stronger offshore component to the implied wind anomalies, as in node 3, to promote low temperatures when the barrier effect is weaker. Other differences, like the northward expansion of node 2 in the models, are less dynamically notable, as node 2 and node 1 both have a negative Z500 anomaly near the grid cell, with positive anomalies to the west, north, and east.

Some regions of spatially coherent node assignments for SLP in Fig. 5 also show spatially coherent node assignments for Z500 in Fig. 7. For example, the region assigned to node 1 for SLP is assigned to node 8 for Z500, while Alaska is assigned to node 7 in both cases. This indicates that these regions exhibit distinct characteristic circulation patterns associated with cold temperature extremes near the surface and in the midtroposphere. As a counterexample, the strip of node 2 assignments along the southwestern coast of the domain for SLP is not reflected at Z500. In this region, extreme cold January temperature days can occur with a Z500 pattern that is also associated with cold extremes at locations further inland, while extreme cold days are associated with a different SLP pattern than extreme cold days further inland.

The composite patterns for the grid cell highlighted over western Canada in Fig. 7 are presented in Fig. 8 in the same format as Fig. 6. Both the reference and the ensemble mean patterns are assigned to node 2 with the reference strongly resembling the pattern for this node (PC of 0.96). In ENS, this grid cell is on the transition from node 2 to node 3, and the composite pattern in Fig. 8 also resembles the pattern for node 3 (PC of 0.95). This is illustrative of the expansion of node 3 northward in ENS compared with reference, as even at this grid cell, the Z500 pattern shows features associated with an amplified wave train, as seen in node 3. Model N most closely resembles the reference pattern, with a PC of 0.95; however, the strong positive Z500 anomalies upstream are also similar to ENS. Model C, while sharing some common features, is more indicative of a progressive

