A Climate-Informed Approach to Create Hourly Future Weather Timeseries for Power System Planning

Erik T. Smith, Delavane B. Diaz, Jacob Mardian†

Abstract

Power system planning tools require hourly weather data to capture the variable conditions that influence electricity supply and demand (e.g., temperature, wind, and solar). It is well-established that using current or historical weather conditions is insufficient for planning a resilient system under climate change and that forward-looking data are needed when conducting energy transition studies through 2050. However, nearly all global climate model (GCM) projections are limited to daily temporal resolution, presenting a data gap that must be addressed to incorporate climtae change projections directly into existing commercial tools for power system modeling. This paper presents an innovative approach to create hourly weather timeseries for future climates. A monthly quantile delta mapping technique is used to produce realistic hourly weather data for a future climate by adding the monthly climate change signal projected by climate models to historical weather data. This method preserves important, real-world characteristics from the historical record that are otherwise missing from climate model output, such as locationally-specific extremes which can be missing from coarse climate projections, natural variability which may not be well represented in the climate models, and important joint correlations among physically-linked variables such as wind, solar, and temperature. This approach has many potential applications in the power sector, including for capacity expansion and production cost modeling where select hourly timeseries are used for complex optimizations or simulations, as well as for resource adequacy assessments that evaluate large samples of realizations to identify possible extremes for stress-testing a future year of interest.

†Energy Systems and Climate Analysis Group, Electric Power Research Institute, Palo Alto, CA, USA. October 2024. Correspondence to: Erik Smith, esmith@epri.com

Introduction

Many power system modeling tools are designed to use hourly timeseries as input data. These tools typically use a representative sequence of hourly weather, often referred to as an 8760 timeseries, to capture the intra-annual conditions that influence power supply and demand, namely measures of effective temperature and renewable resource availability. This data are typically derived from historical meteorological data or synthetic profiles (e.g., a typical meteorological year). Increasingly, power system planners are interested in accounting for climatic trends and potential extreme events in the meteorological inputs to their simulation models (Craig et al., 2022). However, nearly all global climate model (GCM) projections are limited to daily temporal resolution, presenting an important data gap that must be addressed to explicitly represent the effects of climate change in power system planning. Filling this gap will enable more accurate planning of the clean energy transition, ensuring that the impacts of climate change are properly accounted for. Without this, relying solely on historical data could result in plans that fail to anticipate future conditions, with implications for resource adequancy, reliability, and energy costs.

Until hourly climate model projections are widely available, an approach that translates daily resolution information from climate models into hourly timeseries is needed for many forward-looking power system analyses. Currently, several approaches exist, each with their own advantages and drawbacks. GCM interpolation is one approach that estimates hourly values from aggregated daily statistics (Chow & Livermore, 2007), but these methods fail to capture important diurnal fluctuations that have material consequences on electricity demand. Figure 1 illustrates this interpolation approach for the temperature variable, which is typically reported from climate models with a daily maximum, mean, and minimum value, and solar irradiance and wind speeds, which are limited to a single daily average value. Dynamical downscaling approaches use a higher-resolution climate model called a regional climate model (RCM) to increase the spatial and/or temporal resolution of a lower-resolution GCM (Tapiador et al., 2020; Wang, Hamann, Spittlehouse, & Carroll, 2016). However, this approach requires significant resources and expertise to execute (Wood, Leung, Sridhar, & Lettenmaier, 2004). More recently, machine learning has been applied to produce future hourly time series. For instance, Buster, Benton, Glaws, & King (2024) developed a deep learning approach to spatially and temporally downscale climate models based on multiscale relationships in the historical record. Other approaches have been developed, such as a weather classification approach that uses k-nearest neighbor clustering to find the most analogous hourly historical weather sequence given daily GCM data (Hosseini et al., 2021). A limitation of machine learning is its tendency to overfit to historical observational data, constraining their generalizability to future conditions simulated by GCMs. Their relative lack of transparency also necessitates rigorous validation against established downscaling methods to ensure their predictions of climate variability are transferable and reliable.

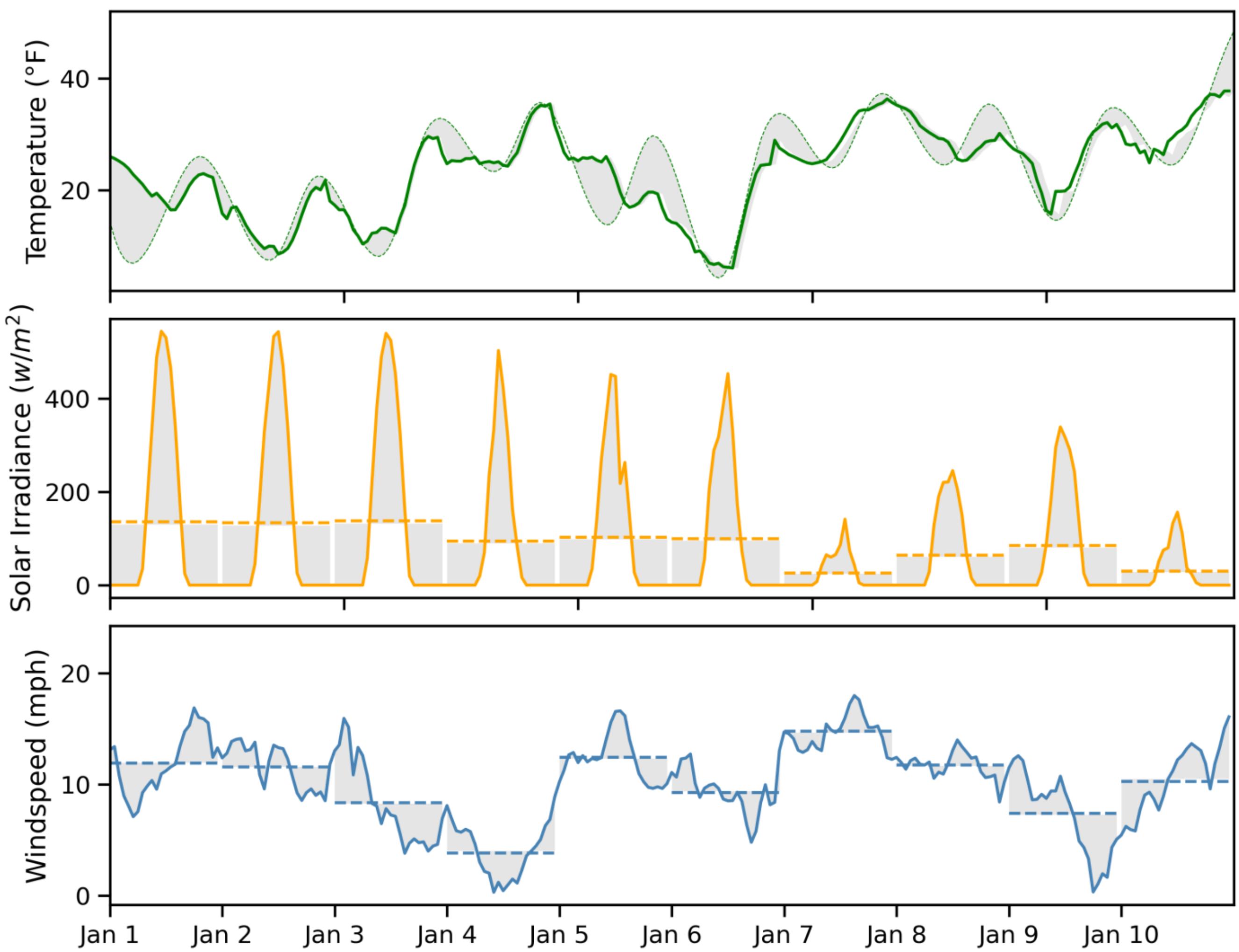

Figure 1: Top - hourly temperature (top) vs daily max and min values from Jan 1-10, 1940, for Nashville, TN. Max and min values are placed at the climatological coldest and hottest part of the day during January and then fit with a spline function. Middle – hourly solar irradiance (solid line) vs daily mean values (dashed line). Bottom – hourly windspeed (solid line) vs daily mean values (dashed line). Gray shading highlights differences between the mean, or max and min, and the hourly values.

Another method that offers a simple yet robust approach to producing future hourly time series is delta mapping. Though typically used for bias correction, delta mapping can be used to scale historical data according to a projected climate change signal. More specifically, an adjustment, or delta, can be calculated using the change in statistical characteristics of a weather variable between a historical baseline period and a projected period from GCM simulations, then used to shift a historical timeseries so that it is more representative of a future climate. Several techniques have been developed based on this premise. Hiruta et al., 2022 utilized a simple mean shifting approach which involved adding the daily delta to a historical hourly time series. However, this carries the assumption that there will be a change in mean climate, but not a change in variability. A more advanced form of delta mapping is quantile delta mapping (QDM). The QDM method computes deltas for different quantiles of the distribution, incorporating changes in variability such as differences in shifts in the upper and lower tails (Bloomfield et al., 2022). This approach is meant to produce realistic patterns of variability as observed in the historical record, but shifted to capture changes in mean, variability, and extremes in GCMs. However, it relies solely on the projected shift in the daily mean values to infer shifts on an hourly basis which may fail to capture changing diurnal patterns that matter to power system planning.

