Received: 18 February 2021 | Revised: 30 August 2021 | Accepted: 5 September 2021 | Published on: 1 November 2021

DOI: 10.1002/joc.7406

RESEARCH ARTICLE

International Journal of Climatology RMetS

Statistical climate model downscaling for impact projections in the Midwest United States

Andrew D. Polasky[1] | **Jenni L. Evans**[1,2] | **Jose D. Fuentes**[1] | **Holly L. Hamilton**[3]

[1]Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania, USA

[2]Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA

[3]Turks and Caicos Islands Airport Authority, Providenciales, Turks and Caicos Islands

Correspondence
Jenni L. Evans, Department of Meteorology and Atmospheric Science, The Pennsylvania State University, 503 Walker Building, University Park, PA 16802, USA.
Email: jle7@psu.edu

Funding information
National Science Foundation, Grant/Award Number: 1639342

Abstract

For future climate projections to be useful they must be actionable at the local level. In this study, we develop daily temperature and precipitation climate scenarios suitable for use in projections of drought, energy use, water use, and crop production. We investigate the magnitude of future changes to air temperature and precipitation in the Midwest United States in response to three future climate change scenarios. Results are used to assess changes to incidence of precipitation extremes and human comfort (using heat index) associated with the anticipated climate changes in the region. We use self-organizing maps and random forest based techniques to generate daily realizations of temperature and precipitation for 279 weather stations in a region centred on Illinois. We determine that the random forest model performs best for maximum and minimum temperatures, while the self-organizing map performs best for precipitation. Using nine models from the Coupled Model Inter-Comparison Project Phase 5, downscaled daily temperature and precipitation values are generated for low, moderate, and high greenhouse gas emissions scenarios for historical and future periods. Based on recent trends, we focus our results on the high emissions scenario, and show an average increase of 4.3°C in maximum daily air temperature across the region for the 2071–2100 period. Precipitation decreases by up to 15% in the southern half of the study region, with a similar percentage increase in the northern half of the region. The regional environmental changes result in an increase of 5.8° in average summer heat index, and increase of 48% in the number of days likely to produce extreme heat, and a decrease in the average value of the standardized precipitation and evapotranspiration index of 1.9 (indicating increased drought) across the region by 2100.

KEYWORDS

climate change, downscaling, drought, heat index, random forest, self-organizing maps

1 | INTRODUCTION

Climate is experienced locally as weather, and as accumulated weather anomalies that result in droughts, heat waves, and cold spells. Climate change will cause disruptions in the patterns of these events, through modifications in the synoptic dynamic and thermodynamic conditions that govern them. Projections of local weather events are important for estimating the costs of climate change, and for communicating the impacts of climate change to local communities.

In the Midwest United States, changes in local weather and climate will be consequential for public, private, and commercial concerns. In particular, agriculture, an approximately $110 billion industry in the region (USDA, 2017), is vulnerable to shifts in temperature and precipitation. The majority of agriculture in the Midwest is rain-fed, leaving it vulnerable to changes in the amount and timing of precipitation and increased evaporation from increased temperature. As a result, much of the variability in historical crop production can be attributed to weather (Liang *et al*., 2017). Thus, being able to predict changes in the frequency, amount, and seasonality of precipitation in the Midwest and similar climate regimes is important for projecting changes to crop yields, and for guiding adaptation decisions.

Agriculture in the Midwest is dominated by corn and soybean production, which occupy 75% of the land under cultivation (Hatfield *et al*., 2018). Climate change will alter the yield and viability of these crops in unpredictable ways: increasing the length of the growing season may increase yields, while decreases in precipitation and increases in heat stress may decrease yields (Angel *et al*., 2018). Studies of the effects of increased temperatures on soy and corn crop yields have found that at average daily temperatures above a threshold of 29°C for corn and 30°C for soy, yields begin to decline rapidly (Schlenker and Roberts, 2009; Hatfield *et al*., 2011). These temperatures currently only occur a few times per year in most of the region, but the frequency will likely increase with increasing temperatures projected for the coming decades (Vose *et al*., 2017). Effective adaptation and long-term planning for a viable agricultural sector is therefore dependent on reliable and localized projections of climate change.

Past events, notably the Dust Bowl of the 1930s, demonstrate the vulnerability of the region to climatic change. The period of the 1930s was significantly warmer and somewhat drier than more recent decades leading to severe droughts in the region (Rosenberg *et al*., 1993). The Dust Bowl climate is well within the range suggested by the global climate models under the most likely scenarios (Hayhoe *et al*., 2017), and would cause a significant reduction in agricultural yields in the region (Glotter and Elliott, 2016).

General circulation models (GCMs), running on grid scales on the order of 100 km, do not capture many of the sub-grid scale processes that drive local weather (Raäisaänen, 2007; Maraun *et al*., 2010; Radić and Clarke, 2011; Taylor *et al*., 2012). This is especially true

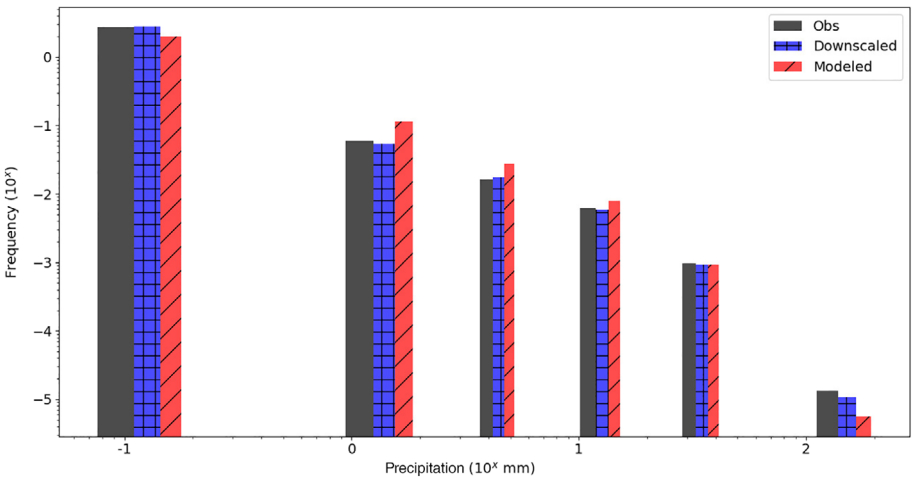

FIGURE 1 Log–log plot of the historical distribution of daily precipitation for an example station located at 42.6° N, 87.8° W; station observations are plotted in blue. Data from the Canadian Centre for Modelling and Analysis (CanESM2) climate model are chosen to illustrate the impact of downscaling (cross hatch) compared to gridpoint (diagonal hatch) fields. Downscaling results for other climate models are consistent with this example. The bins are chosen to match the standard categories used in model output statistics (MOS) for operational forecasts. The CanESM2 climate model produces too few days with heavy precipitation (>5 mm·day^{-1}) and too few 'no precipitation' days, while over-representing days with intermediate precipitation [Colour figure can be viewed at wileyonlinelibrary.com]

for precipitation, where GCMs tend to produce too few examples of extreme precipitation and too few days with no precipitation, while overproducing moderate precipitation days (Stephens *et al*., 2010; Koutroulis *et al*., 2016). Downscaling from the climate model gridded values can correct for this bias, creating a more accurate distribution of values (Figure 1). Accurate representation of the distribution of precipitation events is critical for predicting incidence of the high impact weather experienced at individual locations [such as flooding (Xiao *et al*., 2013), convective initiation (Gerken *et al*., 2018), and drought (Lobell *et al*., 2014)] in a changing climate. Statistical downscaling techniques provide one solution to the shortcomings of low-resolution climate models by exploiting the relationships between coarse resolution reanalyses and local observations in the historical climate record. Several downscaling techniques have been used in the past (Wilby *et al*., 2004), ranging from simple regression analyses (e.g., Murphy, 1999) to stochastic weather generators (e.g., Wilks, 1999; Kilsby *et al*., 2007). Statistical downscaling provides a computationally efficient bridge between the synoptic and local scales.

Statistical downscaling relies on the assumptions that the variability in the local values of the downscaled variable is explainable by the synoptic-scale features, and that the relationships between those synoptic features and downscaled variable will be maintained over the range of climates being considered (Wilby *et al*., 2004). Without the first assumption, downscaling will not produce any useful outputs, but will, at best, recreate the variability of the historical observations. If the second assumption is not valid, the downscaling method will fail to capture the full effects of the changes to the local climate in the downscaled values. Statistical downscaling can also struggle to predict rare and extreme events, especially when the underlying statistics change. These concerns can be mitigated by careful testing of the downscaling results over time periods independent to those used in training the models. We discuss in Section 2.5 the methods used to test and verify the downscaling results. Despite these challenges, statistical downscaling methods can be used to bridge the gap between the information provided by the GCMs and the local changes that will impact a range of applications, from agriculture to household energy use (e.g., Salathé *et al*., 2007; Sinha *et al*., 2018).