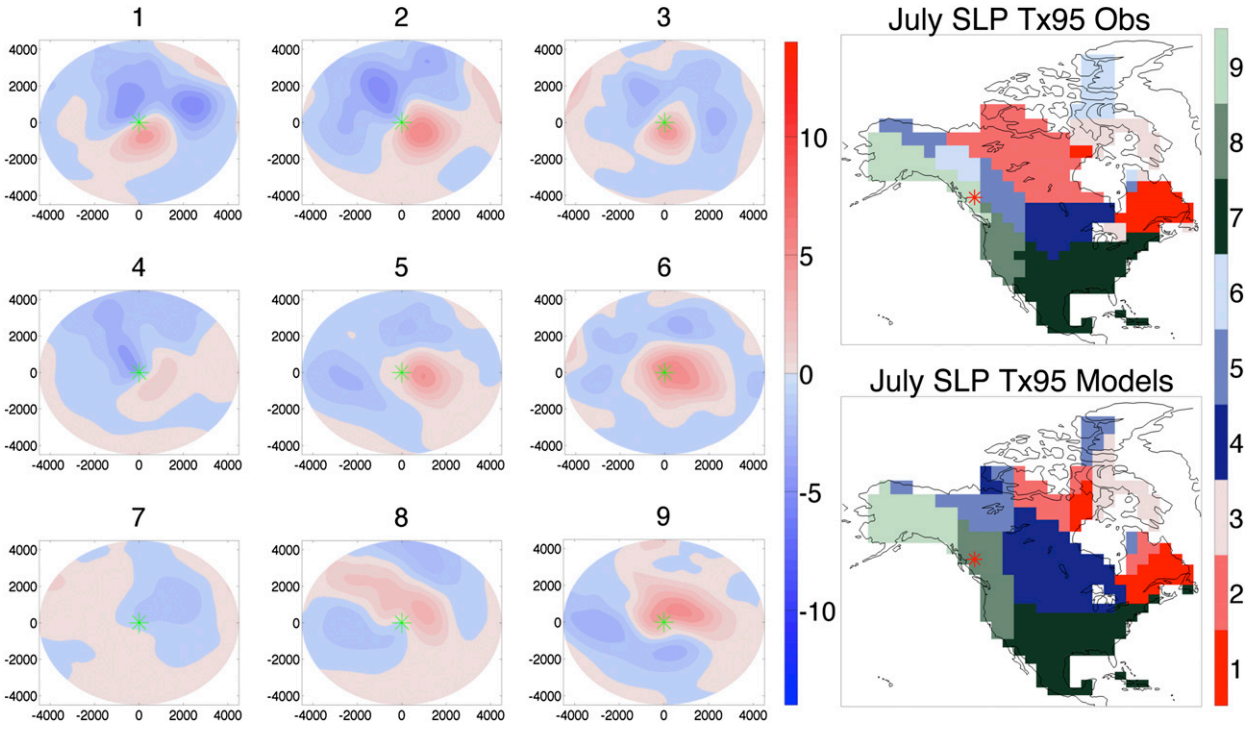

FIG. 9. As in Fig. 5, but for July Tx95.

ridge–trough pattern, as seen in node 3, although the strong positive Z500 anomaly center upstream with weaker center downstream is similar to ENS.

b. *July Tx95*

The SOMs analysis results for July Tx95 SLP patterns are shown in Fig. 9. SLP features associated with extreme warm summer temperatures are subtle over much of the conterminous United States, as indicated by the large area assigned to node 7 in both the reference and models. In the western portion of the domain, extreme warmth is generally associated with meteorological patterns that inhibit marine influence from the Pacific Ocean, as reflected in the patterns for nodes 8 and 9. Here, the SLP gradient is oriented perpendicular to the coast, which inhibits onshore flow at the surface. Such conditions have been shown to be associated with extreme heat along western North America (e.g., Bumbaco et al. 2013; Grotjahn and Faure 2008), and this feature is mostly reproduced by the models. In the northern portion of the domain, southerly advection is strongly associated with warm extremes, as demonstrated by the large area of node 2 assignments. Horizontal temperature gradients are weak over much of the domain during the summer, providing little opportunity for extreme temperatures to be a result of surface advection and inhibiting strong perturbations in the SLP field. This is less true at higher latitudes, where relatively strong positive SLP anomalies to the southeast and negative anomalies to the northwest promote advection of warmer air from lower latitudes. The models do not capture this feature well, with most of the area represented by node 2 in observations being assigned to node 4. Node 4 has some similarities with node 2; however, the anomalies are much weaker. It is possible that many ensemble members do capture the pattern seen in node 2 here, but the anomalies are weakened when computing the ensemble average with other members that do not capture this pattern. The overall weaker agreement between models and observations for July SLP compared with January is consistent with results presented in Fig. 1.

Composites for the grid cell highlighted in Fig. 9 located in western Canada are presented in Fig. 10. The reference pattern is assigned to node 9, while the ENS pattern is assigned to node 8. As discussed above, both patterns inhibit onshore flow, which prevents marine air from moderating surface temperature; however, node 9 has stronger anomalies oriented more east–west, while node 8 has weaker anomalies oriented northwest–southeast. The reference pattern resembles the pattern for node 9 well (PC of 0.80); however, the lower PC coefficient compared with the January examples indicates some level of within-node variability. The ENS pattern assigned to node 8 resembles a hybrid of nodes 8 and 9, with PC coefficients of 0.70 and 0.67, respectively. The key mechanism (inhibition of onshore flow) is

