www.nature.com/scientificdata

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

NASA Earth Exchange Downscaled Climate Projections 30 Arcseconds CMIP6

Bridget Thrasher[1] ✉, Sepideh Khajehei[2], John B. Kim[3] & Ian Brosnan[2]

We describe the latest version of the NASA Earth eXchange Downscaled Climate Projections 30 arcseconds (NEX-DCP30-CMIP6). The archive contains downscaled historical and future projections for 1950–2100 based on output from Phase 6 of the Climate Model Intercomparison Project (CMIP6). The downscaled products were produced using a daily variant of the monthly bias correction/spatial disaggregation (BCSD) method and are at 30-arcsecond horizontal resolution. Four variables (maximum temperature, minimum temperature, precipitation, and vapour pressure) from five CMIP6 experiments (historical, and the four Tier 1 scenarios: SSP1.2-6, SSP2.4-5, SSP3.7-0, and SSP5.8-5) are provided as available from thirty global climate models. The downscaled data cover the coterminous United States (CONUS), extending from approximately 24-50 degrees North and 235–294 degrees East.

Background & Summary

Originally released in 2013, the NASA Earth eXchange Downscaled Climate Projections 30 arcseconds archive[1,2] (NEX-DCP30) was developed to support the scientific community in the examination of climate change impacts over the coterminous United States (CONUS) at local scale. The 30-arcsecond dataset provided high-resolution climate change projections derived from the coarser resolution (1.5–2° resolution) global climate model (GCM) output from the Climate Model Intercomparison Project Phase 5 (CMIP5). Since then, results from the newer CMIP6 experiments have been published by the Earth System Grid Federation (https://aims2.llnl.gov/search/cmip6). While the CMIP5 simulations scenarios encapsulated representative concentration pathways (RCPs), CMIP6 simulations use updated scenarios known as shared socio-economic pathways (SSPs), as well as newer versions of GCMs. These updates have resulted in a strong desire for a new version of NEX-DCP30.

Here we present the latest version of the NEX-DCP30 archive[3] (NEX-DCP30-CMIP6), which will be useful for assessing trends in projected changes in climate over CONUS at a variety of spatial and temporal scales, as well as for driving impacts and adaptation models that might require higher spatial and/or temporal resolution than is available from other climate data products. It contains downscaled daily output of precipitation (pr), maximum temperature (tasmax), minimum temperature (tasmin), and estimated vapour pressure (vpr) at the land surface for 1950–2100 and complements the NEX-GDDP-CMIP6 downscaled archive[4,5], which offers global daily projections at a coarser spatial resolution of one-quarter degree. (See Table 1 for further description of the variables.) Of several downscaled datasets available for parts of North America, this dataset is unique in having the highest spatial resolution while using bias correction in the process, and is comparable with other downscaled datasets[6]. Monthly and annual aggregates are also made available for ease of use.

Methods

The daily BCSD statistical downscaling algorithm (a variant of the monthly BCSD method described in Wood, *et al.*[7,8]) is described in Thrasher, *et al.*[9] and Thrasher, *et al.*[5]. The overall workflow comprises pre-processing and bias correction, followed by spatial disaggregation. The method compares GCM output with an observation-based reference dataset over a common reference period, thereby adjusting the GCM projections toward more consistency with the historical climate record. In addition, the spatial detail provided by the reference dataset is used to interpolate the adjusted GCM outputs to a higher-resolution grid. Downscaling was performed individually for each of 30 published CMIP6 GCM outputs for four Tier 1 SSP scenarios (Table 2).

[1]Climate Analytics Group, Dayton, VA, 22821, USA. [2]Earth Science Division, NASA Ames Research Center, Moffett Field, CA, 94035, USA. [3]USDA Forest Service Western Wildland Environmental Threat Assessment Center, Corvallis, OR, 97331, USA. ✉e-mail: bridget@climateanalyticsgroup.org

Variable	Description	Units	Lower bound	Upper bound
pr	Precipitation (including both liquid and solid phases)	mm	0 mm	1000 mm
tasmax	Maximum near-surface air temperature	degrees C	−70 C	60 C
tasmin	Minimum near-surface air temperature	degrees C	−70 C	60 C
vpr	Vapour pressure	hPa	0 hPa	45 hPa

Table 1. Variables included in the NEX-DCP30-CMIP6 archive.