Using self-organizing maps (SOM) (Kohonen, 1990; Hewitson and Crane, 2006) and random forest (RF) (Breiman, 2001) based statistical downscaling techniques, we examine the projected impacts of climate change on the Midwest. With these downscaling methods, we produce downscaled projections of multiple climate scenarios suitable for use in impact studies, including messaging to residents of the region about climate change, and crop yield modelling of agricultural productivity in future climates (Watkins *et al*., 2019). This research is done as part of broader project studying changes and impacts of household consumption in the Midwest (Watkins *et al*., 2019) including a household level intervention study and a role-playing game for educating on climate change impacts (Agusdinata and Lukosch, 2019). The results from this work will be used to craft messages to households in this region, focusing

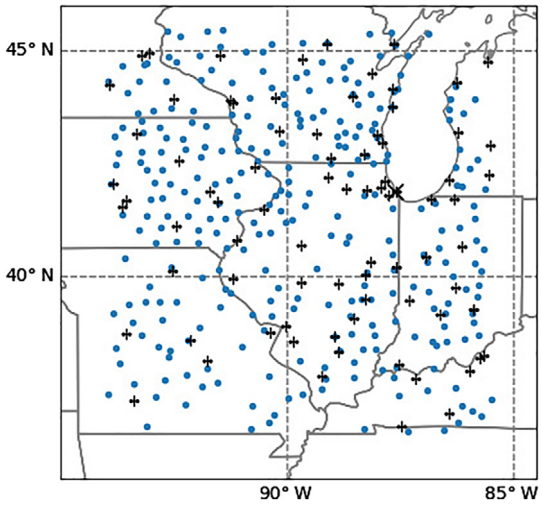

FIGURE 2 Locations of the stations used for the downscaling analysis. In total, 279 stations are used, spread across eight states. Stations with observations missing for more than 10% of the days between 1948 and 2005 were excluded. In addition to the main set of GHCN stations (dots), GSOD stations with dew point temperatures are used for calculating heat index (crosses). Chicago, the largest city in the region, is marked with a black 'X' [Colour figure can be viewed at wileyonlinelibrary.com]

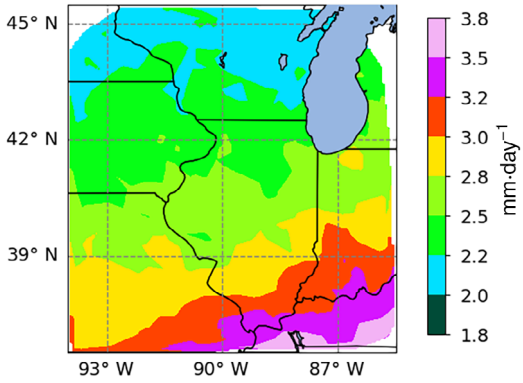

FIGURE 3 Average daily precipitation for observing stations across the region for the model training period (HIST). The southern and eastern portions of the region receive close to twice the average annual rainfall of the northern portion [Colour figure can be viewed at wileyonlinelibrary.com]

on the specific impacts to the region. These messages will be given to homeowners in the study, to give them insights on the impacts of climate change on their local area. By providing localized projections of climate change impact, the work presented here will help to focus messaging in these studies to impacts most relatable to the participants, with a broader goal of determining what types of messages are most useful for shifting consumption patterns of energy, water, and food.

Our study area (Figure 2) is centred on Illinois and includes portions of Indiana, Iowa, Kentucky, Missouri, Michigan Minnesota, and Wisconsin. This region is predominantly plains and small, rolling hills, with some areas of greater topographical variety, especially in southern Indiana and Illinois, and the Driftless Area of northwestern Illinois, southwestern Wisconsin, and northeastern Iowa.

The climate in this region is typically continental with cold winters, warm summers, and frequent rapid fluctuations in temperature, humidity, cloudiness, and wind direction (Changnon *et al*., 2004). Lake Michigan lies northeast of Illinois and tends to increase cloudiness and suppress summer precipitation in northeastern Illinois (Changnon *et al*., 2004). Average precipitation generally increases moving south across the region, with close to 150% as much precipitation at the southern edge than at the northern (Figure 3). Buildings, parking lots, roads and industrial activities in major cities render the urban climate noticeably different than that of the surrounding rural areas. Urban areas may also enhance summertime precipitation downwind of the city (Shepherd *et al*., 2002) leading to changes in humidity, cloudiness, wind speeds and directions (Changnon *et al*., 2004).

Except in summer, the polar jet stream is typically located near or over the region. The polar jet is responsible for the generation and steering of mid-latitude cyclones that move across the region. These cyclone systems account for a significant portion of the cool season precipitation in the region (Changnon *et al*., 2004). In the summer months, an estimated 30–70% of precipitation is driven by mesoscale convective systems (Fritsch *et al*., 1986).

Climate change will impact the Midwest under a range of global climate scenarios. In this study, we use statistical downscaling methods to explore the changes in temperature and precipitation in the Midwest, and the changes in drought and heat index (HI) resulting from those changes. Quantifying these changes will allow greater understanding of the impacts and costs of climate change on the Midwest, and support future research into the downstream impacts of climate change, especially on human health and crop yields.

The specific objectives of this research are:

- Develop machine learning-based downscaling models to determine impacts of atmospheric warming on regional weather and climatic conditions;
- Estimate uncertainties associated with climate change scenarios on downscaled climate projections using climate models and three greenhouse gas (GHG) emissions scenarios; and
- Use metrics such as human comfort and changes to the growing season to deduce the impacts of future climate projections on local populations.

2 | DATA AND METHODS

Three sources of data are used for the downscaling analysis: historical weather station observations, reanalysis data on regular grids, and climate model simulations. Historical station data for daily maximum temperature (Tmax), daily minimum temperature (Tmin), daily precipitation (PRCP) provide the predictands for the downscaling models. The predictors are drawn from the National Centers for Environmental Prediction (NCEP) climate reanalyses. The set of candidate predictors are chosen to capture the seasonal and regional drivers governing the weather in the Midwest. To avoid redundancy, dimension reduction approaches are used to reduce a large set of candidate predictors to a smaller set of independent predictors (Section 2.3.1). Climate model outputs from Coupled Model Inter-Comparison Project Phase 5 (CMIP5) are then used for projecting future climate scenarios. Combined with the SOM and RF methods, this creates a tune-able, multivariate approach that better captures the true distribution of outcomes across changing underlying climate conditions. By evaluating the results of the downscaling season by season, we show that these methods respond to shifts in the underlying climate, and produce realistic distributions of the station level observations.

TABLE 1 Final set of predictors used in the downscaling models. The variables are daily averages from the CMIP5 models. See Section 2.3.1 for discussion of the variable selection method

Variable	Pressure level (hPa)
Temperature	850
Temperature anomaly	850
Geopotential height	500
Geopotential height anomaly	500
Meridional wind	700
Zonal wind	700
Relative humidity	850
Sine of day of year	N/A

2.1 | Training data

The downscaling models are developed using station and reanalysis data from the 1948–2005 (hereafter, TRAIN) period. Station data is taken from the Global Historical Climatology Network (GHCN) from the National Centers for Environmental Information (NCEI, formerly the National Climatic Data Center) (Menne *et al*., 2012). Stations with records for at least 90% of the days in TRAIN are used, giving a set of 279 stations across Illinois and the neighbouring states (Figure 2) for the downscaling procedure. To ensure that days missing from the observed data did not bias the results, the reanalysis temperatures for each station on the days with observations and the days without observations were compared. Differences in the distributions of temperature were not found to be statistically significant for any of the stations, indicating that there was not a consistent bias in the missing days of observation.

We use data from the NCEP reanalysis 2 product (Kalnay *et al*., 1996), available at a 2.5° × 2.5° grids similar to the native resolution of the CMIP5 models used for the downscaling analysis (Table 1). In preparation for use in the downscaling, the reanalyses and the CMIP5 data are regridded from their native resolutions, which range from 1° to 2.5°, to a consistent spatial resolution of 2° × 2°. For each individual station, the training dataset for downscaling is comprised of an area approximately 1,000 km on a side (5 × 5 GCM grid points), and centred on the nearest grid point to each station. The gridded values for each variable are then standardized as Z-scores, using the mean and standard deviation from the TRAIN period, and multiplied by the cosine of the latitude to avoid distortions caused by the size of the grid boxes across the domain. The values for each variable are then combined into a 25-element vector. The vectors for each predictor variable are concatenated to be passed into the downscaling model as a single vector.

TABLE 2 The nine CMIP5 climate models used to construct the downscaled climate projections

Model	Centre
ACCESS1-0	Australian Weather and Climate Research, Australia
CanESM2	Canadian Centre for Modelling and Analysis, Canada
CMCC-CM	Centro Euro-Mediterraneo sui Cambianmenti Climatici (CMCC), Italy
CMCC-CMS	CMCC, Italy
CNRM-CM5	National Centre for Meteorological Research, France
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory (GFDL), USA
GFDL-ESM2M	GFDL, USA
MIROC-ESM-CHEM	University of Tokyo, Japan
MPI-ESM-LR	Max Plank Institute for Meteorology, Germany

2.2 | Future climate scenarios

Representative Concentration Pathways (RCPs) scenarios 2.6, 4.5, and 8.5 for nine CMIP5 models (from a total of seven modelling centres, Table 2) are used to explore the potential range of local climate changes in the region. As with the reanalysis data, the CMIP5 data is regridded to 2° × 2° for consistency, then standardized as Z-scores for use in the downscaling models.

The RCP 2.6, 4.5, and 8.5 scenarios are temporal trajectories of GHG emissions, developed to represent a range of possible pathways for climate change outcomes (van Vuuren *et al*., 2011). The labels 2.6, 4.5, and 8.5 refer to the increase in radiative forcing in W m^{-2} in 2100 compared to the pre-industrial forcing. The scenarios describe differing levels of anthropogenic emissions to achieve these different forcings. The RCP 2.6 scenario—an increase of 2.6 W m^{-2} of radiative forcing by 2100—requires rapid reductions in GHG emissions, with uptake exceeding emissions by the end of the century. The RCP 4.5 scenario corresponds to reductions in GHG emissions leading to a stabilization of atmospheric GHGs by about 2060. The RCP 8.5 scenario, which most closely resembles the path we have been on since 2005, has emissions continuing to increase through 2100 (van Vuuren *et al*., 2011).

From the CMIP5 data, we produce a set of daily downscaled values for each station for four time periods: 1976–2005 (HIST), 2006–2015 (TEST), 2021–2050 (NEAR), and 2071–2100 (FAR). HIST is used to train and test the downscaling methods, and provide a baseline for changes in future periods. TEST is used to validate the trained models, and the NEAR and FAR periods are used to project the daily air temperature and precipitation associated with different climate outcomes in the three RCP scenarios. This gives a total of two historical downscaling realizations (HIST and TEST), and six climate change projections (three RCP scenarios for each of NEAR and FAR).