The QDM-based technique presented in this paper was first developed for bias correcting GCM data (Cannon, Sobie, & Murdock, 2015) and later added a multivariate quantile delta mapping (MQDM) approach to spatial bias-correction (Cannon, 2018), which adapted the N-pdf transform method outlined in Pitie, Kokaram, & Dahyot (2005). Here we present a temporal application of this MQDM approach, repurposed to shift historical hourly weather data to future years. This approach has four distinct advantages over existing approaches. First, we employ a multivariate shift of daily minimum and daily maximum temperature for a more advanced treatment of important diurnal features. This has the benefit of retaining realistic hourly variation critical for modelling power system supply and demand. Second, the climatological shift is conducted for each month, capturing changes in seasonality projected by GCMs. These month-specific changes provide valuable information for changes will have a significant impact on the cost-effective scheduling, operation, and maintenance of future power systems. Third, an important design feature of this method is the emphasis on relative changes projected by GCMs rather than absolute values: the climate model simulations are used as a statistically-representative climatology, not precise forecasts for a future year like 2050. This feature also reduces the impact of local to regional scale biases in GCMs (e.g., models that tend to run hot vs. cold, or wet vs. dry). Fourth, leveraging the historical record conserves much of the original physical link between variables, such as wind, solar, and temperature, and this synchronicity may be critical for power system modeling.

Additionally, this method can be used to expand the representation of annual variability in hourly timeseries, compared to current approaches that focus on a small number of representative years, because all historical years can be detrended and used to create an extended dataset with intrinsic variability over the length of the historical record, yielding a vast array (e.g., hundreds or more) of realistic hourly timeseries that can be used to represent a future climate under different warming scenarios and climate model formulations. Taken all together, this method combines key strengths of both historical and projection datasets.

Approach

We illustrate this method using an example location of Nashville, Tennessee, which experiences a wide range of temperatures, though this method for shifting historical meteorological timeseries to future years can be applied to any location with suitable data as described below.

Data

This method relies on two sources of climate data: one source of representative hourly data for the primary variable (i.e., temperature, as well as any synchronous variables of interest), and a second source for the distributional changes for the primary variable.

In this paper, ERA5 historical reanalysis 2-m temperature data from the European Center for Medium Range Weather Forecasting (ECMWF; Hersbach et al., 2020) is used as the former, the underlying reference data to be shifted to future years. Reanalysis data is advantageous in that it is spatially and temporally complete and often available for an extended historical period for multiple variables of interest to the power system such as temperature, humidity, wind and solar irradiance. ERA5 data has an hourly temporal resolution, a spatial resolution of 31 km x 31 km, and is available from 1940 to present. While ERA5 data is used in this study, other sources of hourly historical data such station observations are also appropriate. This step of reference data identification also establishes the baseline climatology for the exercise; the National Oceanic and Atmospheric Administration (NOAA) defines the current 30-year climate normal period as 1991 – 2020 and that is used as the reference period here.

For the latter source of distributional changes between the reference and the future climate period of interest, projections of daily maximum and daily minimum 2-m temperature are used from five Coupled Model Intercomparison Project phase 6 (CMIP6) climate models. Specifically, a subset of five GCMs featured in the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP; Warszawski et. Al., 2014) are used as these models were preselected to span a range of equilibrium climate sensitivities and the spatial grids and calendars of the raw model data have been adjusted to be consistent across the five models. For the present study, two climate scenarios from each model, SSP1-2.6 (SSP126; hereafter lower emission scenario) and SSP3-7.0 (SSP370; hereafter higher emission scenario), are used to create 10 variants of potential future climate states (5 models x 2 scenarios), with a focus on two future climates: 2036 – 2065 (representing the climate of 2050) and 2071 – 2100 (representing the climate of 2085). This is a design choice for illustration here and the approach can be applied using other model-scenario permutations as well[1].

After the reference and future data sources and climatological periods have been identified, there are four steps that are used in the monthly quantile delta mapping (QDM) technique.

[1] ISIMIP provides data from a range of standardized climate warming scenarios including the CMIP6 Shared Socioeconomic Pathways (SSPs) and the CMIP5 Representative Concentration Pathways (RCPs).

Step 0 [Optional]: Expand the set of representative weather years

This step transforms additional historical weather data to align with the trends of the chosen baseline in order to increase the sample size of representative weather years. If increased representation of variability is a priority this step is needed; otherwise, if the selected baseline climatology (e.g., 1991-2020) provides sufficient variation then this step can be skipped.

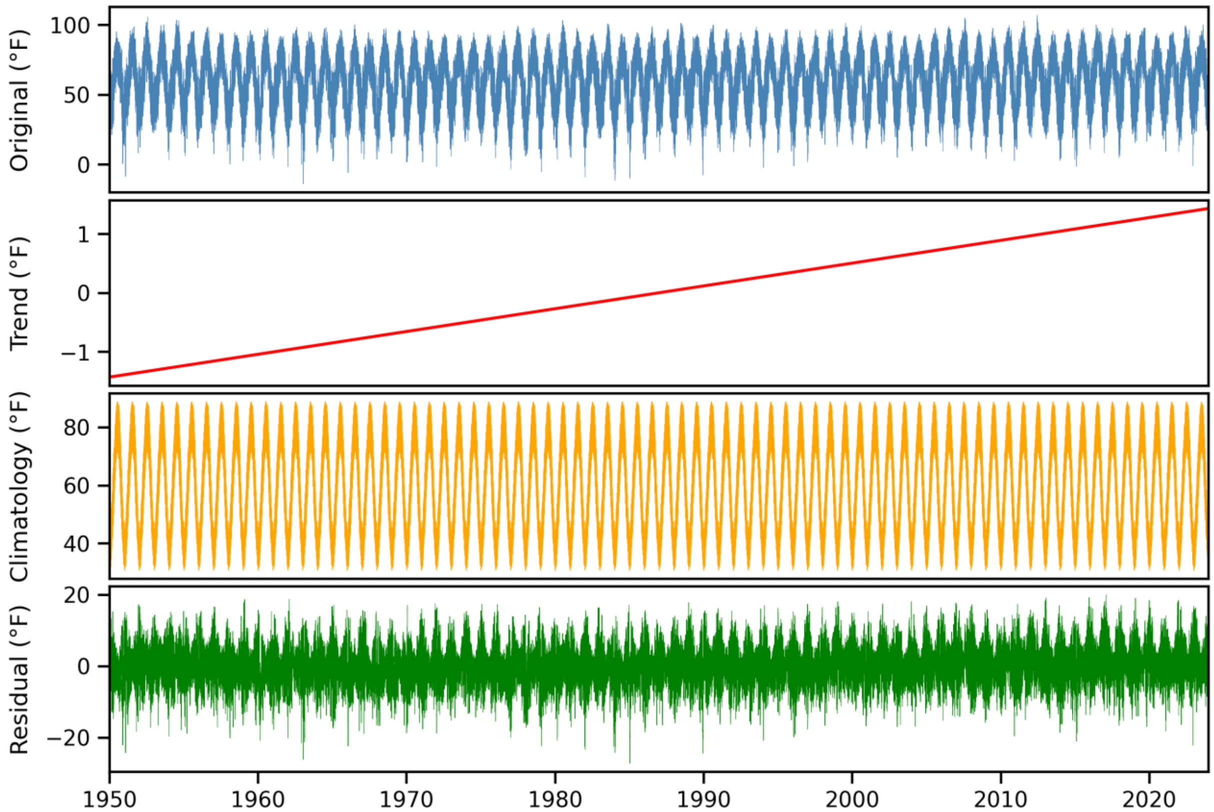

Figure 2: Timeseries decomposition for Nashville, TN.

The long-run historical data is first detrended to remove the warming signal. In this example, ERA5 data going back to 1940 is detrended using a first order polynomial detrending method. The timeseries is decomposed additively, which assumes seasonal periodicity is relatively constant over time (Cannon, Sobie, & Murdock, 2015). Once the trend and seasonality have been removed, only the residuals from the mean remain; thus the mean temperature must be added back to the timeseries. Each component of the decomposition is shown in Figure 2. In this example, the mean from the current climate normal period (1991 – 2020) is added back to the trend-less residuals and seasonal climatology to make all years similarly representative of the present-day climate. While recent years (e.g., 2010s) do not undergo a considerable shift, earlier historical years (e.g., 1940s) are effectively warmed to account for the historical warming signal (Figure 3). In other words, detrending the data and scaling to present-day imposes observed climatic trends onto the earlier candidate weather years to make them equally representative of the baseline climate period.

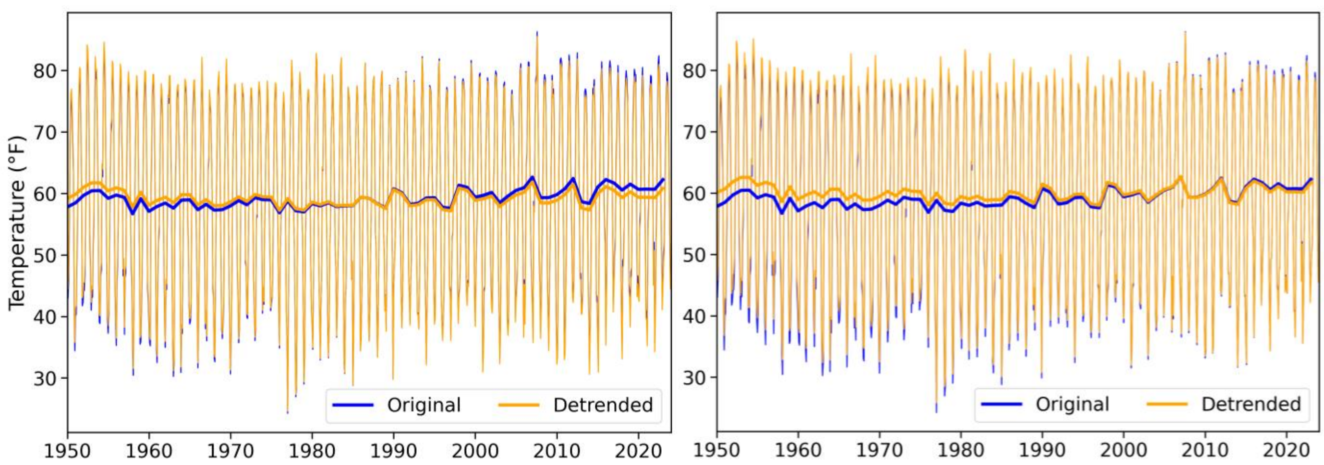

Figure 3: Left: detrending of 2-m temperature for Nashville, TN with the mean for the full historical period (1950 – 2023) used to rescale the timeseries. Right: detrending of 2-m temperature for Nashville, TN with the mean for the current climate normal period (1991 – 2020) used to rescale the timeseries.