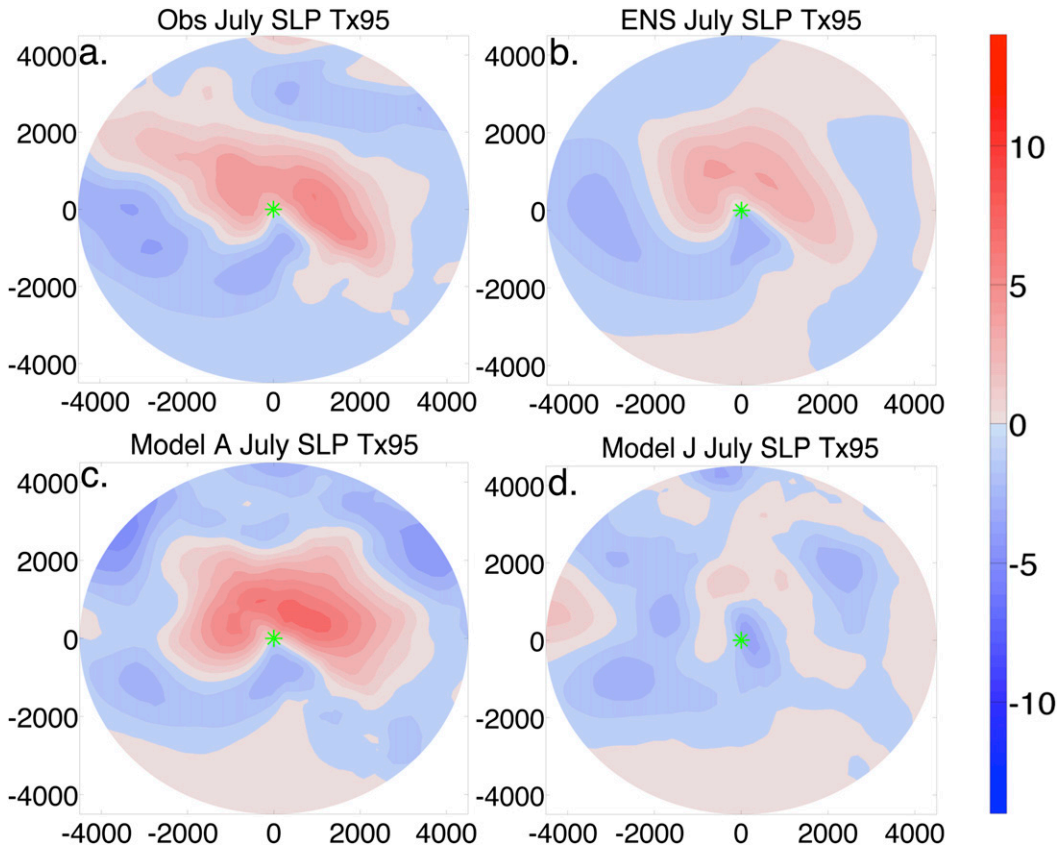

FIG. 10. Examples of Z500 anomaly patterns (in m) for January Tx5 days at the grid cell, highlighted with a red asterisk in Fig 9. Patterns are from (a) observations, (b) the multimodel ensemble mean, (c) model A, and (d) model J. The green asterisk is the point where the extreme temperatures are occurring.

captured by ENS, and the strong similarity between nodes 8 and 9 further suggests that the northward expansion of node 8 for ENS compared with reference may not indicate a fundamental issue with simulating the proper mechanisms. Model A closely resembles the reference pattern (PC of 0.83), with strong similarity to ENS as well, while model J only weakly suggests the same pattern qualitatively but with a PC coefficient of zero. The influence of ensemble members, such as model J, could lead to weaker anomalies in the ensemble mean, nudging the ENS pattern toward the less robust node 8 rather than node 9.