2.3 | Candidate methods

Several downscaling methods are considered, including SOMs, RFs, multiple linear regression, and quantile

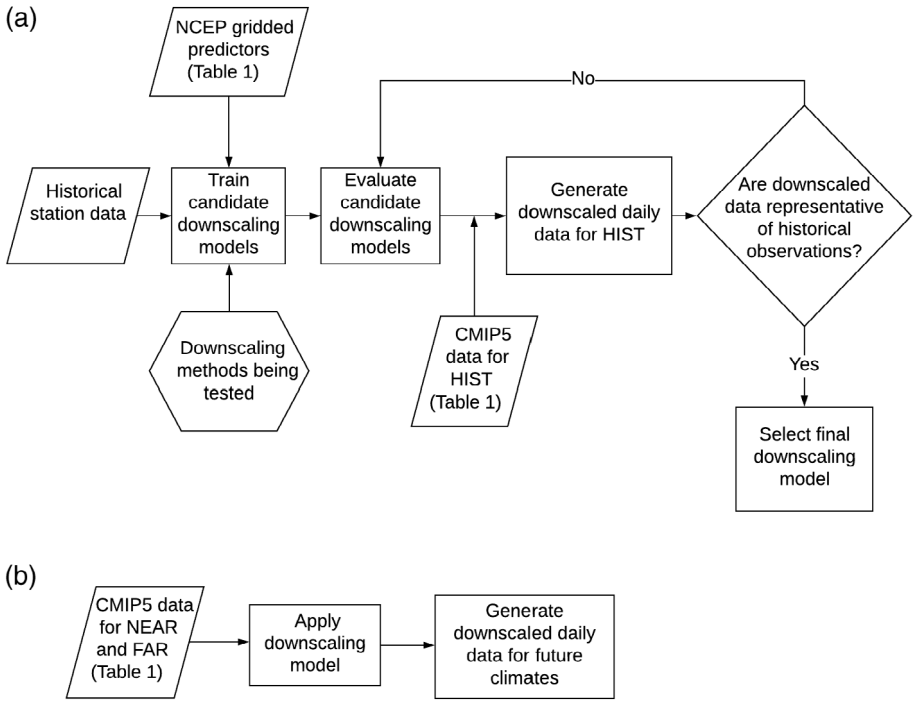

FIGURE 4 Flowchart depicting (a) the development of the downscaling models; and (b) their application to derive downscaled scenarios. The station level observations (Obs) are used along with the NCEP reanalysis data to train models for each of the downscaling methods (Section 2.3). These models are then evaluated on the CMIP5 data for the HIST period, and the outputs are compared to the historical observations at each station. Based on the performance on the HIST period, the final downscaling model is selected for use in the future periods. This process is repeated to develop a separate downscaling model for each of the variables of interest (Tmax, Tmin, PRCP). Using these downscaling models and the CMIP5 data for the NEAR and FAR periods, downscaled data are produced for future climate scenarios (b)

mapping. The downscaling procedure follows the framework shown in Figure 4. NCEP reanalysis data are standardized, then used along with the observed station data to train each of the downscaling models (Figure 4a). The trained downscaling models are then evaluated on the standardized CMIP5 data to generate daily time series for the HIST, TEST, NEAR, and FAR periods (Figure 4b). This approach provides a framework that can be adapted to new methods, variables, and locations.

2.3.1 | Predictor selection

A candidate set of predictors were chosen to capture the elements of the synoptic environment related to local temperature and precipitation variability. This test set included temperature, humidity, wind components, and geopotential height (to capture impacts of synoptic systems). Using dimension reduction, the final set of predictor variables (Table 1) are selected from this candidate set.

The final set of predictors is obtained using sliced inverse regression (SIR), (Li, 1991) to identify the variables contributing with the largest contribution to the variance of the input model data and the downscaling target. This method removes redundant variables from the set of predictors, and is able to capture nonlinear relationships between the predictors. Based on the SIR analysis we select the subset of variables (and pressure levels for each variable) that had the largest effect on the temperature and precipitation. These are used as the predictors in the downscaling model. We also used the dimension reduction approach to examine the importance of anomaly variables, calculated as the difference between the individual value of the variable and the period average for the same location, to the variance of the downscaling targets. Based on this analysis, we included anomaly values for two variables, 500 hPa geopotential height and 850 hPa temperature, to the final set of predictor variables (Table 1).

2.3.2 | SOM based downscaling

The SOMs downscaling method introduced by Hewitson and Crane (2006) uses the SOM clustering method (Kohonen, 1990) to create a set of typical synoptic conditions for a region. This is an objective process similar to synoptic typing, where human observers would classify large scale weather patterns in a region based on the observed patterns. For each identified pattern, the station observations for all of the days in TRAIN matching that pattern are combined, to create a distribution of values for similar days. To produce the downscaled projections, days in the future climate are sampled from the CMIP5 data, and compared to the identified patterns. Once the most similar pattern is identified, we sample from the historical distribution for that pattern, to produce a projection of the station conditions on that day. Variations on this method have been used to produce downscaled

estimates for climate model projections over South Africa (Hewitson and Crane, 2006) and Florida (Sinha *et al*., 2018). A more detailed explanation of the method is provided in Appendix A.

2.3.3 | Random Forest method

Random Forests (RFs) are a widely used machine learning tool (Breiman, 2001). Using an ensemble of decision trees, RFs provide a computationally inexpensive, nonlinear predictor that performs well on a wide variety of tasks with minimal tuning (Hastie *et al*., 2009). To train the model, a large number of individual decision trees are created from subsets of the data. To make predictions, the results from each of the individual trees are averaged together, to smooth out the over-fitting and variability of the individual trees. For downscaling, RFs are trained using the standardized NCEP reanalysis data and the station observations, with a different model being trained for each station and variable. These models are then used to project daily values of Tmax, Tmin, and PRCP for future climate scenarios, using input data from the CMIP5 GCMs. RFs have been used in downscaling low resolution remote sensing products (Hutengs and Vohland, 2016), precipitation on the island of Borneo (Sa'adi *et al*., 2017), and temperature in the Pearl River Basin of China (Pang *et al*., 2017). More information on the RF method and its adaptation to downscaling are provided in Appendix A.

2.3.4 | Additional methods

In addition to the SOM and RF methods, we test a multiple linear regression model and a simple quantile mapping bias corrector method for comparison. The quantile mapping method takes the bias at each quantile between the reanalysis data and the observations, and applies a correction factor for that bias to the climate model values. These methods did not perform well when compared to the SOM and RF however, so we focus on those two methods for the rest of the analysis. Each of these methods train a different model for each station and each downscaling predictand (Tmax, Tmin, PRCP), creating a total of 837 (279 stations × three variables) models per method.

2.4 | Derived outputs

In addition to Tmax, Tmin, and PRCP, we compute the standardized precipitation and evapotranspiration index (SPEI)[1] and HI distributions across the region. These two metrics provide additional information on how the changes to temperature and precipitation will affect individual activity and agriculture in the region. SPEI is calculated using monthly average of temperature and precipitation to create a standardized index for the relative available moisture in an area (Vicente-Serrano *et al*., 2010), as a result of precipitation and evaporation. It has been widely used in analyses of changing drought conditions with climate change (e.g., Vicente-Serrano *et al*., 2012; Cook *et al*., 2014; Naumann *et al*., 2018).

Increases in HI indicate an increased risk of heat-related mortality, already one of the leading weather-related risk factors in the United States (CDC, 2006; Fechter-Leggett *et al*., 2016), and decreases in the available safe outdoor labour hours (Dunne *et al*., 2013). HI is calculated using a combination of air temperature and humidity (Rothfusz, 1990). To project HI in the future climate scenarios, we downscale values for dew point temperature (Tdew), and calculate HI based these values. Details on these calculations can be found in Appendix B. Tdew is not available in the GHCN database that we use as the primary observational data set. To downscale Tdew, we use the Global Surface Summary of Day (GSOD), which provides daily data for stations around the globe (National Climatic Data Center, 2020). This dataset has fewer stations available in the Midwest, only 62 with sufficient data available, but these stations are well distributed across the study region, and the average temperature across these stations is not significantly different than the average of the GHCN stations (Figure 2).

2.5 | Validation and evaluation of methods

To evaluate the variable selection, the statistical distribution of downscaling results calculated from the reanalysis over the HIST period is compared to the statistical distribution of observations for the same period at each station. The downscaling was carried out over the full year, in order to capture the climate extremes and preserve the widest range of values available to the downscaling. For example, a spring day in 2100 could have a synoptic pattern typical of a summer day in 2000. However, to produce a better seasonal separation in the downscaled temperature and precipitation, the sine of the day of the year was included as a predictor. This variable provides information about the seasonal cycle to the downscaling model, without prohibiting shifts in the seasons in future climates.

To test the downscaling approaches, we analyse the differences between the distributions using the probability distribution function (PDF) skill score introduced by

(Perkins *et al*., 2007). The PDF skill score was designed to provide an easily interpretable metric for the similarity between two distributions, and is calculated as

$$S_{\text{score}} = \sum_{1}^{n} \min(Z_{\text{m}}, Z_{\text{o}}), \tag{1}$$

where S_{score} is the skill score, Z_{m} and Z_{o} are the frequency of modelled and observed values in a given bin, respectively, and n is the number of bins used to calculate the PDF. With this formulation, a score of one indicates perfect overlap between the observed and modelled distributions, with a score of zero representing no overlap. The PDF skill score is designed to capture differences in the distribution of the variable, rather than in the mean or individual values, and has the advantages of being easy to calculate and easy to compare across variables. It has been used in a variety of studies evaluating both GCMs and downscaling methods (e.g., Kjellström *et al*., 2010; Koutroulis *et al*., 2016).

This test better aligns with the goals of downscaling than some commonly used metrics such as Pearson correlation or root mean squared error (RMSE), which asses how accurately the downscaling method recreates the conditions on a specific day in the historical record. By focusing on distribution based metrics, we capture how well the downscaling method performs at estimating the range of events based on the synoptic conditions around the station.

Comparisons of the distribution of the HIST observed Tmax for an example station (Figure 5, black) and the downscaled variables (blue) shows small differences over the full year, and in the winter and summer seasons. There are larger differences in the fall and spring distributions, likely due to the increased variability of temperatures in these transition seasons.

To ensure that the methods are not over-fitting to the training data (1948–2005), we apply the downscaling models to an independent historical test dataset of 2006–2015. While there is not a large change in the underlying climate conditions between the two periods, using an independent test set allows us to ensure that the model is not over-fitting to the training data. In the training period, the SOM (Figure 6, left) and RF (6, right) models closely match the observed statistical distribution. The observed climate in 10-year TEST period (red) has a decrease in the frequency of colder days, and a slight increase around the peak of warmer days compared to the HIST period used to train the downscaling models.