Model	Variant	pr	tasmax	tasmin	vpr
ACCESS-CM2	r1i1p1f1				
ACCESS-ESM1-5	r1i1p1f1				
BCC-CSM2-MR	r1i1p1f1				
CanESM5	r1i1p1f1				
CMCC-ESM2	r1i1p1f1				
CNRM-CM6-1	r1i1p1f2				
CNRM-ESM2-1	r1i1p1f2				
EC-Earth3	r1i1p1f1				
EC-Earth3-Veg-LR	r1i1p1f1				
FGOALS-g3	r3i1p1f1				
GFDL-CM4	r1i1p1f1	2-4.5, 5-8.5	2-4.5, 5-8.5	2-4.5, 5-8.5	2-4.5, 5-8.5
GFDL-ESM4	r1i1p1f1				
GISS-E2-1-G	r1i1p1f2				
HadGEM3-GC31-LL	r1i1p1f3	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5
HadGEM3-GC31-MM	r1i1p1f3	1-2.6, 5-8.5	1-2.6, 5-8.5	1-2.6, 5-8.5	1-2.6, 5-8.5
INM-CM4-8	r1i1p1f1				
INM-CM5-0	r1i1p1f1				
IPSL-CM6A-LR	r1i1p1f1				
KACE-1-0-G	r1i1p1f1				
KIOST-ESM	r1i1p1f1	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5	*1-2.6, 2-4.5, 5-8.5
MIROC6	r1i1p1f1				
MIROC-ES2L	r1i1p1f2				
MPI-ESM1-2-HR	r1i1p1f1				
MPI-ESM1-2-LR	r1i1p1f1				
MRI-ESM2-0	r1i1p1f1				
NESM3	r1i1p1f1	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5	1-2.6, 2-4.5, 5-8.5	
NorESM2-LM	r1i1p1f1				
NorESM2-MM	r1i1p1f1				
TaiESM1	r1i1p1f1				
UKESM1-0-LL	r1i1p1f2				

Table 2. CMIP6 models included in NEX-DCP30-CMIP6 archive. Key: Green = historical, and all four SSPs available; yellow = historical & some SSPs available; red = no data available. * Original GCM output for hurs SSP1.2-6 missing year 2023 and SSP2.4-5 missing year 2058, so vpr not derived for those years.

Within each GCM simulation of a scenario, only a single realisation ("variant") was downscaled. Because identical variant labels do not imply identical simulation configurations among GCMs, each was chosen to maximise the level of completeness across SSPs and variables for each GCM.

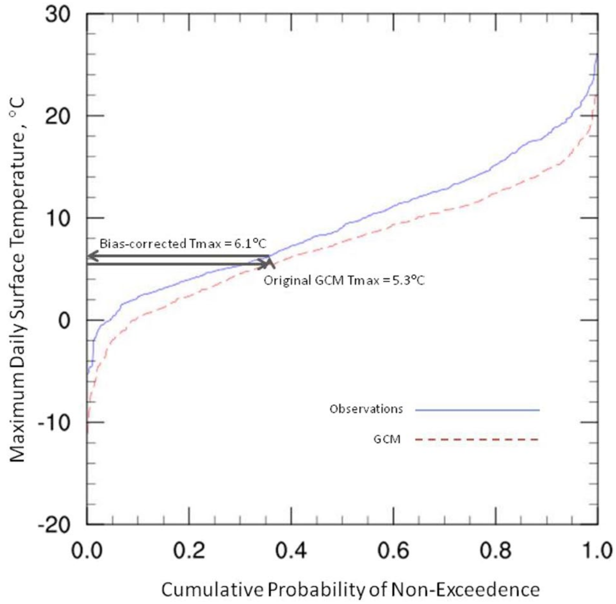

Fig. 1 Example CDF for maximum daily near-surface temperature. Arrows illustrate the quantile mapping process used for bias correction. Adapted from Thrasher *et al.*[9].

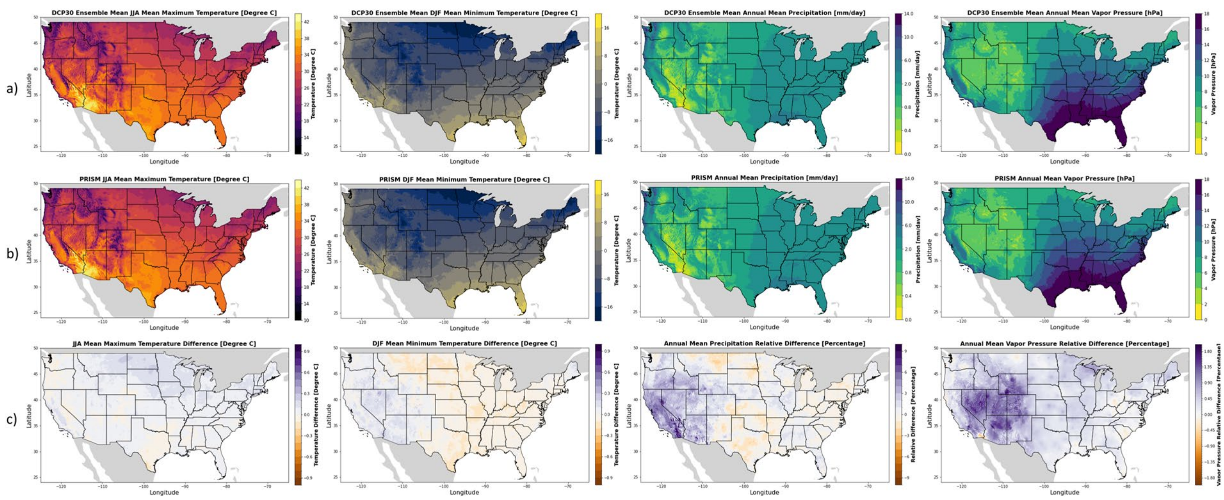

Fig. 2 Seasonal and annual ensemble means. Row (**a**) NEX-DCP30-CMIP6 1985–2014 ensemble means of summer (JJA) maximum temperature, winter (DJF) minimum temperature, annual precipitation, and annual vapour pressure. Row (**b**) PRISM 1985–2014 means of summer (JJA) temperature, winter (DJF), temperature, precipitation, and vapour pressure. Row (**c**) Differences between NEX-DCP30-CMIP6 ensemble and PRISM means.