While it's not always feasible to use the full historical period of data, representation of natural variability is important and the ability of a candidate dataset to capture the range of variability should be considered. One way to determine the number of years needed to capture natural variability is to assess the empirical cumulative distribution function (CDF) of multiple candidate time periods, to identify the cumulative error between the available long-run distribution and the selected representation. (Note that these timeseries should first be detrended so the variability of the data is primarily natural variability as opposed to natural variability + external forcing (warming trend)). Figure 4 illustrates how a single historical year (e.g., 2023) does not capture the full distribution of the 74 (or 84) year period, especially with respect to a significant portion of the cold extremes as

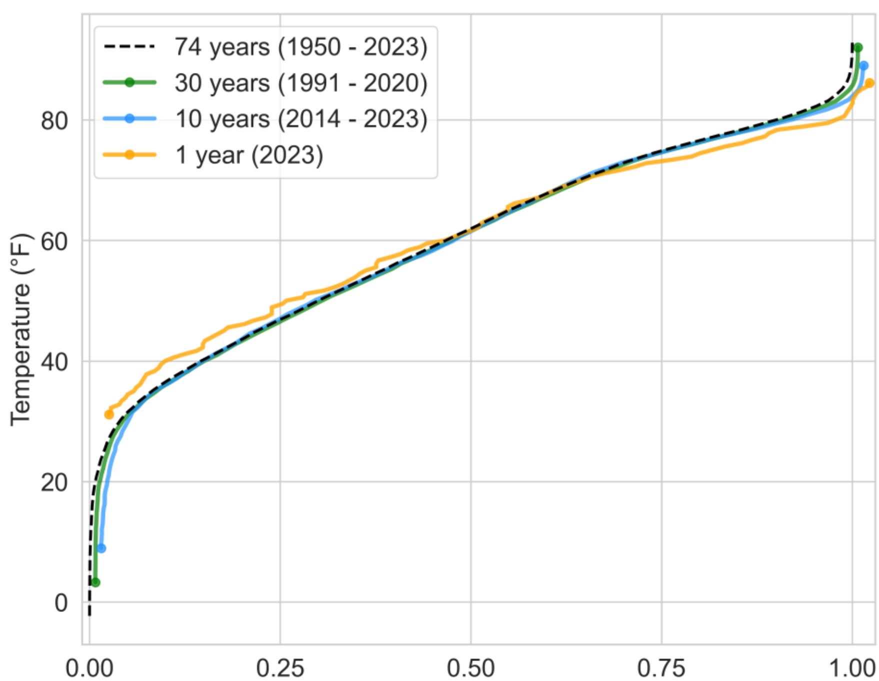

Figure 4: Daily mean detrended CDF for Nashville for 1-, 10-, 30-, and 74-year timeseries. Note, the cold tail of the 74-year CDF is still colder than the 30-year CDF because the 1970s and 1980s are not accounted for and had some of the coldest events in the eastern US.

well as much of the hot extremes observed in longer timeseries. While it doesn't capture the full range of temperatures, the 30-year timeseries (Figure 4, green line) aligns well with the full 74-year timeseries of temperature; this is consistent with the standard climatological definition of a climate "normal" as the three-decade average of meteorological variables.

Step 1: Calculate the projected shift in the climate variable

Central to this approach is the ability to leverage limited (daily) outputs from climate models to characterize distributional shifts that can be applied to the reference data in a way that preserves the monthly, non-stationary shifts that are observed in climate model projections. Most climate models report daily average values over their projection horizon, and in the case of temperature, daily maximum and daily minimum. The main approach presented here treats the daily maximum and daily minimum variables independently (Figure 5); step 2 discusses the multivariate approach that considers projected changes in the joint distribution among multiple variables.

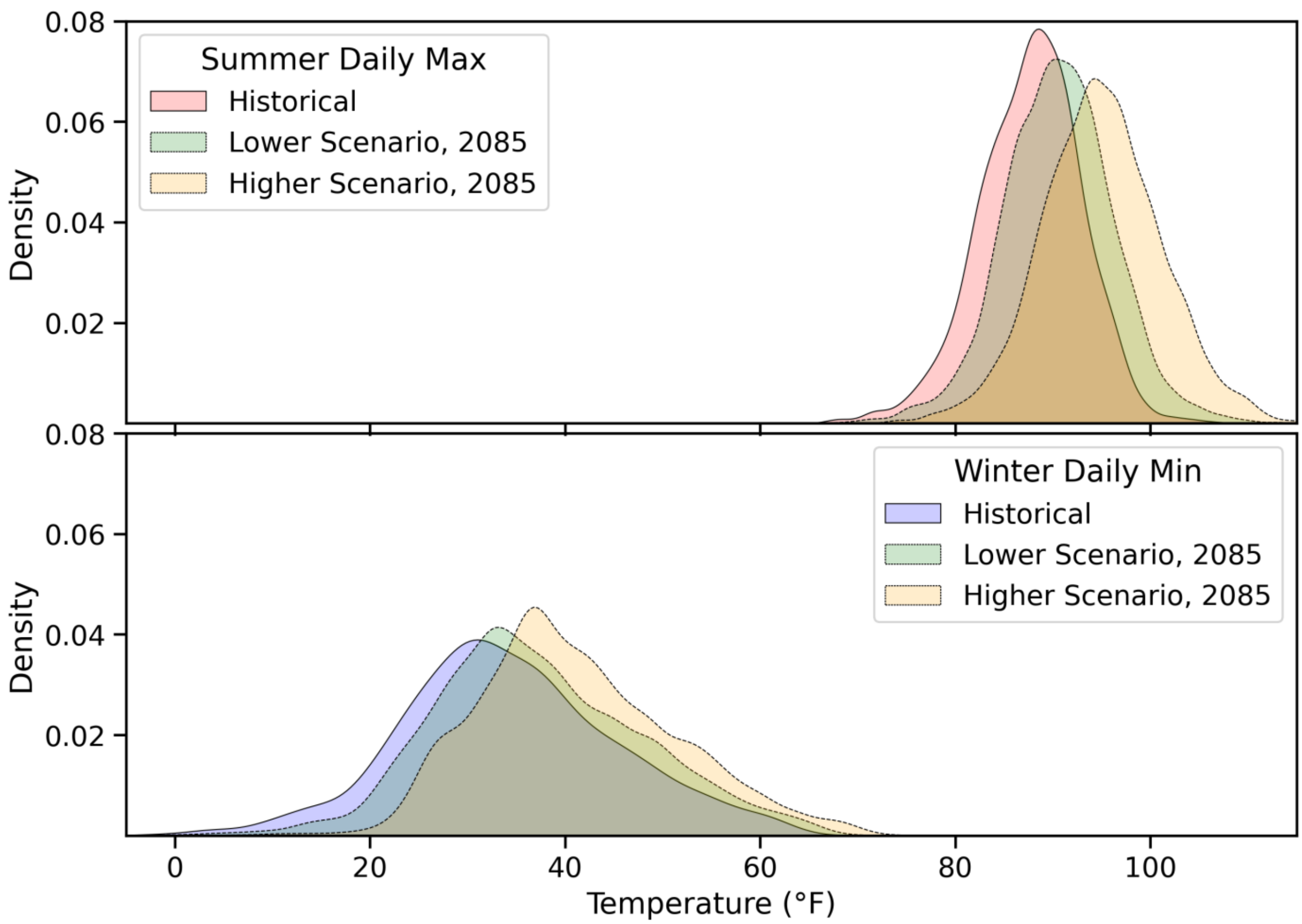

Figure 5: Illustrative non-stationary distributional shift in the detrended daily minimum temperature (left) and daily maximum temperature (right) for the baseline climate (2005) and the projected climate of 2050. The projections include all 5 models.

First, projected changes in the daily maximum temperature from the climate models are used to univariately shift the daily maximum temperature values in the reference data using the quantile delta mapping method outlined by Cannon, Sobie, & Murdock, (2015) and applied on a monthly basis. This is done by calculating the quantile-specific difference (i.e., delta, or anomaly) between the model's simulated historical distribution and its future distribution for each month. This univariate process is then repeated for daily minimum temperatures. Deltas are computed at regular intervals on the distribution (e.g., 20-100 quantiles) each month to account for seasonality and non-stationarity. As shown (Figure 5) for 2085, the extreme quantiles (e.g., 0.01 and 0.99) can have different deltas than the central quantiles, and as expected, the deltas tend to be larger for the higher climate scenario (green shading) than the lower climate scenario (orange shading). While larger deltas in the extreme quantiles may be robust, more testing should be done to explore the implications of quantile specification (e.g., appropriate number and coverage) depending on the sample size of the dataset.