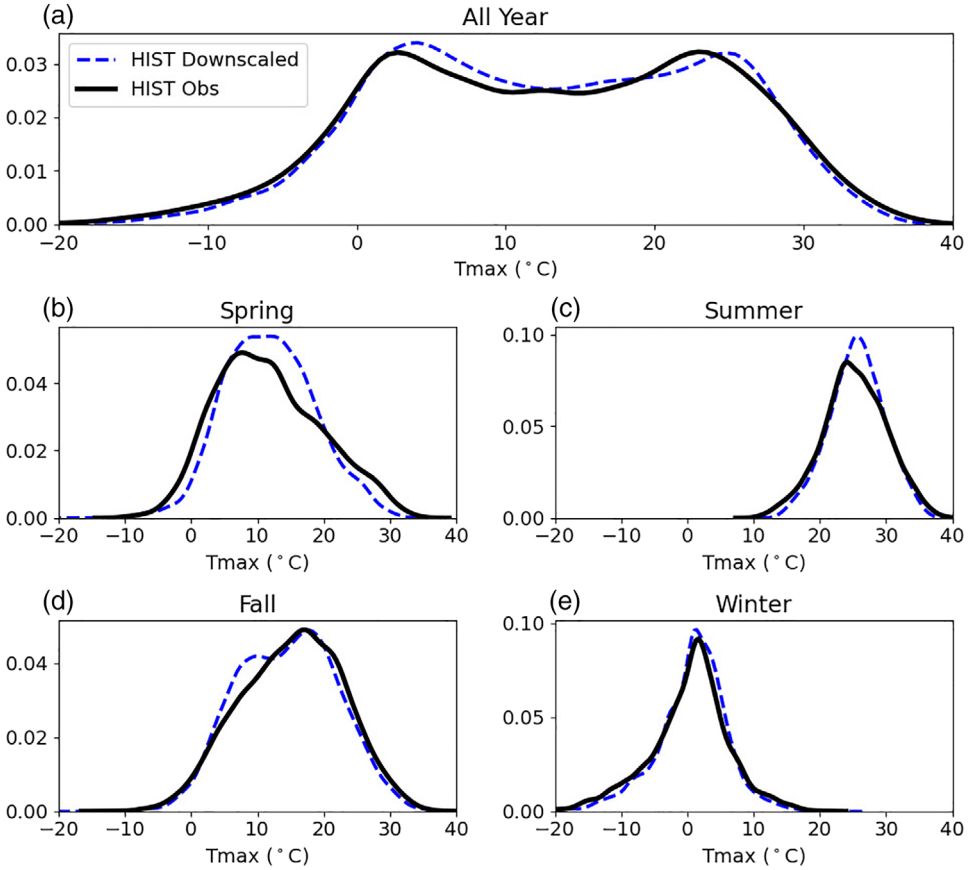

FIGURE 5 Kernel density plots of the distribution of observed (solid) and downscaled Tmax (dashed). Downscaled data are calculated using the RF method applied to the NCEP reanalysis data. In both cases, data are for the same station as in Figure 1. Comparisons are plotted for the full year (top), and each season. The RF downscaling accurately reproduces the annual distribution of Tmax, but struggles to fully match the seasonal distributions [Colour figure can be viewed at wileyonlinelibrary.com]

The downscaled distributions for the SOM and RF models for the TEST period shift in the same ways, but the decrease in colder days is much smaller in the downscaled values than in the observations.

In addition to reproducing the distribution of events for the observed period, the downscaled results must be responsive to shifts in the input data—if the synoptic scale climate shifts, the downscaled data should shift to

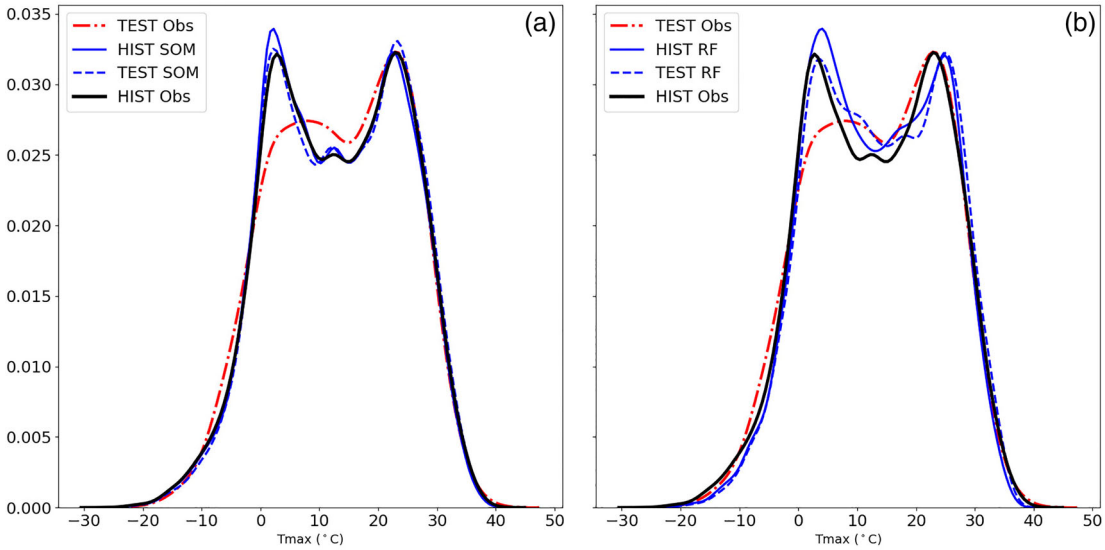

FIGURE 6 Kernel density plot for Tmax of the different downscaling methods for an example station (same station as in Figure 1). Blue lines denote the SOM (left) and RF (right) downscaled estimates of Tmax. Solid lines depict the HIST (1948–2005), with dashed lines for TEST (2006–2015). The thick lines display the station observations (Obs) for HIST (solid) and TEST (dotted). The SOM method most accurately recreates the HIST observations, but does not respond to the differences in TEST as well as the RF model [Colour figure can be viewed at wileyonlinelibrary.com]

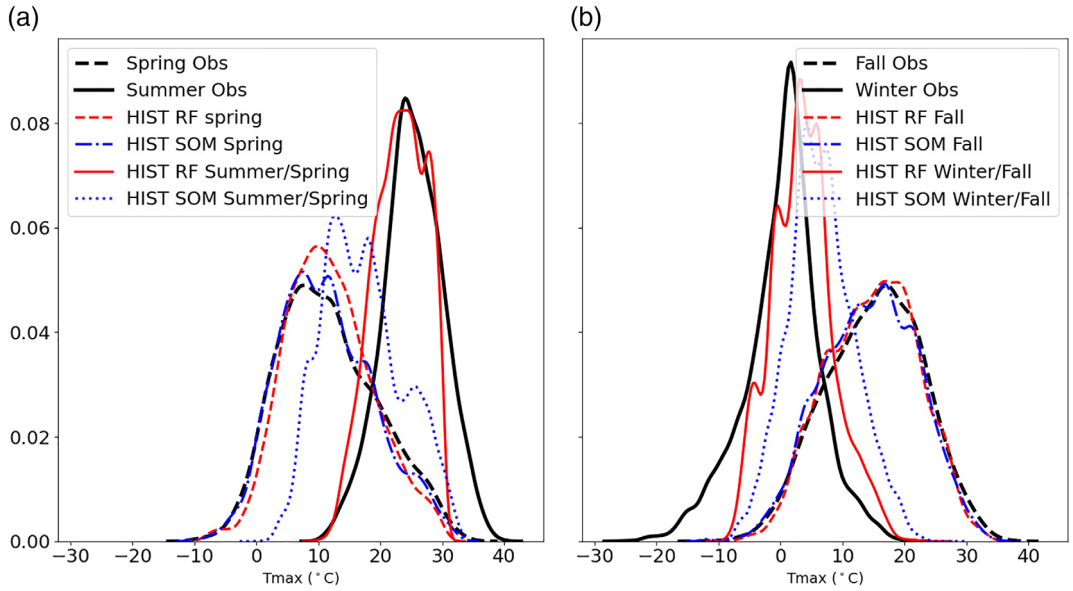

FIGURE 7 Kernel density plot of the downscaled Tmax using the SOM and RF methods for an example station (same station as in Figure 1), compared with the historical observations (Obs). The results plotted represent a test of whether a downscaling model built on data from one climate regime (e.g., spring, fall) can be used to effectively reproduce the statistics of the predictand in a different regime (summer, winter). Details on this test can be found in Section 2.5. Dashed lines are for downscaling models trained and evaluated on the same season: Spring (left) and fall (right). Solid/dotted lines represent the same downscaling models, now applied to input data from the more extreme season: Summer/spring (left) and winter/fall (right). These results demonstrate that these downscaling models can produce a distribution that differs from the training set [Colour figure can be viewed at wileyonlinelibrary.com]

match. To test the ability of the downscaling models to detect a change in the underlying climate conditions, we train several additional models. First, models are trained on the spring (March, April, and May) and summer (June, July, and August) months only. Then, the spring model is evaluated on the climate model data for the summer months in HIST. Using this approach, we test how the downscaling models perform at the edges of their training distributions, and how well they cope with a change in the underlying statistics. This provides an estimate for how the different methods will handle climate changes in future periods, evaluating how the model performs in a climate different from the one it was trained on.

Both the SOM and RF method succeed in capturing the shift towards the higher summer temperatures, with the RF model capturing more of the shift (Figure 7, left). Table 3 shows the PDF skill score for the RF and SOM models, comparing the observed summer distribution to the downscaled values from each method from the model trained on the spring, and evaluated on the summer climate model inputs. The RF model outperforms the SOM, with substantially higher skill scores. The same test was carried out for fall (September, October, and November) and winter (December, January, and February) months (Figure 7, right), with similar results. The RF model trained for the fall input data was able to capture much of the shift to colder temperatures when evaluated on the winter input data, the SOM model produced lower temperatures in the winter from fall test, but did not capture as much of the shift as the RF model. The results shown here are for Tmax, and similar tests for Tmin produced similar results. Based on these results, the RF model performed the best for the temperature downscaling, and was used to produce the future projections of Tmax and Tmin.