The daily BCSD method does not add information beyond that in the original CMIP6 scenarios. It assumes that relative spatial patterns in future climate change will remain consistent with those represented in the observation-based reference dataset. In addition, the downscaled values corresponding to extreme events in the GCM output might be underestimated or overestimated in some locations due to dispersal of the area-averaged value that results from the interpolation from low to high spatial resolution (Hwang and Graham[10], Gutmann *et al.*[11]). However, the frequency of those extreme events within each individual scenario are preserved.

Temperature Pre-processing. Before applying the bias correction, monthly large-scale trends are removed from the SSP temperature variables. In this implementation, the trends were calculated for each individual month as a 9-year running average (e.g., the trend for all Januaries), preserved, and then returned to the adjusted data after the bias correction step.

Bias correction. While the original NEX-DCP30 archive used the monthly-averaged 30-arcsecond gridded fields generated by the PRISM Climate Group at Oregon State University with the Parameter-elevation

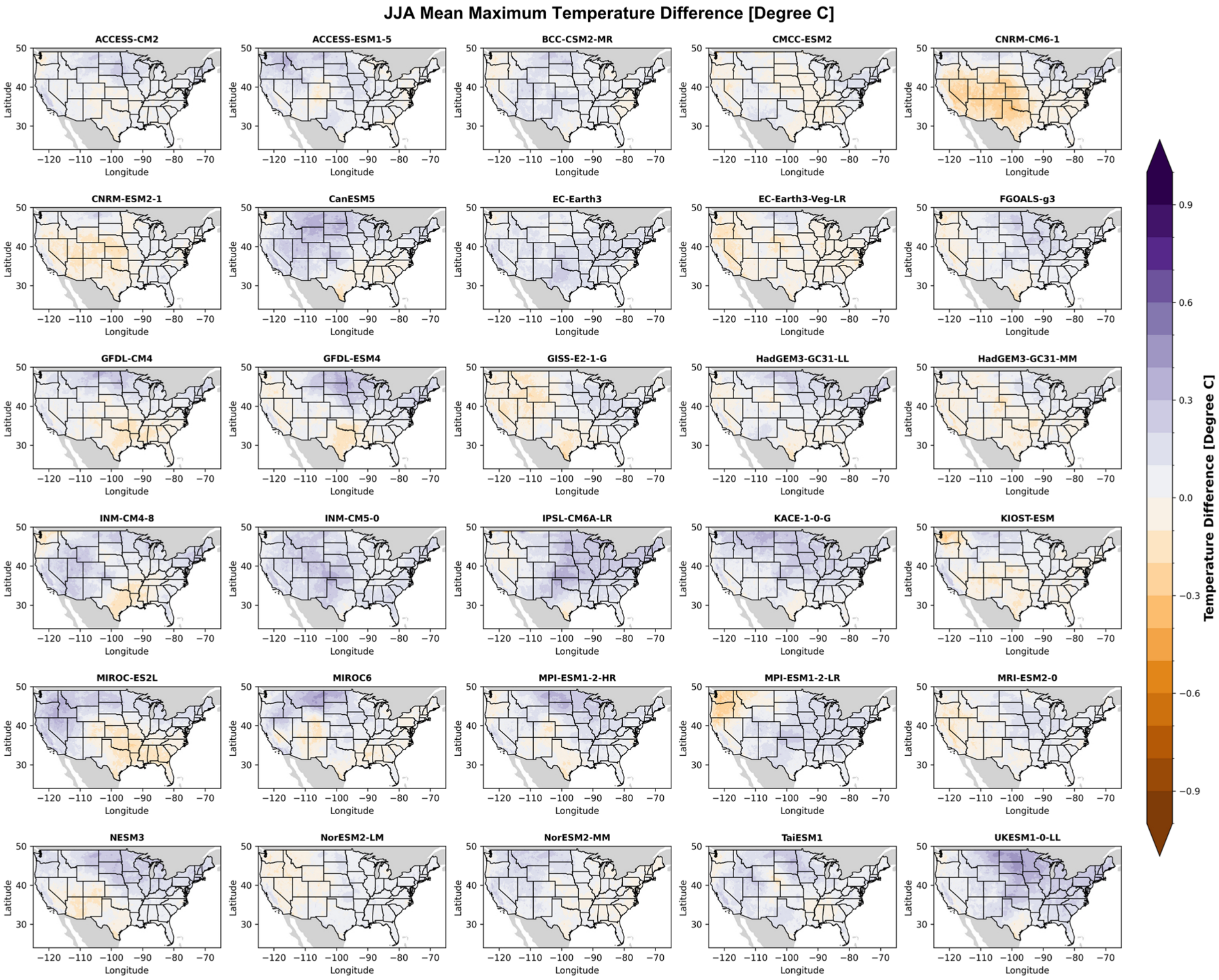

Fig. 3 Individual model mean summer (JJA) maximum temperature differences from PRISM mean.