Step 2: Multivariate Transformation [if shifting 2 or more variables]

If shifting more than one variable, the correlation between the two variables should be considered. Because daily maximum and daily minimum temperatures are used to create a future hourly temperature timeseries, they are treated multivariately. More specifically, a multivariate treatment is needed to ensure the distributions are being shifted consistently and so that there are no instances where daily minimum temperature is warmer after the shift than daily maximum temperatures. While this multivariate step can be used for variables other than just daily maximum and daily minimum temperature, the impact of this statistical process on the physical relationship between variables should be further explored.

While QDM does not emphasize the dependence between variables, the multivariate relationship can be preserved by iteratively transforming the N-dimensional probability density function (pdf) of the variables in which the univariate QDM was applied in the previous step (Cannon, 2018). First, daily maximum and minimum temperature variables are combined into a single array and standardized together, then an N-dimensional pdf transformation, which randomly rotates the arrays in the variable space and applies a univariate adjustment before rotating it back, is conducted. This process results in an array's joint distribution converging toward the target's joint distribution when a large number of iterations is done. The result of this adjustment is a rank structure which is then reordered to obtain the final bias-adjusted series. This process, which builds on the method created by Pitié et al. (2005) and is discussed in further detail in Cannon (2018), effectively corrects the lost dependence between variables in the univariate shift.

Step 3: Apply the deltas to reference daily maximum and daily minimum independently

These month-specific deltas, created in step 2 and modified to account for dependence in step 3, are then applied additively to the corresponding quantile of daily maximum temperature values in the reference data, hence “quantile mapping” (Appendix A; Cannon, Sobie, & Murdock, 2015). This process is then repeated independently for daily minimum temperature. The result is two shifted daily maximum temperature timeseries and daily minimum temperature timeseries. The daily maximum temperature deltas for each model and scenario across each month are illustrated in Figure 6.

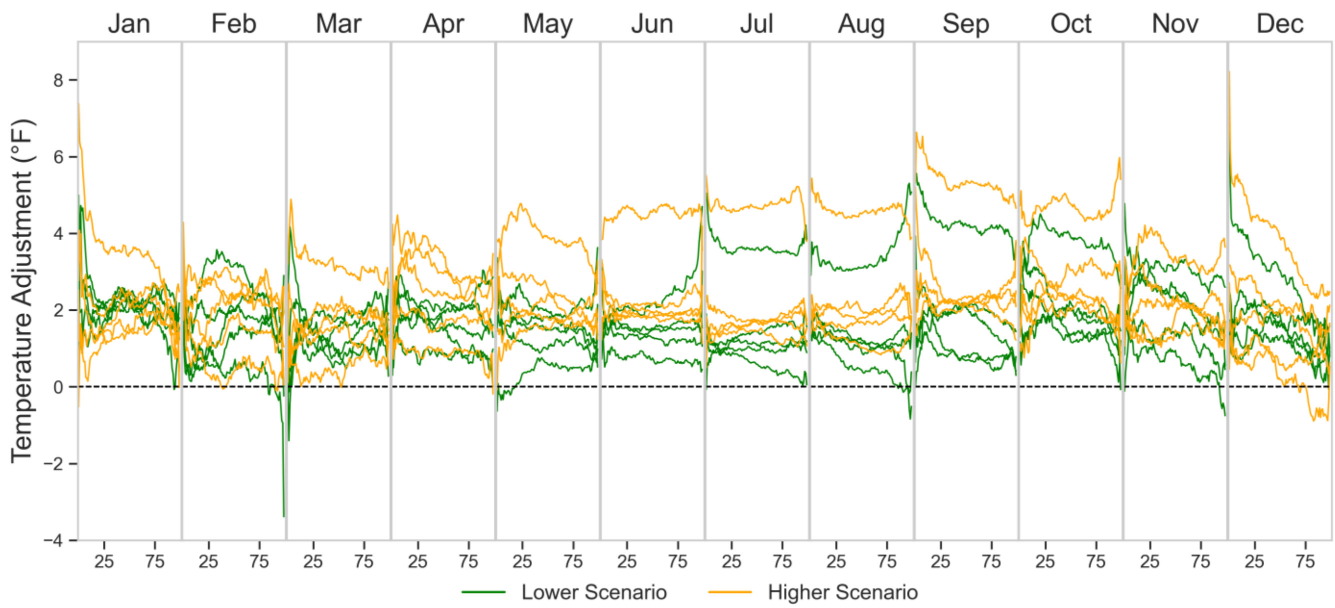

Figure 6: Within-model quantile-specific monthly difference between the baseline climate period (1991-2020) and the 2050 climate period (2036-2065) shown for 5 GCMs and two warming scenarios for Nashville. The x-axis represents the monthly distribution of maximum temperature with 100 quantiles per month.

Step 4: Convert the shifted daily maximum and minimum temperatures to hourly data

To interpolate the diurnal profile of the reference timeseries, scaling factors are used to match the pattern of the other 22 hours between the shifted daily maximum and daily minimum temperatures of each day. The scaling factors are calculated as the fractional value of the difference between reference day’s maximum temperature and the subject hourly temperature divided by the same reference day’s diurnal temperature range (DTR). This results in scaling factors that can be multiplied by the shifted DTR timeseries to produce an appropriately scaled hourly future timeseries as shown in Figure 7. Not only does this technique preserve the diurnal pattern of the reference meteorology but it also preserves the intra-annual chronology from a “real-world” benchmark timeseries.

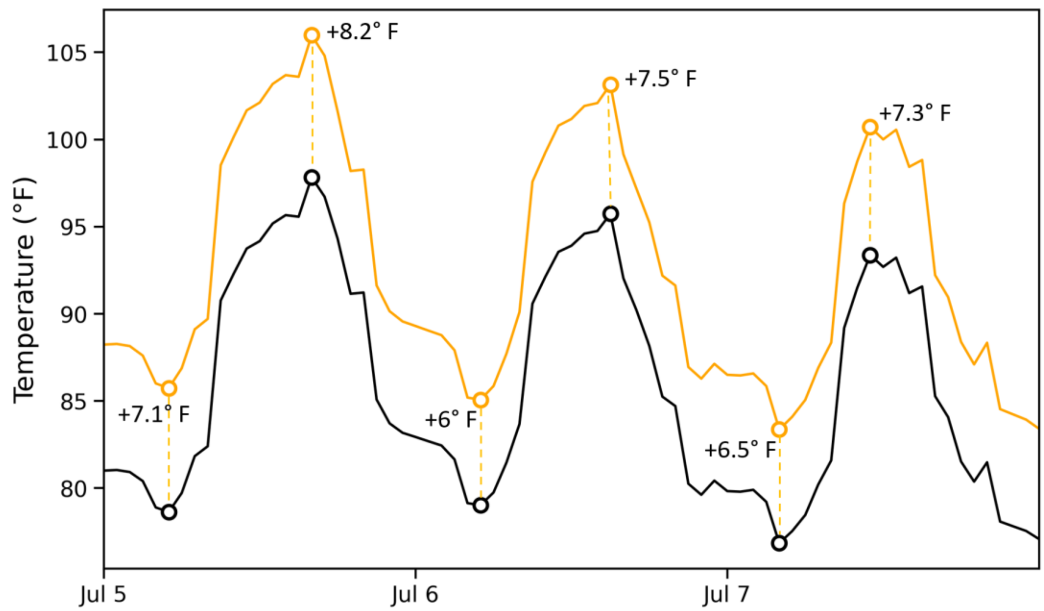

Figure 7: Example of new timeseries created with scaling factors. Data is July 5 – 7, 2022 for Nashville, TN. [This figure can be further annotated to show the DTR of the reference timeseries and the shifted timeseries and how an interim hour, say midnight, might be interpolated.]

Discussion

Once this methodology has been applied across variables of interest, the resulting future hourly timeseries retains the geophysically synchronous and chronological properties of the historical data, but with a seasonally varying non-stationary shift based on individual climate models and scenarios. This preserves realistic weather variability in a way that is forward-looking and consistent with the magnitude and higher-order dimensions of climate model projections. The relatively large spread that is visible across the 10 future climate realizations provides important representation of the uncertainty in possible futures (Figure 8). While the hourly climatology for the two climate scenarios is largely similar when shifted to 2050, the difference is generally larger when shifted to 2085 (Figure 8a). There is also considerable spread across models (Figure 8b), thus multiple models should be considered to account for the model uncertainty. By shifting the historical hourly data to the climate of 2050, the shape of the annual climatology is largely maintained, though there is some seasonal variation in the amount of change. This is consistent with how different climate scenarios are projected to diverge over time, with warming under lower climate scenarios leveling off around mid-century and warming continuing under higher scenarios through the end of the century (Lehner et al., 2020).