The Z500 patterns corresponding to the Tx95 SLP patterns are shown in Fig. 11. All nodes feature a positive Z500 anomaly centered near the grid cell, with the magnitude of this anomaly increasing from node 1 to node 9. Aside from nodes 1, 2, and 4, there is one primary, closed Z500 positive anomaly region, surrounded by weaker negative anomalies. There is little similarity between the cohesive regions for Z500 and SLP, although some similarities exist, such as the western portions of the domain generally being assigned to nodes 7 and 8, where for the SLP patterns these grid cells were assigned to nodes 8 and 9. This feature is not well reproduced in the CMIP5 ensemble. In general, the weakest anomaly patterns are in the south, as indicated by the region assigned to nodes 1, 2, and 4, while the strongest anomaly patterns are to the north in reference. The models capture nodes 1, 2, and to some extent 4 assignments well, but large differences exist elsewhere. One striking difference is the large area assigned to node 5 for the models, while very few grid cells within that region are assigned to node 5 in the reference. Node 5 does share common features with many other nodes, so the model error here may not represent a fundamental bias in the model ensemble; however, this does indicate that the models are unable to capture the spatial variability in the Z500 patterns across this large portion of the domain.

Figure 12 shows example Z500 composite patterns for the southern U.S. grid cell highlighted in Fig. 11. The Z500 anomalies are weak and small in spatial scale, with a positive Z500 anomaly overhead in the reference. The ENS pattern strongly resembles the reference (PC of

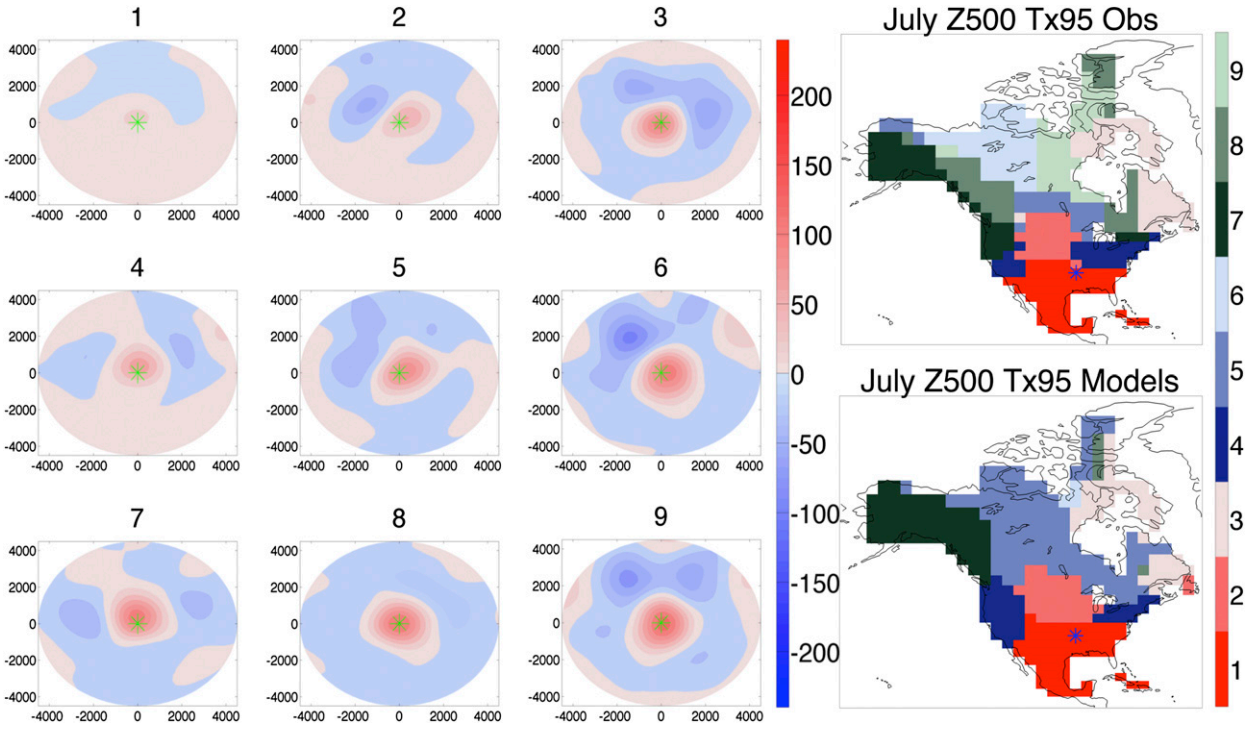

FIG. 11. As in Fig 7, but for July Tx95.