Relationships on Independent Slopes Model[12] (PRISM), this application made use of the latest version of the daily PRISM output. These fields were first interpolated to the horizontal grid of the GCM in process using local area averaging. Then a reference period of 1981–2014 (years common to both PRISM and the CMIP6 historical experiment) served as the basis for the cumulative distribution functions (CDFs) used for bias correction, resulting in 1,054 values (+/-15-day moving window over each of thirty-four years) in each of the two CDFs, one from the PRISM output and one from the GCM historical experiment output. Each day's GCM-based CDF distribution was used to calculate the quantile associated with each original GCM value, then the corresponding value from the PRISM-based CDF distribution at that quantile was calculated, resulting in the bias-corrected version of the GCM output. Figure 1 provides a visual example of this procedure. As noted in Thrasher *et al.*[9], and implemented here, bias-corrected minimum daily temperature is best derived from the difference between the bias-corrected maximum daily temperature and the bias-corrected diurnal temperature range fields.

Spatial disaggregation. Upon completion of the bias correction, a spatial disaggregation was performed as follows. First, the daily climatology of the 30-arcsecond PRISM output was calculated using a Fast Fourier Transform. This PRISM climatology was then aggregated to the low-resolution grid of the GCM being processed and removed from the bias-corrected fields either through subtraction (temperature variables) or division (all other variables). The resulting residuals were then bilinearly interpolated to the PRISM 30-arcsecond horizontal grid. Lastly, the original high-resolution PRISM climatology was returned either through addition or multiplication. This results in the final downscaled product.

Vapour pressure. While vapour pressure (vpr) was not included in the original NEX-DCP30 archive, the variable was downscaled for CMIP5 as a supplement to that dataset[13,14], so it has also been downscaled using the method above and included in the NEX-DCP30-CMIP6 archive. Because vpr is not a variable archived in either the CMIP6 output or the PRISM output, we used other variables provided in those archives to estimate it before downscaling. The PRISM output includes daily average dewpoint temperature (tdmean), therefore we used the equation for actual vapour pressure (e) from the United States National Weather Service (NWS;