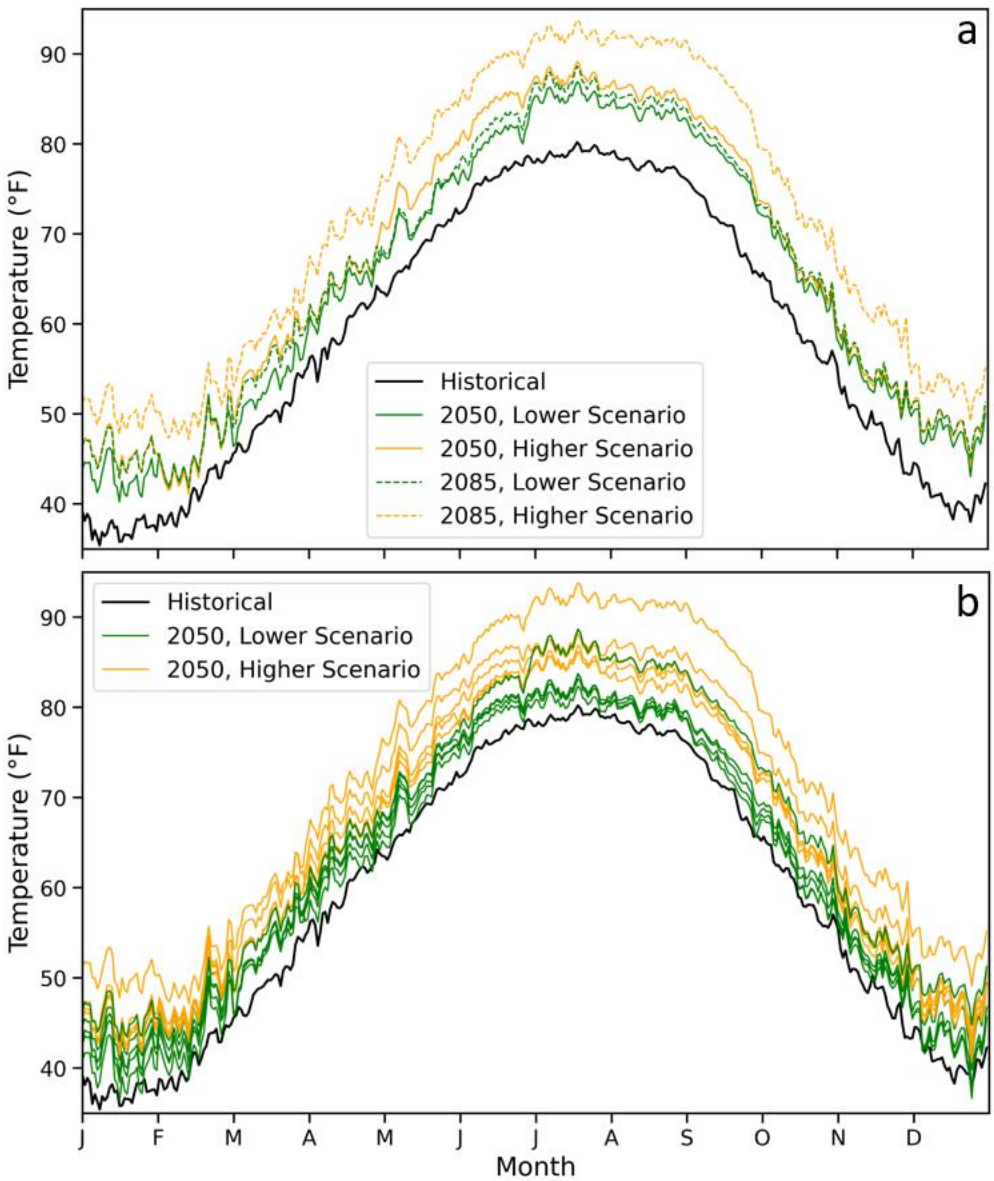

Figure 8: a.) Historical (2005 climate) and shifted daily mean temperatures for a single model (UKESM) and 2 climate scenarios for 2 future climate years (2050 and 2085). b.) Historical (2005 climate) and shifted daily mean temperatures for the 2085 climate corresponding to projections from 5 models and 2 climate scenarios.

Visualizing the shifted dataset during selected events of extreme temperatures helps demonstrate the utility of this method for developing forward-looking data to use in modeling the power system during moments of potential stress (Figure 9). During the December 2022 cold air outbreak in Nashville, temperatures reached -1°F the morning of December 23rd (panel a, b). Shifting the detrended event timeseries (which also shows a minimum temperature of -1°F) with a single model, the resultant low temperature on December 23rd is found to be 1°F. In other words, if this weather event were to "re-occur" in the climate of 2050, the event would generally be +2°F warmer. Like extreme cold, extreme heat is projected to increase in a warming climate, adding new levels of stress to the power grid during the summer season. The June 2012 heatwave in Nashville is generally warmer in the climate of 2050 (panel c, d); while not every hour of the day is warmer than the historical timeseries, the nighttime low temperature and daytime high temperatures are as much as +4°F warmer. More specifically, the warmest temperature is shifted from 107°F to 108°F.

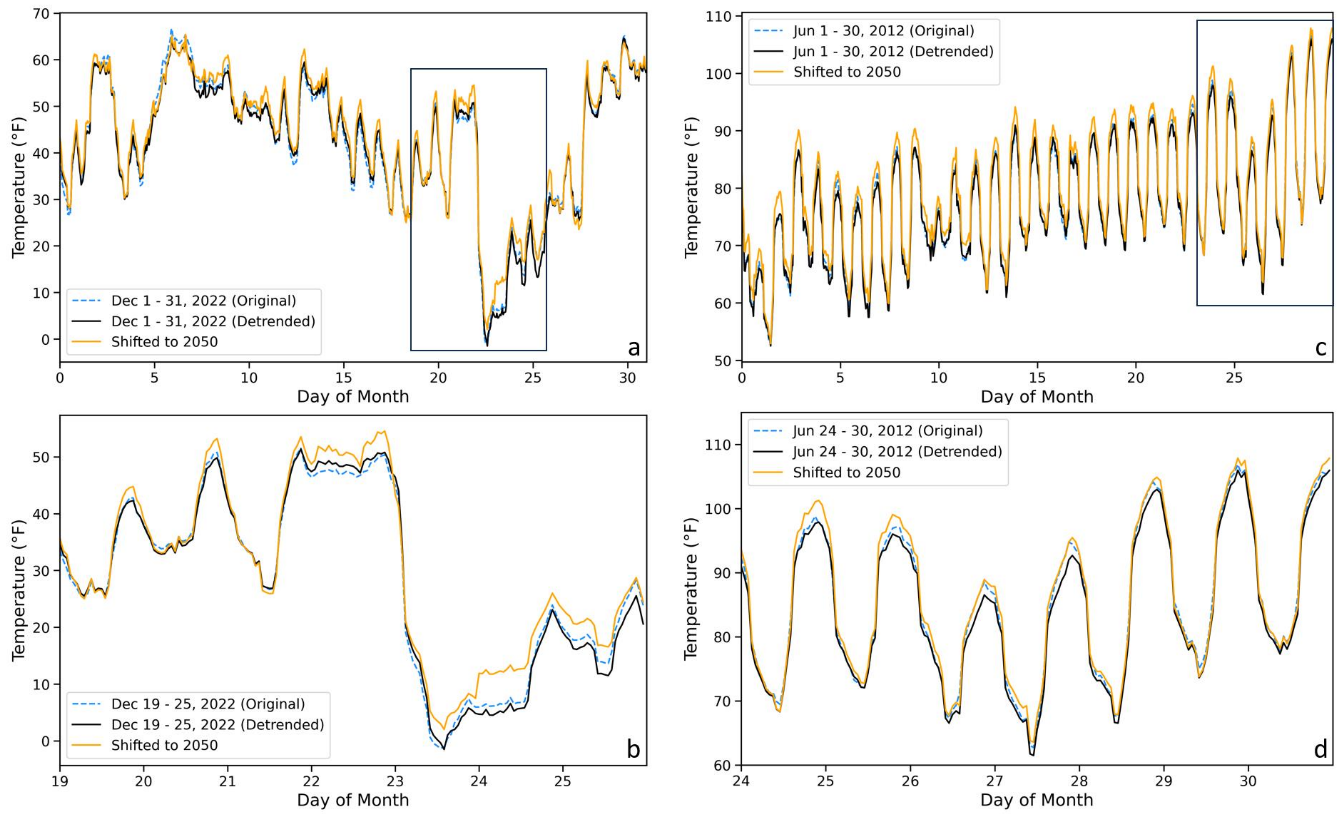

Figure 9: Example extreme temperature events for Nashville, TN showing the historical reference ~~(both original and detrended)~~ and the MQDM-shifted timeseries for the 2050 climate corresponding to projections from a single model (GFDL) and scenario (SSP370, higher climate scenario). a.) December 2022, b.) 7-day window from December 19 – 25, 2022 to capture an extreme cold event, c.) June 2012, d.) 7-day window from June 24 – June 30, 2012, to capture an extreme heat event. Note, time of day is GMT.

The two shifted extreme events will be different across different models and different scenarios. The GFDL model, for instance, tends to be cooler than other models and generally shows less warming, or smaller shifts in the timeseries (Figure 10). The shifted distributions can be compared to the detrended historical distributions to get a better idea of how this method effects the timeseries cumulatively.