0.86), with a similar positive Z500 anomaly overhead and even with weak positive and negative anomalies beyond this feature. During July, horizontal temperature gradients are negligible in this region, and the atmosphere is more equivalent barotropic than in the winter, resulting in weak and relatively small-scale atmospheric circulation. It should also be noted that other factors, such as anomalously low soil moisture, may also be important for influencing extreme warm temperatures in the summer that are not captured in these composites (Berg et al. 2014; Loikith and Broccoli 2014). Model K captures the local positive Z500 anomaly, while model G has anomalies that are too positive. In both cases, features that are removed from the central positive anomaly center do not match well; however, these features may be relatively unimportant for the occurrence of extreme warm temperatures.

c. *Individual model skill*

Table 2 summarizes the ability of the models to realistically simulate the features identified by the SOMS analysis. For each model and the ensemble, the percentage of the 315 grid cells that that are assigned to the same node as the reference SOMs is computed. In other words, if the pattern at every grid cell in a given model were assigned to the node assignment for that grid cell in the reference SOMs, the value would be 100% in Table 2. Overall, while there is considerable model-to-model variability, the node assignments match the reference SOMs better in January than in July, consistent with the result of Fig. 1. Interestingly, the percentages are higher for SLP than Z500 for all models in July and the majority in January. While overall pattern-to-pattern disagreement is larger for SLP than Z500, the spatial variability as captured by the SOMs is better. This is also true for the ensemble mean.

Models that perform well for one type of extreme or variable do not necessarily perform well for others. For example, model D shows relatively high percentages for January but low percentages for July. Model E shows low percentages for January Z500 and high percentages for January SLP. In some cases, the model performance is systematic, such as the systematically low percentages for model J and high percentages for model K. In all four cases, the ensemble mean shows percentages that are as high or are higher than the best models.

5. Summary and concluding remarks

The results from a systematic evaluation of a suite of 17 state-of-the-art climate models from the CMIP5 database indicate that most models capture the key features of the primary atmospheric circulation patterns associated with extreme temperature days. Substantial