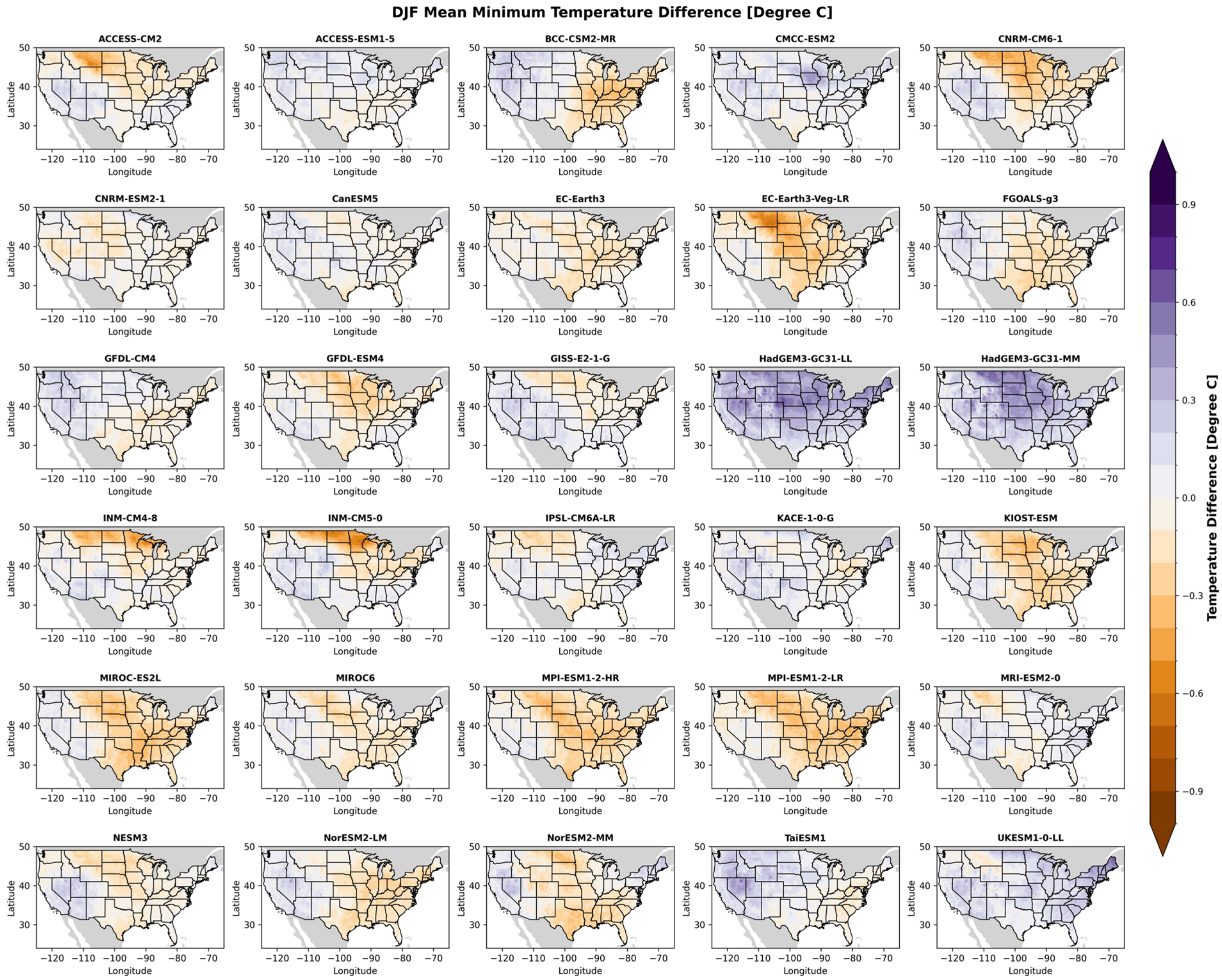

Fig. 4 Individual model mean winter (DJF) minimum temperature differences from PRISM mean.

https://www.weather.gov/media/epz/wxcalc/vaporPressure.pdf) to estimate the PRISM-based daily average vapour pressure (e_{PRISM}) as a function of tdmean as follows. This is an empirical formula known as the Magnus-Tetens approximation derived from the Clausius-Clapeyron equation relating saturation water vapour pressure with temperature:

$$e_{PRISM} = e(tdmean) \simeq 6.11 * 10^{\left(\frac{7.5*tdmean}{237.3+tdmean}\right)}$$

Meanwhile, the CMIP6 output includes daily average maximum temperature (tasmax), minimum temperature (tasmin), and relative humidity (hurs), so by combining the NWS equations for saturated vapour pressure (e_s) and relative humidity (rh), we estimated the CMIP6-based daily average vapour pressure (e_{CMIP}) in hPa as follows:

$$e_{CMIP} = e(tasmax, tasmin, hurs) \simeq \frac{e_{s(tasmax)} + e_{s(tasmin)}}{2.0} * \frac{hurs}{100.0}$$

Data Records

The NEX-DCP30-CMIP6 archive[3] is made available through a Creative Commons Zero (CC0) licence and is hosted by the NASA Advanced Supercomputing (NAS) Data Portal at https://doi.org/10.25966/427f-8370. Table 2 lists the current contents, including the variables available for each model and notes on any inventory discrepancies between models. These are the same models and variants contained in the NEX-GDDP-CMIP6 archive[4], excluding those models that did not provide tasmax and tasmin. The downscaled output is contained in yearly files in compressed netCDF-4 classic format, with file sizes ranging from ~8–11 GB for daily output, ~350–390 MB for monthly output, and ~30 MB for annual output. The current data volume of the entire archive is approximately 480 TB, broken down as follows: 459 TB of daily output, 17 TB of monthly output, and 1.5 TB of annual output.