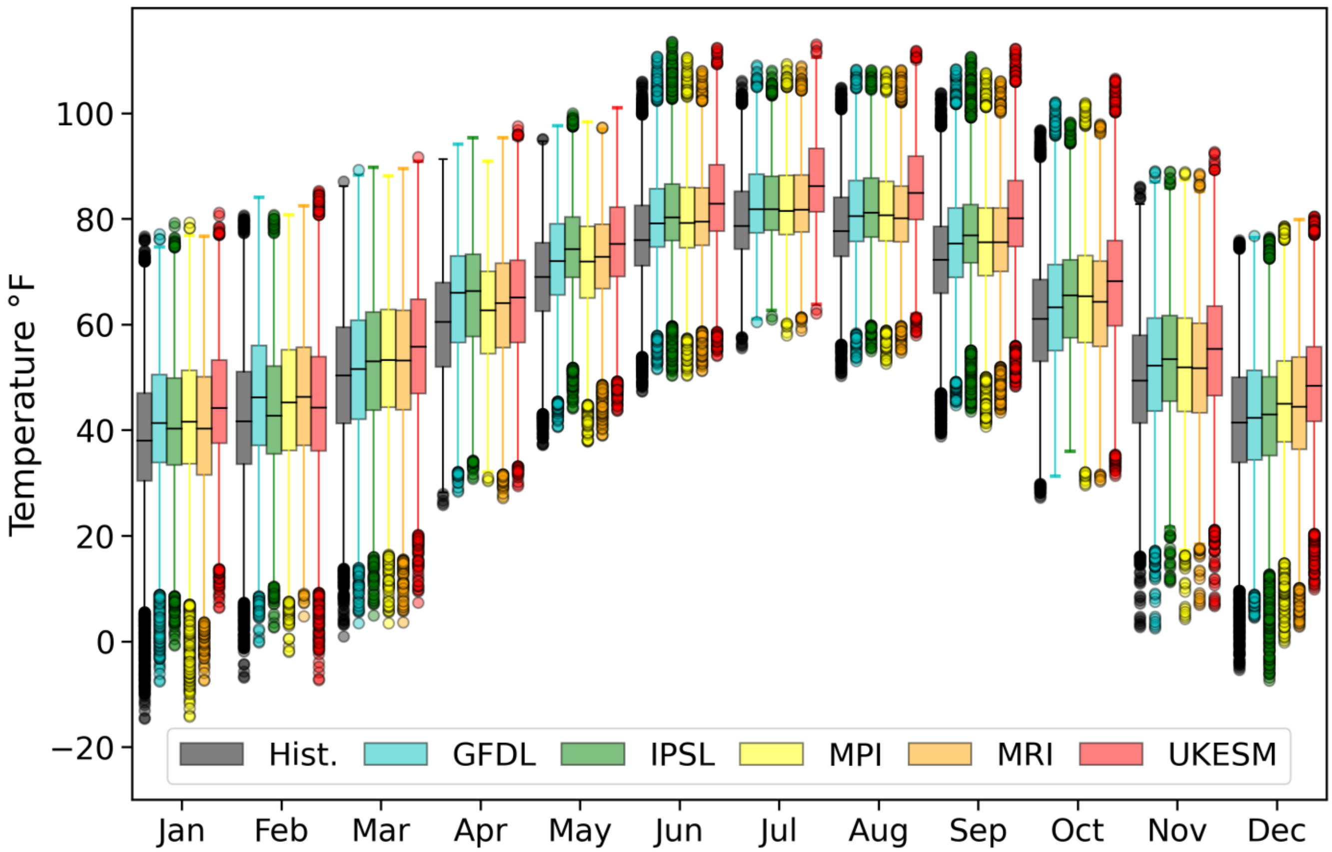

Figure 10: Boxplot of detrended historical hourly timeseries (1950 - 2023, gray) and the full reference timeseries MQDM-shifted to the 2085 climate for each of the five models under a higher climate scenario. Boxplot whiskers show the 1st to 99th percentiles, with dots showing outliers.

To convey the changes in key meteorological statistics between the reference and shifted hourly timeseries more directly, six benchmark temperature metrics are shown in Figure 11. For both the mean temperature and 1st percentile, the projected shift is at least 2°F or more from the historical baseline, with a larger shift in the higher climate scenario relative to the lower climate scenario. The shift ranges from 2°F in the coldest model in the lower scenario, to 6.5°F in the warmest model for the higher climate scenario. The projected shift in the 99th percentile is more variable than the mean and 1st percentile, with ranges from 0.5°F, in the coldest model in the lower scenario, to nearly 10°F in the warmest model for the higher climate scenario. Shifts in the most extreme values, which are the record minimum and record maximum temperatures for the full historical and shifted periods, show similar variability across models.

For the record minimum temperature, under the lower climate scenario (green points) all models project a range of warmer values, while for the higher climate scenario the projected period minimum temperature is tightly grouped with one single model even projecting a minimum temperature lower than the historical record. In this case, the GFDL model, which is generally a colder model over the continental United States, has a negative delta for the 1st quantile (refer to Figure 6). This negative delta leads to a decrease in the 1st quantile of the shifted timeseries. This outlier is somewhat counter-intuitive but should not be interpreted as a result that extreme cold is expected to become more frequent or more intense, rather the majority of models for both scenarios show a warming of the most extreme cold temperatures (Blackport, Fyfe, & Screen, 2022; Smith and Sheridan, 2020).

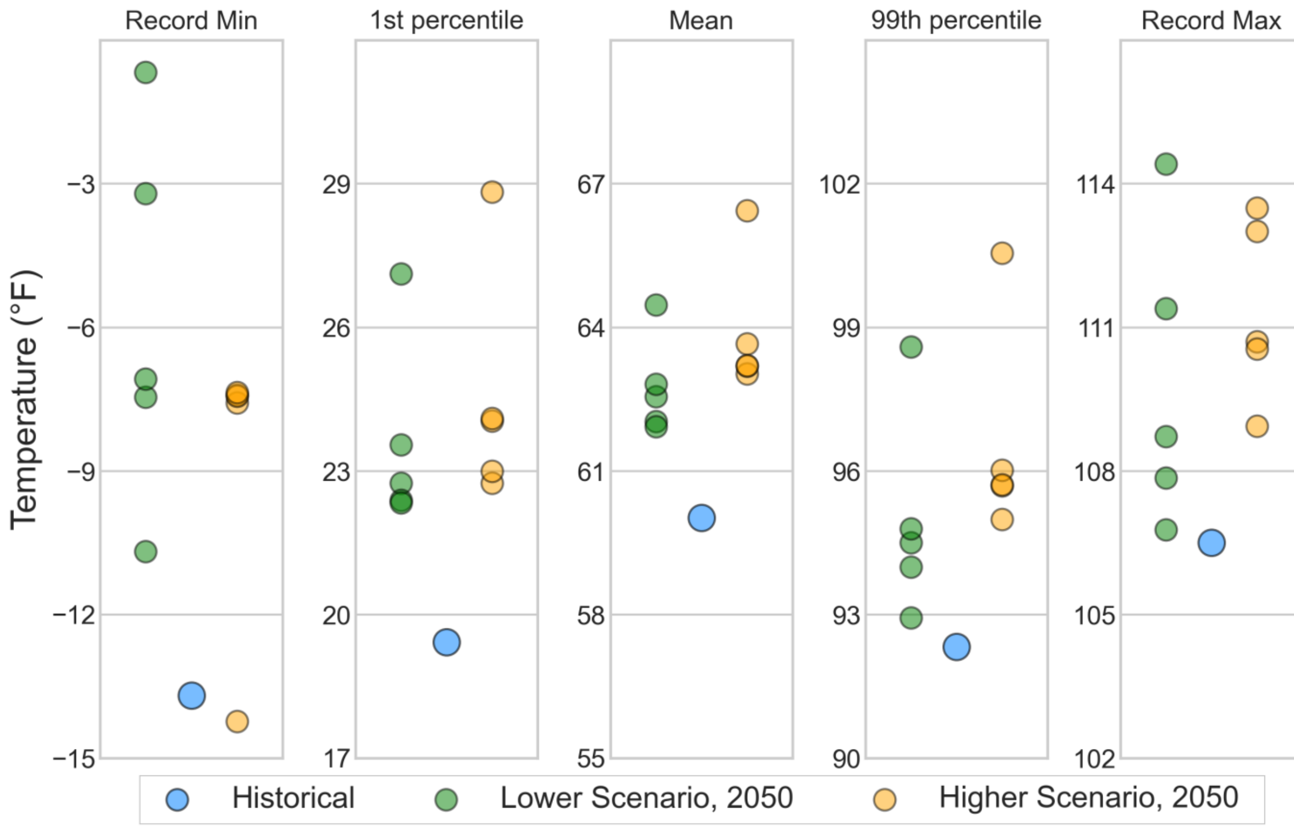

Figure 11: Changes in five temperature metrics for two climate scenarios for Nashville, TN: 1.) mean temperatures, 2.) 1st percentile of hourly temperatures, 3.) 99th percentile of hourly temperatures, 4.) record minimum temperature, or coldest temperature over the period of record in the hourly data, 5.) the record maximum temperature, or warmest temperature over the period of record in the hourly data.

Representation of discrete extreme events

Another way of data validation directly addresses the importance of representation of discrete extreme events that could cause stress to the power system. Here we provide a comparison of the future-shifted (multi-model ensemble median) and raw ISI-MIP CMIP6 (multi-model ensemble) annual number of hot and cold days, where hot (cold) days are defined as the 95th (5th) percentile of maximum (minimum) temperature over the 1991-2020 reference period (Figure 12).