In addition to testing the accuracy of the downscaled distributions for temperature, we analyse the results for

TABLE 3 Comparison of the PDF skill score for the distribution of values produced by the self-organizing map (SOM) and random forest (RF) downscaling methods to the distribution of observed values over the HIST period. Higher values indicate more similarity between the two distributions, with a value of one being perfect correspondence. Three different sets of input variables are compared, using only 850 hPa temperature, and the full set of variables with and without day of year, as described in Table 1. Based on these results, the RF outperforms the SOM

	T_{850}	No. day of year	All variables	Summer/spring	Winter/fall
SOM	0.951	0.944	0.948	0.561	0.628
RF	0.916	0.906	0.925	0.678	0.723

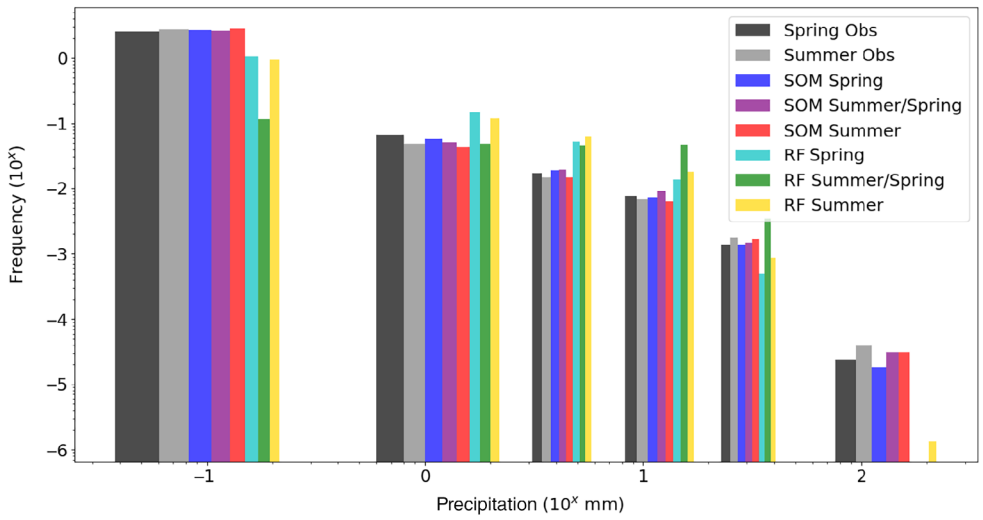

FIGURE 8 Histogram of PRCP amount comparing the SOM and RF methods. The distribution is plotted as a histogram with log axes to capture the large number of no-precipitation days and the long tail of large precipitation events. Bars from left to right are: Obs (Spring, Summer), SOM (Spring, Summer/Spring, Summer), RF (Spring, Summer/Spring, Summer). The SOM models generally follow the observed spring and summer values, while the RF model underestimating the number of dry days and the number of large precipitation events. The difference between the spring and summer distributions are much smaller than for Tmax. The SOM summer/spring model captures most of the shift between the spring and summer precipitation, with the RF model demonstrating less skill on this test [Colour figure can be viewed at wileyonlinelibrary.com]

the same downscaling methods now applied to daily precipitation (Figure 8). The observed change between the stations is much smaller than for Tmax, but the SOM model in both seasons does a much better job of capturing the observed distributions. The RFs in both seasons underestimate the number of days with no precipitation, and the number of large precipitation events, similar to the pattern displayed by the PRCP values from the GCMs (Figure 1). This bias towards middling values of precipitation is likely a consequence of the RF method, which averages the results from all the individual trees in the model. The summer/spring RF model shows a very large shift towards higher average precipitation that is not shown in the observations, indicating that this method may produce unphysical results for precipitation under a changing climate. This error is reflected in the PDF skill score values, with the RF model producing significantly lower scores than the SOM model (Table 4). Based on the results of our validation tests, the SOM method produces the most realistic projections for precipitation across the region, and we expect that it will continue to do so in projections of the future climate. As a result, we use the SOM model to produce the future projections of PRCP.

3 | RESULTS

We produce downscaled projections for each downscaling method for the HIST, TEST, NEAR, and FAR periods. We consider 30-year NEAR (2021–2050), and FAR (2071–2100) periods for each of the three (RCP 2.6, 4.5, 8.5) climate scenarios. The Tmax and Tmin results use the RF model, while the PRCP results use the SOM method. We produce downscaled projections for each of the CMIP5 models, which are combined to create a single projection for each scenario and period. This effectively creates a larger dataset for each period, and average out individual variations in models that may be present over a 30 year period. A trend towards increasing temperatures, mirroring the global changes for each scenario, is seen across the region.

3.1 | Temperature

In the RF projections, all three RCP scenarios have a nearly uniform increase in average Tmax and Tmin across the region. For the 2071–2100 period, the downscaled daily Tmax increases by an average of 1.1 (RCP 2.6), 2.4 (RCP 4.5), and 4.3°C (RCP 8.5). The gridpoint CMIP5 models produced an average increase in the temperature in the target region of 1.3, 2.9, and 4.8°C in the 2.6, 4.5, and 8.5 scenarios, respectively compared to the HIST period. The RF downscaling produces a slightly weaker warming signal across all three scenarios. The downscaling in the TEST period also (correctly) produces a slightly weaker warming signature than the CMIP5 GCMs. In addition to the increase in average

	T_{850}	No. day of year	All variables	Summer/spring	Winter/fall
SOM	0.97	0.973	0.973	0.949	0.949
RF	0.776	0.817	0.813	0.383	0.716

TABLE 4 As in Table 3, but for PRCP. Here the SOM outperforms the RF model

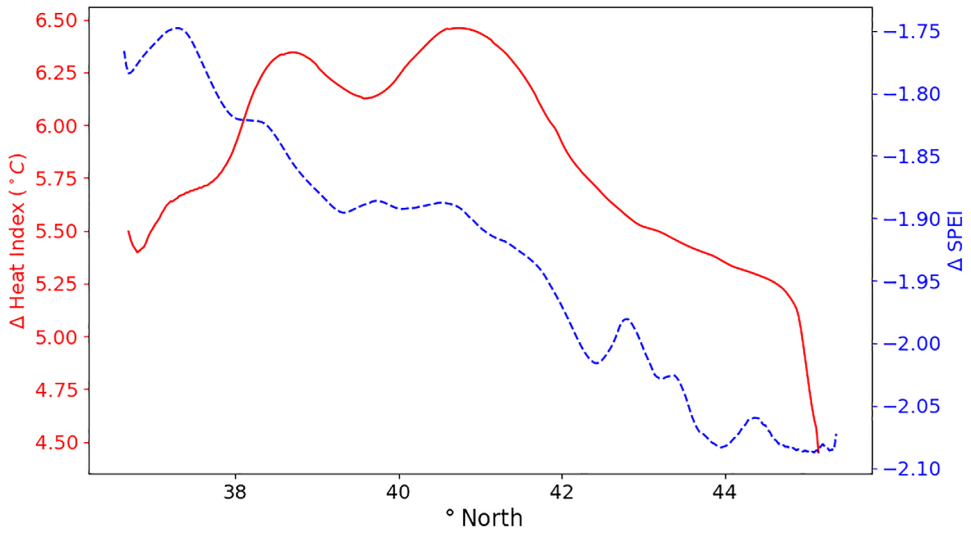

FIGURE 9 Change in average summer heat index (HI, solid, left) and average year-round SPEI (dashed, right) between the HIST and FAR future period under the RCP 8.5 scenario, averaged east–west across the domain. More positive HI indicates higher temperatures in the FAR period, while more negative SPEI indicates increased drought. HI is calculated using downscaled dew point values at the smaller set of GSOD stations (see Figure 2) [Colour figure can be viewed at wileyonlinelibrary.com]

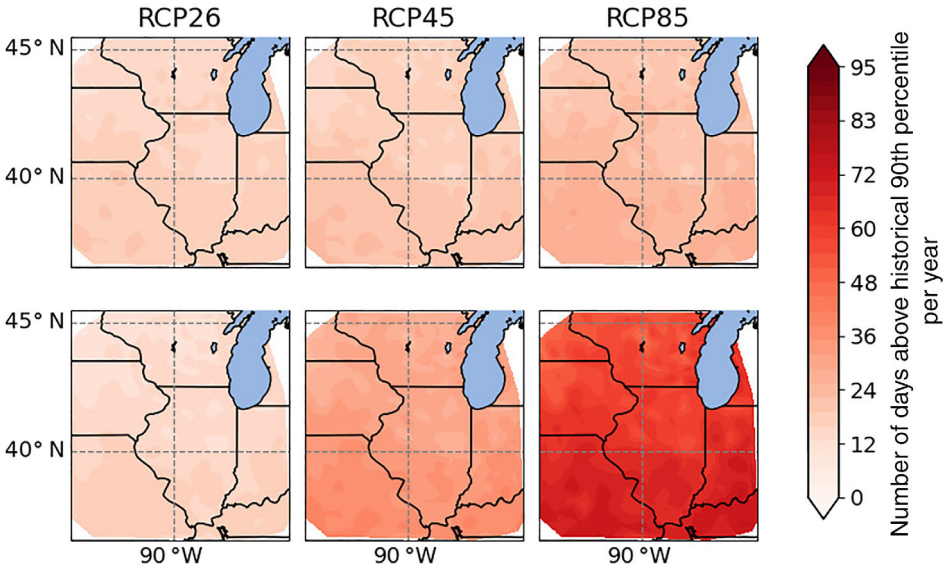

FIGURE 10 Change in the number of days with extreme heat. Extreme heat is defined here as Tmax above the 90th percentile in the historical period for each station individually. Difference between the HIST and each of the two future periods, NEAR (top) and FAR (bottom) is plotted for the three RCP scenarios. The number of extreme heat days close to triples across the region in the FAR period under the RCP 8.5 scenario. In no case does downscaling indicate a decrease in the number of days of extreme heat at a station [Colour figure can be viewed at wileyonlinelibrary.com]

temperature, we calculate the change in the length of the summer season, computed as the length of time the average Tmax is above the June 1st and August 31st temperatures from the historical period. The summer length in the FAR period increases by 10, 26, and 48 days for the RCP 2.6, 4.5, and 8.5 scenarios, respectively.

The Tmax changes projected by our downscaling models agree with the general trend of consistent warming across the region shown in other downscaling studies (Hayhoe *et al.*, 2010; Byun and Hamlet, 2018). However, our results give a generally smaller increase in temperature by the end of the century. There remains significant uncertainty around the magnitude of the future changes, especially on the local scale, but based on the results of our validation tests, we conclude our results are reasonable projections for the different RCP scenarios.

3.2 | Heat index

To calculate HI, we first downscale Tdew using the RF method and the same set of input variables. We find that the downscaling accurately reproduces the historical distribution of Tdew. Averaging across these stations, Tdew increases by an average of 1.0, 2.5, and 4.3°C in the FAR period for the RCP 2.6, 4.5, and 8.5 scenarios. Calculating the change in HI based on the Tmax and Tdew, we find an average increase in the summer HI of 1.6, 3.3, and 5.8°C for the three RCP scenarios (Figure 9).