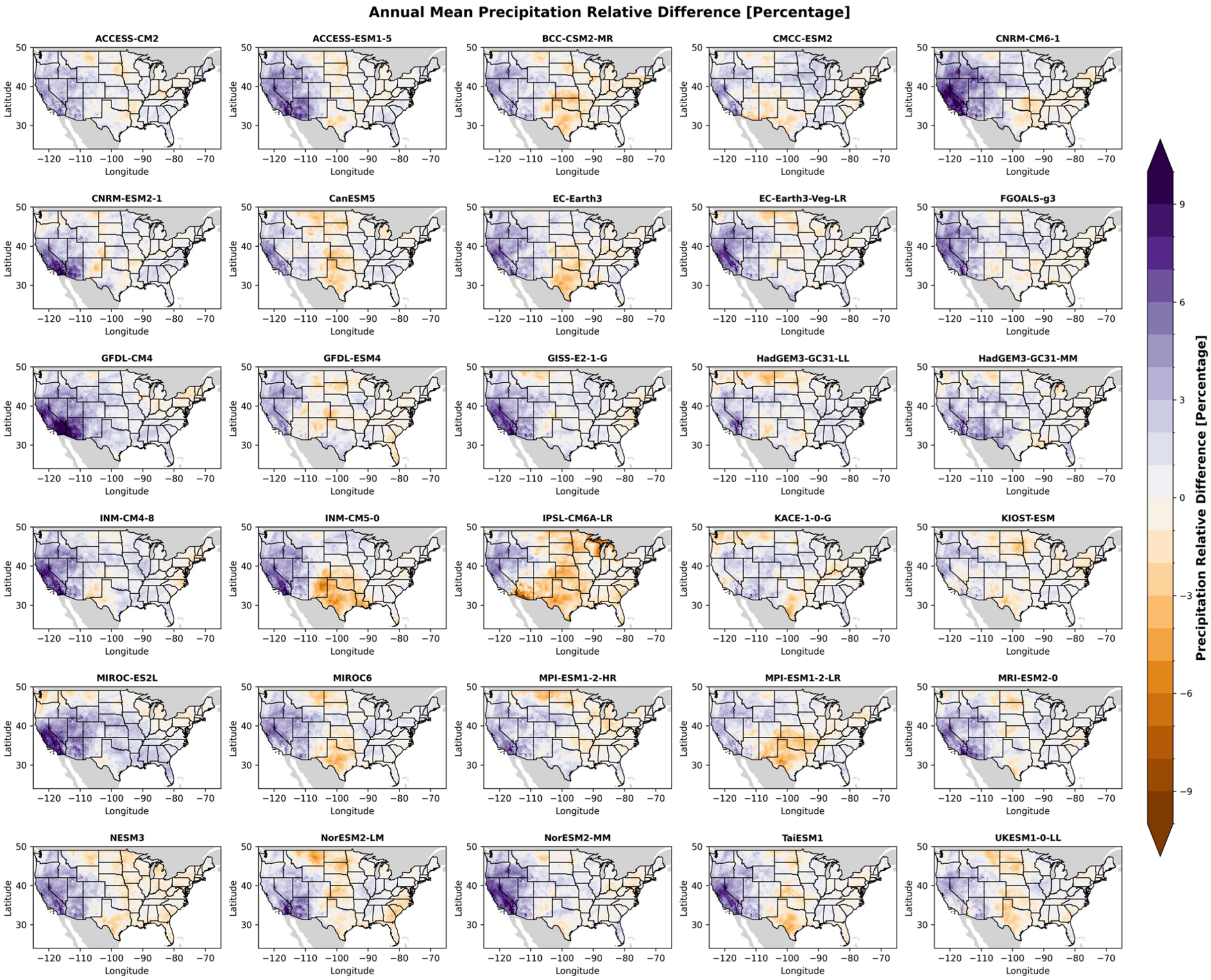

Fig. 5 Individual model mean annual precipitation differences from PRISM mean.

Technical Validation

To confirm that values fell within a realistic range, a quality control process was applied to all downscaled output. Table 1 shows the bounding values used to evaluate each variable. While values falling outside these ranges are not necessarily erroneous, they should not occur very often, and in fact, these outlying values occur in less than 0.001% of the entire archive. In addition, the geographic correlations between the NEX-DCP30-CMIP6 ensemble annual means and the PRISM annual means for all four variables were calculated. As expected, the R-squared, P-value, and correlation metrics are identical across the variables (1.00, 0.00, and 1.00, respectively), while the Root Mean Square Error (RMSE) ranges from 0.02 to 0.08 (not shown).

Further validation was performed through various comparisons of the downscaled output to its reference dataset over the period 1985–2014. Figure 2 shows (a) NEX-DCP30-CMIP6 ensemble means for summer (JJA) maximum daily temperature, winter (DJF) minimum daily temperature, annual precipitation, and annual vapour pressure, (b) the means from PRISM, and (c) the differences (arithmetic differences for tasmax and tasmin, percent differences for pr and vpr) between between the two. The mean values are geographically distributed as expected (e.g., coldest winter temperatures occur in the central north, highest vapour pressures occur along the Gulf of Mexico). Differences between DCP30 and PRISM are primarily concentrated in different regions depending on the variable. For example, the largest differences in summer precipitation occur in California, where minimal precipitation leads to large percentage differences with only trace rainfall.

Figures 3–6 show these same differences from PRISM for individual models (JJA temperature, DJF temperature, annual precipitation, and annual vapour pressure, respectively). Most models show similar patterns of distribution, though the magnitude of the differences vary across models. Figure 7 shows a range of NEX-DCP30-CMIP6 ensemble mean and PRISM percentiles for maximum temperature, minimum temperature, precipitation, and vapour pressure (25th, 50th, 75th, and 90th, respectively, except for minimum temperature rows, which show 75th, 50th, 25th, and 10th). Again, the percentile values are geographically distributed as expected. Any additional validation required for a particular application of these data is left as an exercise for the user.