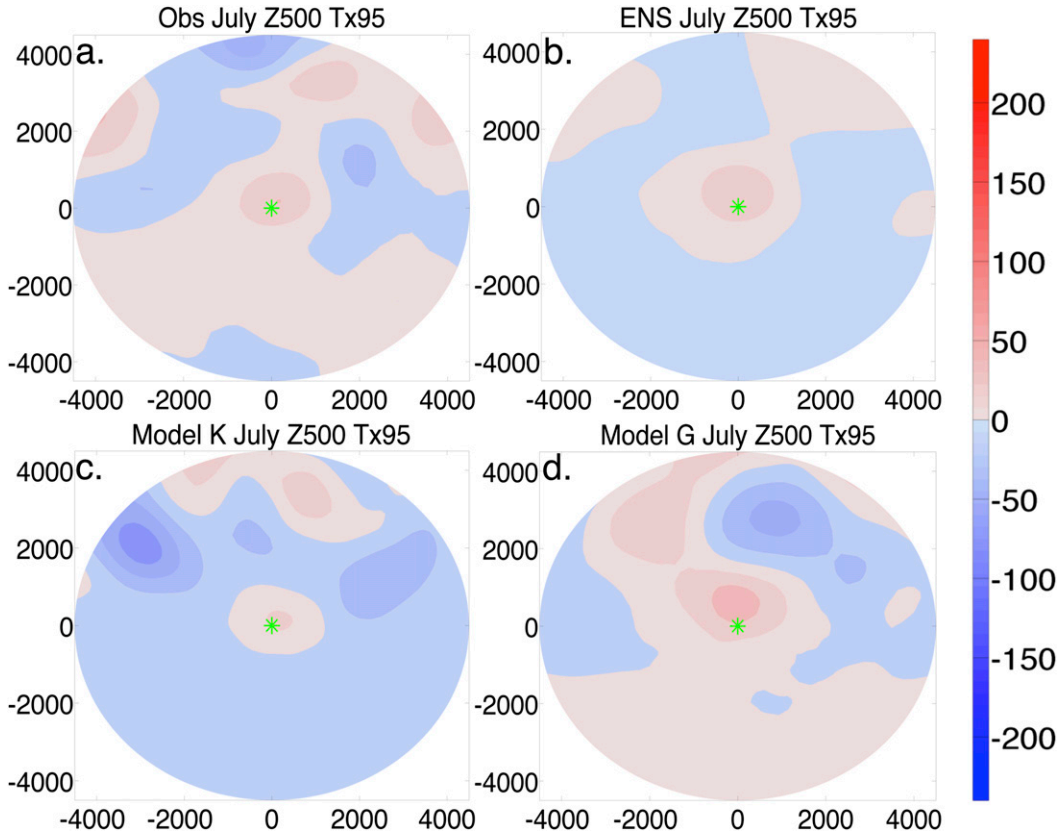

FIG. 12. Examples of Z500 anomaly patterns (in m) for July Tx95 days at the Alaska grid cell, highlighted with a blue asterisk in Fig 11. Patterns are from (a) observations, (b) the multimodel ensemble mean, (c) model K, and (d), model G. The green asterisk is the point where the extreme temperatures are occurring.

variability amongst models within the 17-member suite does suggest that some models may be better suited for analysis of future changes in extremes than others. Model agreement tends to be weakest in regions where complex topography, influence from large bodies of water, or other strong asymmetries are present. Overall, the comparisons presented here suggest that this suite of models may be well suited for analysis of the extent to which, if any, changes in circulation patterns will have on future temperature extremes.

Winter patterns are more realistically simulated than summer patterns, in general. The less prominent role of large-scale circulation and advection for the occurrence of temperature extremes in the summer and the more prominent role of smaller-scale processes and land–atmosphere interactions may explain the lower model fidelity in the summer. Patterns at Z500 are often more realistically simulated compared with SLP patterns, where variability is inherently greater. When the suite of models is considered in the form of a multimodel ensemble, results agree better with observations than results from most individual model members, suggesting that patterns of model error may be random rather than systematic across models.

The generally realistic representation of these patterns in the models used in this work allows for some confidence to be placed in future simulations of these important mechanisms in global warming simulations. To gain a comprehensive understanding of the mechanisms associated with projected changes in extreme temperature days at the end of the twenty-first century using global warming simulations requires analysis of known, important mechanisms for the occurrence of extreme temperature events in future simulations. The relationship between changes in extremes and circulation likely falls on a spectrum of possibilities. One end of this spectrum would be a scenario where extremes only change as a result of a shift in the mean temperature toward warmer conditions. In this scenario, circulation patterns would likely remain the same in a future climate; only surface temperatures will be warmer. At the other end of the spectrum would be a scenario where all changes in temperature extremes result from changes in atmospheric circulation patterns relative to those

TABLE 2. The percentage of grid cells that share the same SOM assignment as the reference SOMs for all GCMs, labeled by their corresponding letter, and the multimodel ensemble (ENS). January values are for Tx5 and July for Tx95.

Model	January Z500	January SLP	July Z500	July SLP
A	59	64	45	59
B	59	67	47	57
C	44	48	51	56
D	64	67	37	54
E	35	63	34	46
F	58	57	48	57
G	65	60	50	51
H	48	61	50	51
I	59	61	46	55
J	33	45	28	34
K	64	76	48	56
L	54	63	43	59
M	65	62	55	59
N	59	64	46	55
O	59	63	50	50
P	60	49	46	52
Q	53	58	40	45
ENS	70	77	51	62

patterns identified in this work and in LB12. To fully understand where reality lies and how variable the results are regionally, future work should focus on examining these models for systematic changes in circulation under future global warming conditions.