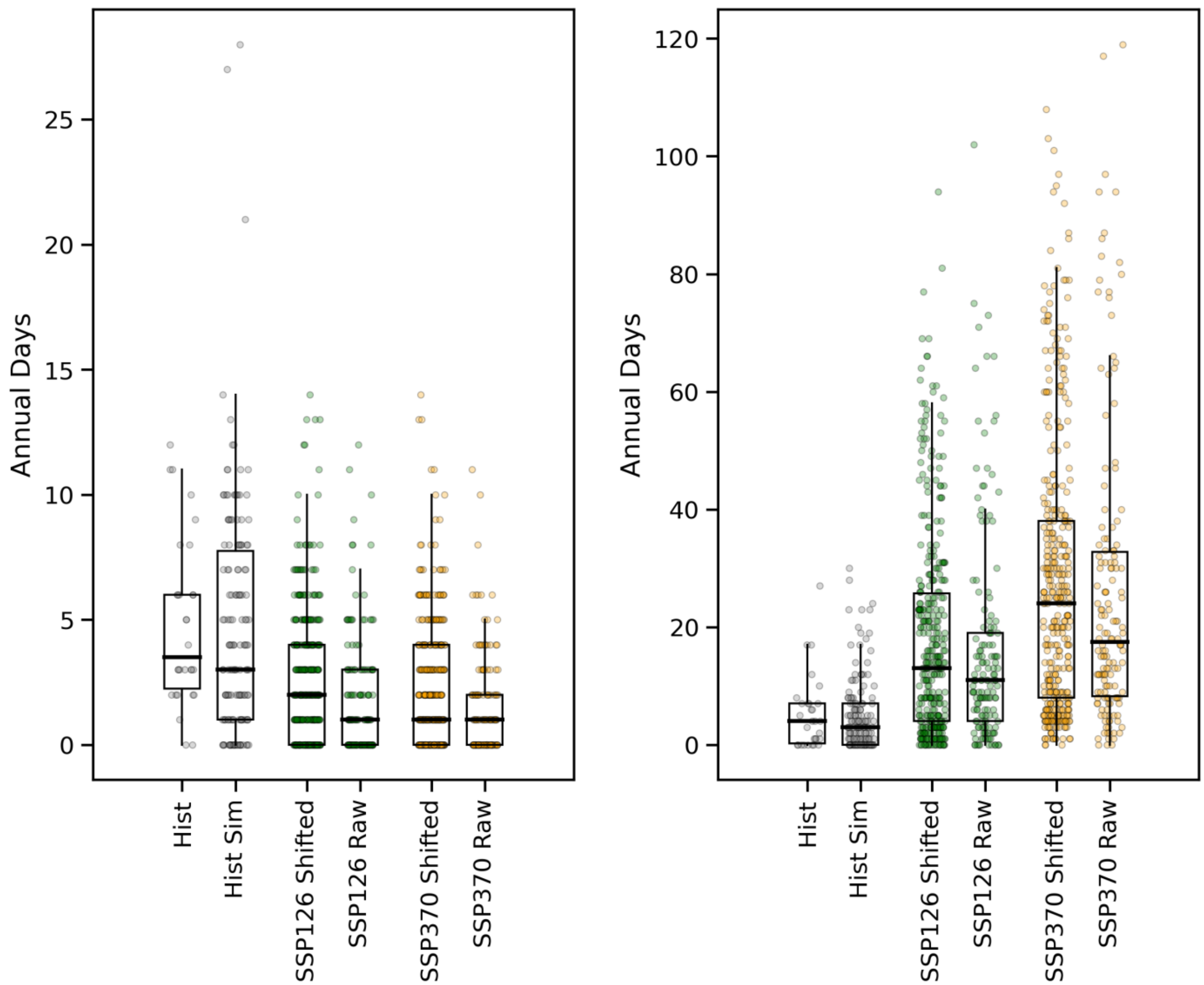

Figure 12: The number of annual cold days (left) and hot days (right) across datasets. Gray values show the historical frequency of cold and hot days in the ERA5 data (Hist) and model simulations (Hist Sim). Green values show the frequency under a lower climate scenario (ssp126) for 2050 in the shifted (SSP126 Shifted) and raw climate model (SSP126 Raw) data. Orange values show the frequency under a higher climate scenario (ssp370) for 2050 in the shifted (SSP370 Shifted) and raw climate model (SSP370 Raw) data. A boxplot is overlaid on each of the scatterplots to better depict the distribution. The scatter is randomly jittered for each category. Due to the fact that the historical period ends in 2014 in the CMIP6 models, we use SSP1-2.6 to provide data for 2015-2020.

Overall, yearly hot and cold day counts within the MQDM-shifted timeseries correspond well with the raw ISI-MIP GCM output. As expected, the median number of hot days increases in 2050 climate conditions and across emissions scenarios, while the median number of cold days decreases in 2050 climate conditions. For this reason, there is an argument to be made to retain historical extreme cold events as conservative future planning scenarios, as there is ready access to many relevant data streams compared to simulations. Lastly, it can include historical years in future scenarios as a lower bound for risk assessments particularly concerned with the impact of extreme cold on the power system.

Finally, to investigate whether the chronological profile of heat events is comparably represented in the shifted future-climate time series, we compare the top three peak heat and cold events by plotting daily extreme temperatures for the two-week period surrounding the peak. We find that the MQDM-shifted and raw GCM events have similar magnitudes (Figure 13) and temporal profiles. This provides additional justification that the hourly shifted weather years maintain the magnitude of extreme events in the original (daily) climate model data.

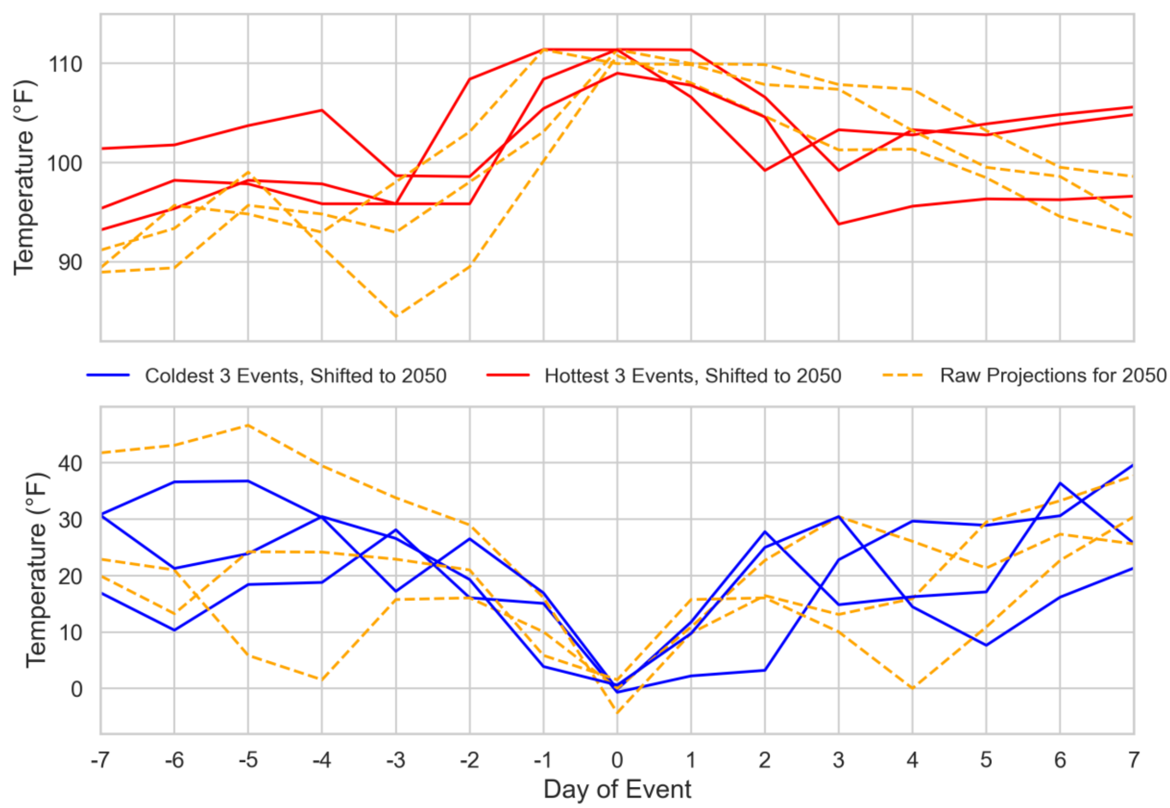

Figure 13: Timeseries of temperature extremes for the top three peak heat and cold events in the 2050 climate, where the MQDM timeseries (red for extreme heat and blue for extreme cold) is compared to the a single CMIP6 model output, IPSL (dashed orange). The week trailing and proceeding the event is shown to compare how the event unfolds in both the MQDM-shifted and raw GCM projections.

Considerations when applying to non-temperature variables

This paper has used the temperature variable to illustrate and discuss the MQDM method as adapted for hourly resolution power system applications, though it can be used for other variables as well. However, not all variables are projected to change significantly, and some variables, like wind speeds and solar irradiance, have considerable uncertainty in the projected changes. In the case of statistically insignificant or discontinuous changes, the user may choose to use the unshifted historical data. In the case of high uncertainty in the projected changes of certain variables or disagreement across models, users should take care in implementing this method, especially in a multi-variate context with temperature, wind and solar, as over-manipulation may violate physical relationships between correlated variables. Furthermore, if no historical trend is evident, the detrending step can be skipped to not introduce additional uncertainty.

In the example laid out in this paper, there is justification for shifting temperatures because there is a clear warming signal in the historical period (refer to Figure 2) and in the climate model projections (refer to Figure 6). Other variables, like solar irradiance and wind speeds, may not necessarily need to be shifted. Thus, the resulting forward-looking dataset with synchronous hourly timeseries of temperature, wind and solar would include 10 variants of shifted temperature data (reflecting 5 climate models and 2 warming scenarios) each paired with a single wind and solar realization (the reference timeseries). While there is uncertainty introduced when shifting only a single variable in a multivariate dataset, in terms of preserving the physical relationship between the variables, the added uncertainty is likely regionally dependent. More research would be needed to determine the amount of uncertainty introduced as well as if the added uncertainty is comparable to the components of uncertainty in raw ISI-MIP climate model projections (Lehner et al., 2020).