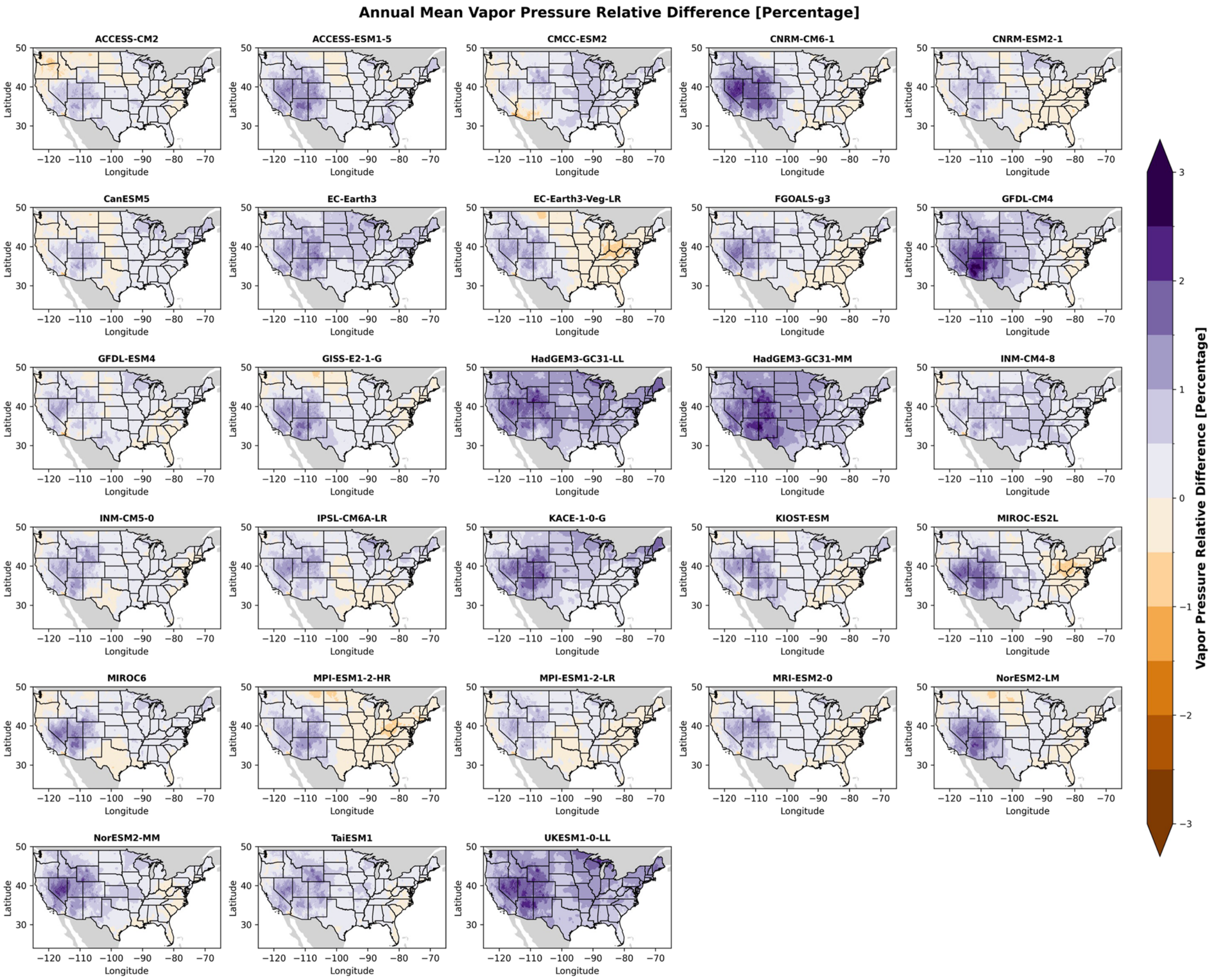

Fig. 6 Individual model mean annual vapour pressure differences from PRISM mean.

Usage Notes

NEX-DCP30-CMIP6[3] includes both SSP5-8.5 and SSP3-7.0. While some question the plausibility of SSP5-8.5[15] and may consider using SSP3-7.0 as an alternative high warming scenario, SSP 3-7.0 represents a special aerosol emissions case and should be accompanied by SSP5-8.5[16]. In addition, vapour pressure modelled by GCMs has been found to differ from recent observations at coarse resolutions, especially in arid and semi-arid regions[17]. DCP30 vapour pressure values have been bias corrected using PRISM-based vapour pressure values, but having been derived from GCM outputs, may still exhibit similar patterns. Therefore, it should be considered an estimation and as such used with discretion. NEX-DCP30-CMIP6 is intended for use in scientific research only. Its use for other purposes, such as commercial applications, engineering, or design studies is not recommended without expert engagement. NASA datasets, such as NEX-DCP30-CMIP6, may be replicated by external providers, and users are advised to confirm that they are using the most current release.

In addition to HTTP access through the Data Portal, NAS also provides open access to NEX-DCP30-CMIP6 through a Thematic Real-time Environmental Distributed Data Services (THREDDS) Data Server (TDS) at https://data.nas.nasa.gov/thredds/catalog/nex-dcp30-cmip6/cmip6Catalog.html. TDS is a web server that provides metadata and access for scientific datasets, using OPeNDAP, OGC WMS and WCS, HTTP, and other remote data access protocols. Given the several methods of data access TDS provides, files can be download directly to a local system via the HTTPServer access method, or data can be subset and read directly off the server, which may be a more tenable option for users that cannot download the full files to their local systems due to the dataset size.