US National Weather Service (NWS) defines extreme heat stress events as a HI of 105°F (40.6°C) for 3 hr or more (NWS, 2009). Our calculation of HI using Tmax does not give a duration, but the average number of days per year with a maximum HI exceeding 40.6°C increases from an average of 1 day in the HIST period to 19 days in the FAR period under RCP 8.5, averaged across all stations. An increase of this size will have significant health consequences for residents of the Midwest (Wehner *et al.*, 2016; Matthews *et al.*, 2017). Mukherjee *et al.* (2021), calculating HI from grid scale changes in Tmax and Tdew for future climate scenarios, found a similar dramatic increase in the frequency and maximum severity of heat stress events in the Midwest under the RCP 8.5 scenario.

3.3 | Extreme heat

To examine the prevalence of extreme heat in the future scenarios, we estimate the frequency of hot days as the number of days in each period exceeding the 90th percentile Tmax in HIST. The increase in the number of days of extreme heat is relatively uniform across the region (Figure 10). To produce the maps of the full region, the individual station values were linearly interpolated to fill the remaining space. For the NEAR period, the increase is about 15 (RCP 2.6), 21 (RCP 4.5), and 23 (RCP 8.5) days per year. In the FAR period the frequency of these hot days increases further for the RCP 4.5 and RCP 8.5 scenarios, to 32 and 61 days respectively. As with the average Tmax, the change in extreme temperature days is near uniform across the region, indicating the shift in temperature is generally consistent across the Midwest.

In addition to the frequency of days above the historical 90th percentile, the SOM method can be used to identify the synoptic signatures of days that have produced extreme values of temperature. In the reanalysis data the HIST period, the days that tend to correspond to the

highest station temperatures fall into a small number of SOM nodes. We compare the number of days in the CMIP5 models that fall on these nodes in the HIST and future scenarios to estimate the frequency of potential extreme days for each station. Taking an example station (Kenosha, WI; 42.6° N, 87.8° W), and examining the nodes that have at least 10% of their days in the hottest 1% of days overall (daily Tmax above 33.9°C), the number of days falling in those nodes increases from 491 (over the 30 years) in the historical period, to 724 in FAR under the RCP 8.5, corresponding to experiencing these extreme temperatures for an additional week or more per year. This change will increase the likelihood of record setting temperatures, as the synoptic conditions will more frequently be in place for conditions that have historically produced the highest station Tmax. Across all stations, this number of days meeting these criteria increases by an average of 48% in the RCP 8.5 scenario. This value is likely to be a conservative estimate, as the statistical downscaling does not account for various feedback mechanisms, notably soil-moisture, that are known to increase the severity of Midwest heat waves (Durre *et al*., 2000; Zhang *et al*., 2020). The increase in frequency of extreme heat will pose a significant health risk to populations across the region; the 2018 National Climate Assessment projects 2,000 additional deaths per year due to extreme heat by 2090 under this scenario (Angel *et al*., 2018).

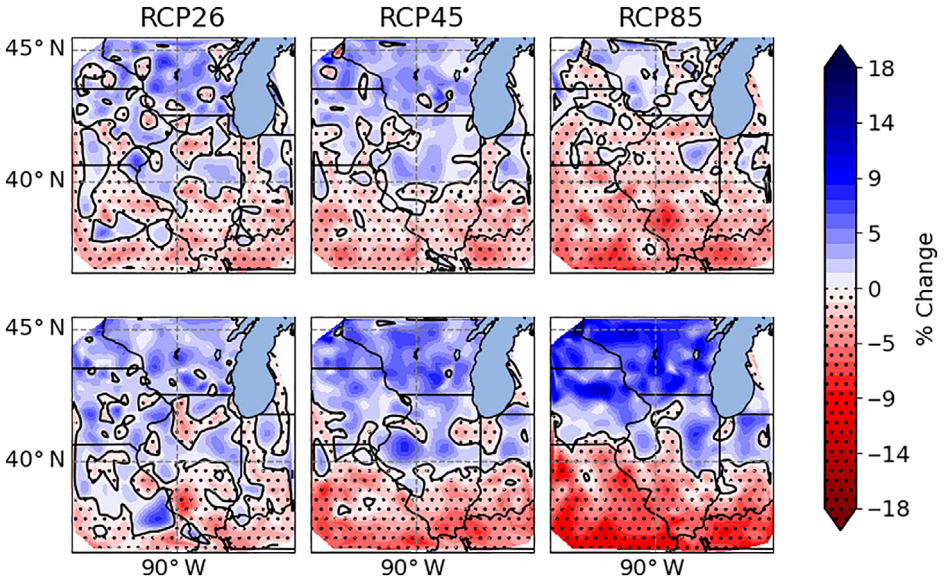

FIGURE 11 Percentage change in precipitation across the study region for the three RCP scenarios for the two future periods (NEAR: top, FAR: bottom). Stippling represents areas with a negative change in precipitation. In all cases, there is a general trend of decreased precipitation to the south, and increased precipitation to the north. These trends are most pronounced in the FAR (2071–2100) period and in the higher emissions scenarios [Colour figure can be viewed at wileyonlinelibrary.com]

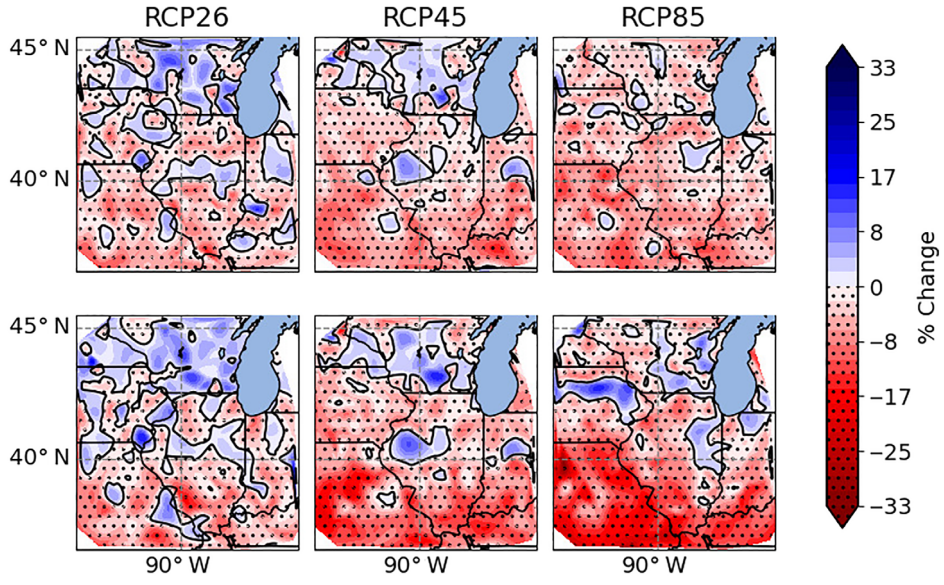

FIGURE 12 Percentage change in summer (June, July, and August) months precipitation from the HIST baseline. There is a larger percentage decrease in precipitation in the summer months than for the annual rainfall. This is balanced by a general increase in precipitation in winter (December, January, and February) precipitation [Colour figure can be viewed at wileyonlinelibrary.com]

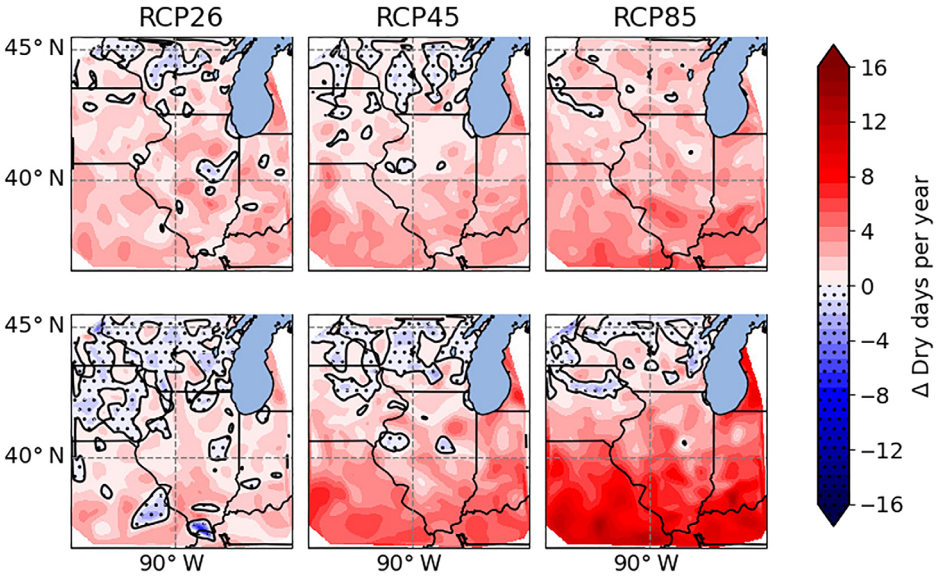

FIGURE 13 Average change in the number of days with daily precipitation less than a trace (0.1 mm·day^{-1}) for the three RCP scenarios and the two future periods (NEAR: top, FAR: bottom). Many of the areas with an increase in the average precipitation (Figure 11) show an increase in the number of dry days, indicating an increase in the precipitation amount on the days where it is present [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | Precipitation

Average daily precipitation exhibits a general trend of increasing precipitation north of 40° N and decreasing precipitation to the south (Figure 11) compared to HIST for both the NEAR and FAR future periods across all three RCP scenarios. For precipitation, we use the kriging method for interpolation to create the regional maps of average changes. Our study region currently receives an average of between 2 and 4 mm·day^{-1} (70 and 150 cm·year^{-1}, Figure 3). The most extreme changes, to both the north and south of the study area in the FAR future period of the RCP 8.5 scenario, represent a 10–15% increase (decrease) in the northern (southern) average precipitation compared to HIST. The decrease in precipitation falls predominantly in the summer months (Figure 12), with some areas of Missouri experiencing decreases in summer precipitation of 1 mm·day^{-1}. These changes contrast with a general increase in winter precipitation across the region. Decreased precipitation in the summer growing season is especially worrisome as most of the agriculture in the region is rain-fed. The decrease in precipitation, combined with increased evaporation due to the projected increase in temperature, will significantly reduce the soil moisture available to growing crops.