Acknowledgments. This study was supported by the Office of Science (BER), U.S. Department of Energy, Award DE-SC0005467. This work was done as a private venture and not in the author's capacity as an employee of the Jet Propulsion Laboratory, California Institute of Technology.

REFERENCES

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.,* **111,** D05109, 10.1029/2005JD006290.

Beniston, M., 2004: The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations. *Geophys. Res. Lett.,* **31,** L02202, doi:10.1029/2003GL018857.

Berg, A., B. R. Lintner, K. L. Findell, S. Malyshev, P. C. Loikith, and P. Gentine, 2014: Impact of soil moisture–atmosphere interactions on the surface temperature distribution. *J. Climate,* **27,** 7976–7993, doi:10.1175/JCLI-D-13-00591.1.

Brown, P. J., R. S. Bradley, and F. T. Keimig, 2010: Changes in extreme climate indices for the northeastern United States, 1870–2005. *J. Climate,* **23,** 6555–6572, doi:10.1175/2010JCLI3363.1.

Bumbaco, K. A., K. D. Dello, and N. A. Bond, 2013: History of Pacific Northwest heat waves: Synoptic pattern and trends. *J. Appl. Meteor. Climatol.,* **52,** 1618–1631, doi:10.1175/JAMC-D-12-094.1.

Caesar, J., L. Alexander, and R. Vose, 2006: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *J. Geophys. Res.,* **111,** D05101, doi:10.1029/2005JD006280.

Cassano, J. J., P. Uotila, and A. Lynch, 2006: Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries. Part 1: Arctic. *Int. J. Climatol.,* **26,** 1027–1049, doi:10.1002/joc.1306.

Cavazos, T., 2000: Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *J. Climate,* **13,** 1718–1732, doi:10.1175/1520-0442(2000)013<1718:USOMTI>2.0.CO;2.

Christidis, N., P. A. Stott, S. Brown, G. Hegerl, and J. Caesar, 2005: Detection of changes in temperature extremes during the second half of the 20th century. *Geophys. Res. Lett.,* **32,** L20716, doi:10.1029/2005GL023885.

——, ——, and ——, 2011: The role of human activity in the recent warming of extremely warm daytime temperatures. *J. Climate,* **24,** 1922–1930, doi:10.1175/2011JCLI4150.1.

Dole, R., and Coauthors, 2011: Was there a basis for anticipating the 2010 Russian heat wave? *Geophys. Res. Lett.,* **38,** L06702, doi:10.1029/2010GL046582.

Donat, M. G., and L. V. Alexander, 2012: The shifting probability distribution of global daytime and night-time temperatures. *Geophys. Res. Lett.,* **39,** L14707, doi:10.1029/2012GL052459.

Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. M. G. Klein Tank, and T. Peterson, 2002: Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Res.,* **19,** 193–212, doi:10.3354/cr019193.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *J. Geophys. Res.,* **113,** D06104, doi:10.1029/2007JD008972.

Griffiths, M. L., and R. S. Bradley, 2007: Variations of twentieth-century temperature and precipitation extreme indicators in the northeast United States. *J. Climate,* **20,** 5401–5417, doi:10.1175/2007JCLI1594.1.

Grotjahn, R., and G. Faure, 2008: Composite predictor maps of extraordinary weather events in the Sacramento, California region. *Wea. Forecasting,* **23,** 313–335, doi:10.1175/2007WAF2006055.1.

Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res.,* **22,** 13–26, doi:10.3354/cr022013.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kharin, V. V., and F. W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *J. Climate,* **13,** 3760–3788, doi:10.1175/1520-0442(2000)013<3760:CITEIA>2.0.CO;2.