Shifting atmospheric moisture variables

A warmer atmosphere will be capable of holding more moisture, thus for many locations, the increase in atmospheric moisture may need to be accounted for when applying this method. However, atmospheric moisture is quantified with both absolute (e.g., dewpoint) and relative (e.g., relative humidity) variables, thus if the relationship between temperature and atmospheric moisture is projected to remain constant in the future, then relative moisture variables may be leveraged to shift absolute atmospheric moisture variables. For example, unshifted relative humidity may be used alongside the shifted temperature values to calculate the shifted absolute atmospheric moisture metrics, like dewpoint or wet-bulb temperatures. More specifically, once the shifted temperature (i.e., dry-bulb temperature) values have been computed, the shifted dry-bulb temperature and unshifted relative humidity values can be plugged into an equation to compute shifted dewpoint and/or wet-bulb temperature for each climate model and scenario. This is a straight-forward process as the time dimension of the unshifted relative humidity values will align with the shifted dry-bulb temperature values.

However, to justify this approach, it must first be shown that the relationship between temperature and moisture is projected to remain fairly constant. In other words, the absolute moisture content should be projected to change at a similar rate as temperature such that the relative humidity is not projected to change. While it is highly unlikely that relative humidity will not change at all in future climate model simulations, even if just due to natural variability, the changes may be negligible. Significance testing, such as Theil-Sen slope estimation of daily or annual mean relative humidity, can be used to inform this decision. The fractional contribution of the change in relative humidity to absolute atmospheric moisture can also be used. More specifically, if the relative humidity is projected to increase by 2%, but that change only contributes to a small increase in the wet bulb calculation (e.g., 0.1°F), then perhaps this increase in relative humidity can be ignored. While this method of using the unshifted relative humidity and shifted dry-bulb temperature to shift absolute atmospheric moisture metrics (e.g., dewpoint and wet-bulb temperature) will introduce caveats, caveats will also be introduced if using the quantile-delta mapping method to shift atmospheric moisture. There is inherent subjectivity introduced, but multiple steps can be taken to limit the subjectivity and create informed projections of both temperature and atmospheric moisture, and potentially other climate variables.

Conclusion

While climate model projections lack the hourly resolution that is often necessary for planning in the energy sector and beyond, they can be leveraged to create realistic synthetic future timeseries that preserve the critical characteristics of the data, namely the trends and extremes. There are several benefits to doing this in that it captures real-world variability from the historical record and creates a large sample of realistic climate-adjusted profiles. This method also conserves much of the original physical link between synchronous meteorological variables which is critical as many extreme weather hazards manifest via multiple variables. Historical data sources have the advantage of more complete coverage of secondary variables like water temperature and stream flow, which are often missing from climate model projections.

There are limitations to this approach that should be considered. Firstly, this method is best suited for single points as the impact on spatial coherence by shifting each grid point individually has not been assessed. Before applying this method independently to contiguous grid cells, more research is needed to determine if the relationship between neighboring locations is sufficiently preserved when the quantile-delta mapping is conducted at the individual cell-level. Alternative approaches for regional application could explore the joint distribution of surrounding cells to better reflect the spatial relationship. Another limitation may be introduced when shifting mean variables. While climate models generally have daily maximum, mean, and minimum temperatures, they tend to only report other variables at a daily mean resolution. A univariate shift of the daily means has the drawback that extremes are not expected to shift uniformly with means, thus the tails of the shifted distribution may not be shifted appropriately. Lastly, to the extent that this approach is applied to individual variables (i.e., shifting one variable but not others), this may impact the relationship between variables (e.g., correlations). While these changes may be negligible, certain variables may have a higher interdependence, making a multivariate adjustment more appropriate. Additional research is warranted to quantify the uncertainty introduced by combining shifted variables with non-shifted variables for analysis or modeling purposes.

The detrending method presented in Step 0 is a relatively simple approach based on additive decomposition into linear trend, seasonal climatology, and residuals. This is just one of several possible methods that could be applied in this optional step. We acknowledge that the linear form includes the implicit assumption of stationarity, and plan to explore other detrending methods in future work, including the application of a QDM method for shifting historical years to the baseline climatological period. Ultimately, the detrending method can be the choice of the user.

Future work will focus on validation of these shifted timeseries profiles against the raw ISI-MIP climate models as well as other dynamical, statistical, and hybrid approaches described in the introduction. Specifically, a more detailed comparison of the statistical properties (e.g., variability and extreme event characteristics) of these synthetic profiles versus the raw ISI-MIP daily GCM projections, as well as a deeper dive into how well the synthetic profiles represent extremes like heat waves, cold events, and droughts, is needed. It will also be beneficial to explore alternative approaches to help identify when different methods are more appropriate for the application at hand. Some of these approaches are rather simple, like applying a fixed stationary adder (e.g., +5°F) to all historical 8760s to represent a future year when temperatures are projected to be 5°F higher on average, or temporal downscaling by linear interpolation between the climate models' daily min, max, and mean values to create hourly data. Others are more complex, such as the method outline in this article, or the use of regional climate models (RCM) initialized with daily climate model data to produce output variables with hourly resolution.

References

1. Alavoine, M., & Grenier, P. (2021). The distinct problems of physical inconsistency and of multivariate bias potentially involved in the statistical adjustment of climate simulations.
2. Blackport, R., Fyfe, J. C., & Screen, J. A. (2022). Arctic change reduces risk of cold extremes. Science, 375(6582), 729-729.
3. Bloomfield, H. C., Brayshaw, D. J., Deakin, M., & Greenwood, D. (2022). Hourly historical and near-future weather and climate variables for energy system modelling. Earth System Science Data, 14(6), 2749-2766.
4. Buster, G., Benton, B. N., Glaws, A., & King, R. N. (2024). High-resolution meteorology with climate change impacts from global climate model data using generative machine learning. Nature Energy, 1-13.
5. Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes?. Journal of Climate, 28(17), 6938-6959.

6. Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. Climate dynamics, 50(1), 31-49.
7. Chow, D. H. C., & Levermore, G. J. (2007). New algorithm for generating hourly temperature values using daily maximum, minimum and average values from climate models. Building Services Engineering Research and Technology, 28(3), 237–248. https://doi.org/10.1177/0143624407078642.
8. Craig, M. T. *et al.* Overcoming the disconnect between energy system and climate modeling. *Joule* **6**, 1405–1417 (2022).
9. Hersbach, H. *et al.* The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
Cannon, A. J., Sobie, S. R. & Murdock, T. Q. Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *J. Clim.* **28**, 6938–6959 (2015).
10. Hiruta, Y., Ishizaki, N. N., Ashina, S., & Takahashi, K. (2022). Hourly future climate scenario datasets for impact assessment of climate change considering simultaneous interactions among multiple meteorological factors. Data in Brief, 42, 108047. https://doi.org/10.1016/j.dib.2022.108047.
11. Hosseini, M., Bigtashi, A., & Lee, B. (2021). Generating future weather files under climate change scenarios to support building energy simulation – A machine learning approach. Energy and Buildings, 230, 110543. https://doi.org/10.1016/j.enbuild.2020.110543.
12. Jain, S., Scaife, A. A., Shepherd, T. G., Deser, C., Dunstone, N., Schmidt, G. A., ... & Turkington, T. (2023). Importance of internal variability for climate model assessment. npj Climate and Atmospheric Science, 6(1), 68.
13. Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E.M., Brunner, L., Knutti, R. and Hawkins, E., (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. Earth System Dynamics, 11(2), 491-508.
14. Loikith, P. C., Waliser, D. E., Lee, H., Neelin, J. D., Lintner, B. R., McGinnis, S., Mearns, L., & Kim, J. (2015). Evaluation of large-scale meteorological patterns associated with temperature extremes in the NARCCAP regional climate model simulations. Climate Dynamics, 45, 3257-3274.
15. Maraun, D. et al. Towards process-informed bias correction of climate change simulations. Nat. Clim. *Change* **7**, 764–773 (2017).
16. Pierce, D. W., Barnett, T. P., Santer, B. D., & Gleckler, P. J. (2009). Selecting global climate models for regional climate change studies. Proceedings of the National Academy of Sciences, 106(21), 8441-8446.
17. Pitie, F., Kokaram, A. C., & Dahyot, R. (2005). N-dimensional probability density function transfer and its application to color transfer. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 (Vol. 2, pp. 1434-1439). IEEE.
18. Smith, E. T., & Sheridan, S. C. (2020). Where do cold air outbreaks occur, and how have they changed over time?. Geophysical Research Letters, 47(13), e2020GL086983.

19. Tapiador, F. J., Navarro, A., Moreno, R., Sánchez, J. L., & García-Ortega, E. (2020). Regional climate models: 30 years of dynamical downscaling. Atmospheric Research, 235, 104785.

20. Vrac, M. & Friederichs, P. Multivariate—Intervariable, Spatial, and Temporal—Bias Correction. *J. Clim.* **28**, 218–237 (2015).

21. Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PloS one*, *11*(6), e0156720.

22. Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2014). The inter-sectoral impact model intercomparison project (ISI–MIP): project framework. Proceedings of the National Academy of Sciences, 111(9), 3228-3232.

23. Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic change, 62, 189-216.