The maximum downscaled value for precipitation is limited by the maximum observed in HIST, due to the limitations of the SOM method. While new maximum values will not be produced, we can examine the frequency at which days with large amounts of precipitation occur in the future periods. The downscaled values across all stations do not show a significant increase in the number of days with more than 12 mm (0.5 in.) of precipitation in a day. However, even in areas near the Illinois/Wisconsin border that show a small increase in precipitation there is an increase in the number of days with no precipitation (Figure 13), signalling a shift towards larger amounts of precipitation, when precipitation does occur.

Our results agree with prior studies (Hayhoe *et al.*, 2010; Byun and Hamlet, 2018) in showing a greater decrease in precipitation during the summer months compared to the annual average. However, there are notable differences in the spatial distribution and magnitudes of the changes, with our results displaying a pattern of increased precipitation in the north, decreased precipitation in the south that is not seen in these other studies. For the equivalent time period and climate scenario, our results also show generally greater decreases in precipitation in the region than these studies.

3.5 | Drought

To quantify potential changes in drought, we use the SPEI to provide an estimate of drought conditions based on average temperature and precipitation (Vicente-Serrano *et al.*, 2010). SPEI combines precipitation and potential evapotranspiration based on temperature to estimate water balance for a location. Using the downscaled projections for temperature and precipitation, we calculate monthly values for SPEI for each station in the region. Over HIST, SPEI values calculated from the downscaled projections based on the NCEP reanalysis correlate with SPEI values calculated from the observations with an r^2 value of .81, indicating that the

downscaled values do a good job of recreating the historical record on the monthly time scale. For the future periods, we calibrate the SPEI calculation using the model values for HIST to establish the baseline for drought at each station, and use this calibration for the future periods.

Across the region, SPEI is likely to decrease (indicating drier conditions) in the future periods, with these changes most pronounced in the RCP 8.5 scenario. Averaging over the 30 year FAR future period, the average change in SPEI for the region is −1.9 (Figure 9), indicating that droughts in the region have the potential to become more severe, with important consequences for agriculture in the region. An SPEI value of −2.0 or less is typically defined as extreme drought (Paulo *et al*., 2012). We project the frequency of values below this threshold will increase from once every 4 years in the HIST period, to six times per year in the FAR period under RCP 8.5, or three times per year under RCP 4.5.

4 | CONCLUSIONS

Many of the effects of a changing climate will be realized on scales too small to be resolved by the current generation of climate models. Statistical downscaling methods bridge the gap between the resolution of the GCMs, and the local scales at which information is needed for understanding a wide variety of climate change impacts. We develop downscaling models using two machine learning approaches, SOMs and RFs, and evaluate them against historical observations. Of these methods, the RF performs best for Tmax and Tmin, matching the historical distribution and generalizing well to new climate conditions, while the SOM method produces the best results for precipitation. We evaluate the ability of these downscaling models to produce realistic estimates at the extremes of the historical data used to develop them as follows: based on observations from the spring months to produce skilful summer temperatures and precipitation; similarly, we evaluate the skill of downscaling models developed on data from fall in reproducing realistic winter values. While these methods have limitations, based on their performance on these and other validation tests, we find that these methods provide a physically realistic, high resolution representation of the station-level weather in future climate scenarios.

Based upon the downscaled temperature and precipitation projections for the Midwest, precipitation will increase in the northern half of the region by up to 15% and decrease in the southern half by a similar amount. Temperatures show a nearly uniformly increase across the region, with very little spatial variation in the amount of warming. The magnitude of these changes scales with the radiative forcing applied in the three RCP scenarios. Increases in HI and the frequency of excessive heat are likely to lead to increases in the incidence of heat-related illness and mortality (Basu, 2009; Matthews *et al*., 2017). Increased heat stress and drought are likely to decrease crop yields in much of the region (Schlenker and Roberts, 2009; Angel *et al*., 2018), especially in the higher emissions scenarios.

Climate change will bring significant disruption to the Midwest, with increasing temperatures and shifting patterns of precipitation. In the RCP 8.5 scenario, the number of days with maximum temperatures above the current 90th percentile will almost triple, along with an average increase of 4.3°C across the region, and the length of summer weather will increase by 48 days, an increase of more than 50%. Precipitation will shift northward across the region, increasing water stress, particularly in the summer months and across the southern portion of the region. Droughts will become more frequent and more severe, as precipitation decreases and evaporation increases due to increased temperature. These changes will pose significant challenges for agriculture and health in the region.

ACKNOWLEDGEMENTS

Thanks to Palash Sinha for discussions and data for the initial downscaling work. Thanks as well to Chad Bahrmann and Justin Pettuci for computing support, and to the Penn State Institute for Computational and Data Sciences for computational resources. We would also like to thank two anonymous reviewers who provided excellent comments and feedback that lead to an improved manuscript, and the journal editor for timeliness in processing the manuscript. This work was supported by the National Science Foundation under grant 1639342.

AUTHOR CONTRIBUTIONS

Andrew Polasky: Data curation, formal analysis, investigation, methodology, software, validation, visualization, writing - original draft preparation. **Jenni Evans:** Conceptualization; funding acquisition; investigation; resources; supervision; writing – review and editing. **Jose D. Fuentes:** Conceptualization; funding acquisition; investigation; project administration; supervision; writing – review and editing. **Holly Hamilton:** Conceptualization; data curation; investigation.

ORCID

Andrew D. Polasky https://orcid.org/0000-0003-1640-6630
Jenni L. Evans https://orcid.org/0000-0003-3425-0932
Jose D. Fuentes https://orcid.org/0000-0002-6177-6326

ENDNOTE

[1] SPEI values were calculated using the climate indices Python library (Adams, 2017).

REFERENCES

Adams, J. (2017) Climate indices, an open source python library providing reference implementations of commonly used climate indices. Available at: https://github.com/monocongo/climate_indices [Accessed 10 Jan 2020].

Agusdinata, D.B. and Lukosch, H. (2019) Supporting interventions to reduce household greenhouse gas emissions: a transdisciplinary role-playing game development. *Simulation & Gaming*, 50(3), 359–376. https://doi.org/10.1177/1046878119848135.

Angel, J., Swantson, C., Boustead, B.M., Conlon, K.C., Hall, K.R., Jorns, J.L., Kunkel, K.E., Lemos, M.C., Lofgren, B., Ontl, T.A., Posey, J., Stone, K., Takle, G. and Todey, D. (2018) Midwest. In: Reidmiller, D., Avery, C., Easterling, D., Kunkel, K., Lewis, K., Maycock, T. and Stewart, B. (Eds.) *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment* [Chapter 21], Vol. II. Washington, DC: U.S. Global Change Research Program, pp. 872–940. https://doi.org/10.7930/NCA4.2018.CH21.

Basu, R. (2009) High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environmental Health*, 8(1), 1–13. https://doi.org/10.1186/1476-069X-8-40.

Breiman, L. (2001) Random forests. *Machine Learning*, 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Byun, K. and Hamlet, A.F. (2018) Projected changes in future climate over the Midwest and Great Lakes region using downscaled CMIP5 ensembles. *International Journal of Climatology*, 38, 531–553. https://doi.org/10.1002/joc.5388.

CDC. (2006) *Heat-Related Deaths—United States, 1999–2003*. Atlanta, GA: US Centers for Disease Control and Prevention, pp. 796–798.

Changnon, S., Angel, J., Kunkel, K. and Lehmann, C. (2004) *Climate Atlas of Illinois*. Champaign, IL: Illinois State Water Survey.

Cook, B.I., Smerdon, J.E., Seager, R. and Coats, S. (2014) Global warming and 21st century drying. *Climate Dynamics*, 43(9), 2607–2627. https://doi.org/10.1007/s00382-014-2075-y.

Dunne, J.P., Stouffer, R.J. and John, J.G. (2013) Reductions in labour capacity from heat stress under climate warming. *Nature Climate Change*, 3(6), 563–566. https://doi.org/10.1038/nclimate1827.

Durre, I., Wallace, J.M. and Lettenmaier, D.P. (2000) Dependence of extreme daily maximum temperatures on antecedent soil moisture in the contiguous United States during summer. *Journal of Climate*, 13(14), 2641–2651. https://doi.org/10.1175/1520-0442(2000)013<2641:DOEDMT>2.0.CO;2.

Fechter-Leggett, E.D., Vaidyanathan, A. and Choudhary, E. (2016) Heat stress illness emergency department visits in national environmental public health tracking states, 2005–2010. *Journal of Community Health*, 41(1), 57–69. https://doi.org/10.1007/s10900-015-0064-7.

Fritsch, J.M., Kane, R. and Chelius, C. (1986) The contribution of mesoscale convective weather systems to the warm-season precipitation in the United States. *Journal of Climate and Applied Meteorology*, 25(10), 1333–1345. https://doi.org/10.1175/1520-0450(1986)025<1333:TCOMCW>2.0.CO;2.

Gerken, T., Bromley, G.T. and Stoy, P.C. (2018) Surface moistening trends in the northern North American Great Plains increase the likelihood of convective initiation. *Journal of Hydrometeorology*, 19(1), 227–244. https://doi.org/10.1175/JHM-D-17-0117.1.

Glotter, M. and Elliott, J. (2016) Simulating US agriculture in a modern dust bowl drought. *Nature Plants*, 3(1), 1–6. https://doi.org/10.1038/nplants.2016.193.

Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning*, 2nd edition. New York, NY: Springer.

Hatfield, J.L., Boote, K.J., Kimball, B., Ziska, L., Izaurralde, R.C., Ort, D., Thomson, A.M. and Wolfe, D. (2011) Climate impacts on agriculture: implications for crop production. *Agronomy Journal*, 103(2), 351–370. https://doi.org/10.2134/agronj2010.0303.

Hatfield, J.L., Wright-Morton, L. and Hall, B. (2018) Vulnerability of grain crops and croplands in the Midwest to climatic variability and adaptation strategies. *Climatic Change*, 146(1), 263–275. https://doi.org/10.1007/s10584-017-1997-x.

Hayhoe, K., VanDorn, J., Croley, T., II, Schlegal, N. and Wuebbles, D. (2010) Regional climate change projections for Chicago and the US Great Lakes. *Journal of Great Lakes Research*, 36, 7–21. https://doi.org/10.1016/j.jglr.2010.03.012.

Hayhoe, K., Edmonds, J., Kopp, R., LeGrande, A., Sanderson, B., Wehner, M. and Wuebbles, D. (2017) Climate models, scenarios, and projections. In: Wuebbles, D., Fahey, D., Hibbard, K., Dokken, D., Stewart, B. and Maycock, T. (Eds.) *Climate Science Special Report: Fourth National Climate Assessment* [Chapter 4], Vol. I. Washington, DC: U.S. Global Change Research Program, pp. 133–160. https://doi.org/10.7930/J0J964J6.