——, and ——, 2005: Estimating extremes in transient climate change simulations. *J. Climate,* **18,** 1156–1173, doi:10.1175/JCLI3320.1.

——, ——, X. Zhange, and G. C. Hegerl, 2007: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *J. Climate,* **20,** 1419–1444, doi:10.1175/JCLI4066.1.

Kodra, E., K. Steinhaeuser, and A. R. Ganguly, 2011: Persisting cold extremes under 21st-century warming scenarios. *Geophys. Res. Lett.,* **38,** L08705, doi:10.1029/2011GL047103.

Liu, Y., R. H. Weisberg, and C. N. K. Mooers, 2006: Performance evaluation of the self-organizing map for feature extraction. *J. Geophys. Res.,* **111,** C05018, doi:10.1029/2005JC003117.

Loikith, P. C., and A. J. Broccoli, 2012: Characteristics of observed atmospheric circulation patterns associated with temperature extremes over North America. *J. Climate,* **25,** 7266–7281, doi:10.1175/JCLI-D-11-00709.1.

——, and ——, 2014: The influence of recurrent modes of climate variability on the occurrence of winter and summer extreme temperatures over North America. *J. Climate,* **27,** 1600–1618, doi:10.1175/JCLI-D-13-00068.1.

Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science,* **305,** 994–997, doi:10.1126/science.1098704.

——, and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–846. [Available online at http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter10.pdf.]

——, C. Tebaldi, G. Walton, D. Easterling, and L. McDaniel, 2009: Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S. *Geophys. Res. Lett.,* **36,** L23701, doi:10.1029/2009GL040736.

Morak, S., G. C. Hegerl, and J. Kenyon, 2011: Detectable regional changes in the number of warm nights. *Geophys. Res. Lett.,* **38,** L17703, doi:10.1029/2011GL048531.

Otto, F. E. L., N. Massey, G. J. van Oldenborgh, R. G. Jones, and M. R. Allen, 2012: Reconciling two approaches to attribution of the 2010 Russian heat wave. *Geophys. Res. Lett.,* **39,** L04702, doi:10.1029/2011GL050422.

Rahmstorf, S., and D. Coumou, 2011: Increase of extreme events in a warming world. *Proc. Natl. Acad. Sci. USA,* **108,** 17 905–17 909, doi:10.1073/pnas.1101766108.

Rowe, C. M., and L. E. Derry, 2012: Trends in record-breaking temperatures for the conterminous United States. *Geophys. Res. Lett.,* **39,** L16703, doi:10.1029/2012GL052775.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature,* **427,** 332–336, doi:10.1038/nature02300.

Sheridan, S. C., and C. C. Lee, 2011: The self-organizing map in synoptic climatological research. *Prog. Phys. Geogr.,* **35,** 109–119, doi:10.1177/0309133310397582.

Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature,* **432,** 610–614, doi:10.1038/nature03089.

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, doi:10.1175/BAMS-D-11-00094.1.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, and G. A. Meehl, 2006: Going to the extremes: An intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change,* **79,** 185–211, doi:10.1007/s10584-006-9051-4.

Vavrus, S., J. E. Walsh, W. L. Chapman, and D. Portis, 2006: The behavior of extreme cold air outbreaks under greenhouse warming. *Int. J. Climatol.,* **26,** 1133–1147, doi:10.1002/joc.1301.

Vesanto, J., J. Himberg, E. Alhoniemi, and J. Parhankangas, 2000: SOM Toolbox for Matlab 5. SOM Toolbox Team Tech. Rep. A57, Helsinki University of Technology, Helsinki, Finland, 59 pp. [Available online at http://www.cis.hut.fi/somtoolbox/package/papers/techrep.pdf.]

Zwiers, F. W., X. Zhang, and Y. Feng, 2011: Anthropogenic influence on long return period daily temperature extremes at regional scales. *J. Climate,* **24,** 881–892, doi:10.1175/2010JCLI3908.1.