Hewitson, B.C. and Crane, R.G. (2006) Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *International Journal of Climatology*, 26(10), 1315–1337. https://doi.org/10.1002/joc.1314.

Hutengs, C. and Vohland, M. (2016) Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, 178, 127–141. https://doi.org/10.1016/j.rse.2016.03.006.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, D., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3), 437–472. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kilsby, C.G., Jones, P.D., Burton, A., Ford, A.C., Fowler, H.J., Harpham, C. and Wilby, R.L. (2007) A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, 22(12), 1705–1719. https://doi.org/10.1016/j.envsoft.2007.02.005.

Kjellström, E., Boberg, F., Castro, M., Christensen, J.H., Nikulin, G. and Sánchez, E. (2010) Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. *Climate Research*, 44(2–3), 135–150. https://doi.org/10.3354/cr00932.

Kohonen, T. (1990) The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464–1480. https://doi.org/10.1109/5.58325.

Koutroulis, A.G., Grillakis, M., Tsanis, I. and Papadimitriou, L. (2016) Evaluation of precipitation and temperature simulation

performance of the CMIP3 and CMIP5 historical experiments. *Climate Dynamics*, 47(5), 1881–1898. https://doi.org/10.1007/s00382-015-2938-x.

Li, K.-C. (1991) Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association*, 86(414), 316–327. https://doi.org/10.1080/01621459.1991.10475035.

Liang, X-Z., Wu, Y., Chambers, R.G., Schmoldt, D.L., Gao, W., Liu, C., Liu, Y-A., Sun, C. and Kennedy, J.A. (2017) Determining climate effects on US total agricultural productivity. *Proceedings of the National Academy of Sciences*, 114(12), E2285–E2292. https://doi.org/10.1073/pnas.1615922114.

Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M. and Hammer, G.L. (2014) Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science*, 344(6183), 516–519. https://doi.org/10.1126/science.1251423.

Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M. and Thiele-Eich, I. (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48(3), 1–34. https://doi.org/10.1029/2009RG000314.

Matthews, T.K., Wilby, R.L. and Murphy, C. (2017) Communicating the deadly consequences of global warming for human heat stress. *Proceedings of the National Academy of Sciences*, 114(15), 3861–3866. https://doi.org/10.1073/pnas.1617526114.

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E. and Houston, T.G. (2012) An overview of the Global Historical Climatology Network—daily database. *Journal of Atmospheric and Oceanic Technology*, 29(7), 897–910. https://doi.org/10.1175/JTECH-D-11-00103.1.

Mukherjee, S., Mishra, A., Mann, M. and Raymond, C. (2021) Anthropogenic warming and population growth may double US heat stress by the late 21st century. *Earth's Future*, 9, e2020EF001886. https://doi.org/10.1029/2020EF001886.

Murphy, J. (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate*, 12(8), 2256–2284. https://doi.org/10.1175/1520-0442(1999)012<2256:AEOSAD>2.0.CO;2.

National Climatic Data Center (2020) Global summary of day. Available at: https://catalog.data.gov/dataset/global-surface-summary-of-the-day-gsod [Accessed 24 March 2020].

Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R.A., Carrao, H., Spinoni, J. and Feyen, L. (2018) Global changes in drought conditions under different levels of warming. *Geophysical Research Letters*, 45(7), 3285–3296.

NWS. (2009) *Excessive Heat Warning*. Silver Spring, MD: National Weather Service. Available at: https://w1.weather.gov/glossary/index.php?word=excessive+heat+warning.

Pang, B., Yue, J., Zhao, G. and Xu, Z. (2017) Statistical downscaling of temperature with the random forest model. *Advances in Meteorology*. 2017, 7265178. https://doi.org/10.1155/2017/7265178.

Paulo, A.A., Rosa, R.D. and Pereira, L.S. (2012) Climate trends and behaviour of drought indices based on precipitation and evapotranspiration in Portugal. *Natural Hazards and Earth System Sciences*, 12(5), 1481–1491. https://doi.org/10.5194/nhess-12-1481-2012. Available at: https://nhess.copernicus.org/articles/12/1481/2012/.

Perkins, S.E., Pitman, A., Holbrook, N. and McAneney, J. (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate*, 20(17), 4356–4376. https://doi.org/10.1175/JCLI4253.1.

Raäisaänen, J. (2007) How reliable are climate models? *Tellus A: Dynamic Meteorology and Oceanography*, 59(1), 2–29. https://doi.org/10.1111/j.1600-0870.2006.00211.x.

Radić, V. and Clarke, G.K. (2011) Evaluation of IPCC models' performance in simulating late-twentieth-century climatologies and weather patterns over North America. *Journal of Climate*, 24(20), 5257–5274. https://doi.org/10.1175/JCLI-D-11-00011.1.

Rosenberg, N.J., Crosson, P.R., Frederick, K.D., Easterling, W.E., Mckenney, M.S., Bowes, M.D., Sedjo, R.A., Darmstadter, J., Katz, L.A. and Lemon, K.M. (1993) The MINK methodology: background and baseline. In: *Towards an Integrated Impact Assessment of Climate Change: The MINK Study*. New York, NY: Springer, pp. 7–22. https://doi.org/10.1007/978-94-011-2096-8_2.

Rothfusz, L.P. (1990) *The heat index equation*. National Weather Service Technical Attachment (SR 90–23).

Sa'adi, Z., Shahid, S., Chung, E.-S. and Bin Ismail, T. (2017) Projection of spatial and temporal changes of rainfall in Sarawak of Borneo Island using statistical downscaling of CMIP5 models. *Atmospheric Research*, 197, 446–460. https://doi.org/10.1016/j.atmosres.2017.08.002.

Salathé, E.P., Jr., Mote, P.W. and Wiley, M.W. (2007) Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States Pacific Northwest. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 27(12), 1611–1621. https://doi.org/10.1002/joc.1540.

Schlenker, W. and Roberts, M.J. (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37), 15594–15598. https://doi.org/10.1073/pnas.0906865106.

Shepherd, J.M., Pierce, H. and Negri, A.J. (2002) Rainfall modification by major urban areas: observations from spaceborne rain radar on the TRMM satellite. *Journal of Applied Meteorology*, 41(7), 689–701. https://doi.org/10.1175/1520-0450(2002)041<0689:RMBMUA>2.0.CO;2.

Sinha, P., Mann, M.E., Fuentes, J.D., Mejia, A., Ning, L., Sun, W., He, T. and Obeysekera, J. (2018) Downscaled rainfall projections in South Florida using self-organizing maps. *Science of the Total Environment*, 635, 1110–1123. https://doi.org/10.1016/j.scitotenv.2018.04.144.

Stephens, G.L., L'Ecuyer, T., Forbes, R., Gettlemen, A., Golaz, J-C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P. and Haynes, J. (2010) Dreary state of precipitation in global models. *Journal of Geophysical Research: Atmospheres*, 115(D24), 1–13. https://doi.org/10.1029/2010JD014532.

Taylor, K.E., Stouffer, R.J. and Meehl, G.A. (2012) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1.

USDA. (2017) *Census of agriculture 2017*. United States Department of Agriculture. Technical report.

Vicente-Serrano, S.M., Beguería, S. and López-Moreno, J.I. (2010) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1.

Vicente-Serrano, S.M., Beguería, S., Lorenzo-Lacruz, J., Camarero, J.J., López, J.I., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E. and Sanchez-Lorenzo, A. (2012) Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interactions*, 16(10), 1–27.

Vose, R., Easterling, D., Kunkel, K., Legrande, A. and Wehner, M. (2017) Temperature changes in the United States. In: Wuebbles, D., Fahey, D., Hibbard, K., Dokken, D., Stewart, B. and Maycock, T. (Eds.) *Climate Science Special Report: Fourth National Climate Assessment* [Chapter 6], Vol. I. Washington, DC: U.S. Global Change Research Program, pp. 133–160. https://doi.org/10.7930/J0N29V45.

van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J-F., Masui, T., Meinshausen, M., Makicenovic, N., Smith, S.H. and Rose, S.K. (2011) The Representative Concentration Pathways: an overview. *Climatic Change*, 109(1), 5–31. https://doi.org/10.1007/s10584-011-0148-z.

Watkins, D., Shwom, R., Schelly, C., Agusdinata, D., Floress, K. and Halvorsen, K. (2019) Understanding household conservation, climate change, and the food–energy–water nexus from a transdisciplinary perspective. In: Halvorsen, K., Schelly, C., Handler, R., Pischke, E. and Knowlton, J. (Eds.) *A Research Agenda for Environmental Management*. Cheltenham, USA: Edward Elgar Publishing.

Wehner, M., Stone, D., Krishnan, H., AchutaRao, K. and Castillo, F. (2016) The deadly combination of heat and humidity in India and Pakistan in summer 2015. *Bulletin of the American Meteorological Society*, 97(12), S81–S86. https://doi.org/10.1175/BAMS-D-16-0145.1. Available at: https://journals.ametsoc.org/view/journals/bams/97/12/bams-d-16-0145.1.xml.

Wilby, R.L., Charles, S., Zorita, E., Timbal, B., Whetton, P. and Mearns, L. (2004) *Guidelines for use of climate scenarios developed from statistical downscaling methods*. Supporting material of the Intergovernmental Panel on Climate Change. Available from the DDC of IPCC TGCIA, 27.

Wilks, D.S. (1999) Multisite downscaling of daily precipitation with a stochastic weather generator. *Climate Research*, 11(2), 125–136.

Xiao, Y., Wan, J. and Hewings, G.J.D. (2013) Flooding and the Midwest economy: assessing the Midwest floods of 1993 and 2008. *GeoJournal*, 78(2), 245–258. https://doi.org/10.1007/s10708-011-9415-9.

Zhang, Z., Li, Y., Chen, F., Barlage, M. and Li, Z. (2020) Evaluation of convection-permitting WRF CONUS simulation on the relationship between soil moisture and heatwaves. *Climate Dynamics*, 55(1), 235–252. https://doi.org/10.1007/s00382-018-4508-5.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Polasky, A. D., Evans, J. L., Fuentes, J. D., & Hamilton, H. L. (2022). Statistical climate model downscaling for impact projections in the Midwest United States. *International Journal of Climatology*, *42*(5), 3038–3055. https://doi.org/10.1002/joc.7406