Another option for data access is to download a set of files of interest from the TDS and then use a software package or application to read the netCDF files on the local system. Some tools can read subsets of the data off the server directly via the OpenDAP protocol. There are numerous high-level scripting languages, such as Python, and higher-level open source tools (e.g., Grid Analysis and Display system (GrADS), NetCDF Operators (NCO), Climate Data Operators (CDO), NetCDF Viewer (Ncview), and Panoply) readily available to meet most users'

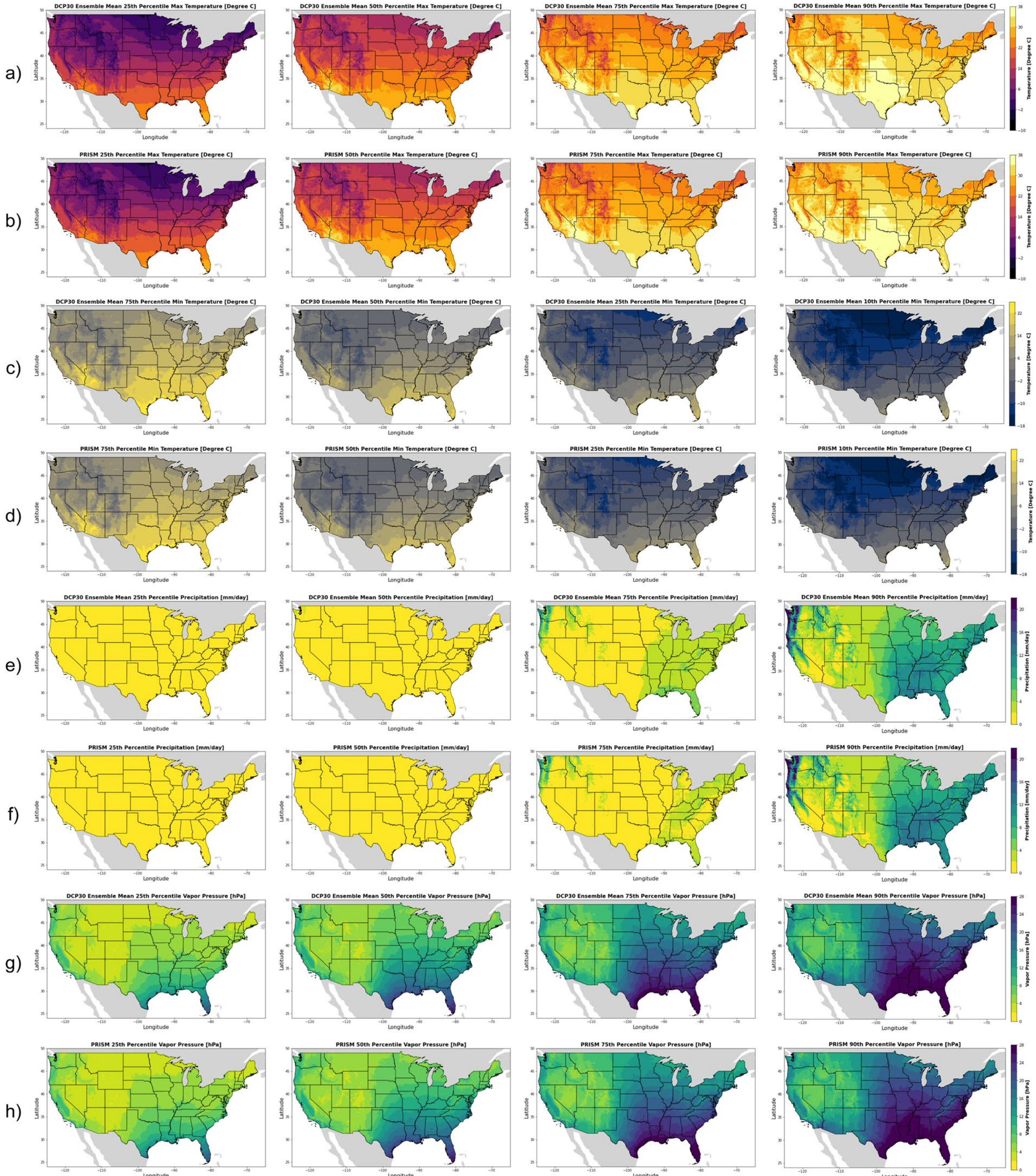

Fig. 7 Ensemble mean percentiles. Row (**a**) NEX-DCP30-CMIP6 ensemble mean percentiles of daily maximum temperature. Row (**b**) PRISM percentiles of daily maximum temperature. Row (**c**) NEX-DCP30-CMIP6 ensemble mean percentiles of daily minimum temperature. Row (**d**) PRISM percentiles of daily minimum temperature. Row (**e**) NEX-DCP30-CMIP6 ensemble mean percentiles of daily precipitation. Row (**f**) PRISM percentiles of daily precipitation. Row (**g**) NEX-DCP30-CMIP6 ensemble mean percentiles of daily vapour pressure. Row (**h**) PRISM percentiles of daily vapour pressure.

needs for data exploration and analytics. The NetCDF files are written with the CF-1.7 metadata conventions, providing an opportunity to work with the data in GIS applications, such as the open source QGiS application.

Code availability

The NCL code used to generate the downscaled products can be found at https://github.com/NEX-Climate-Downscaling/daily_BCSD.

Received: 27 August 2024; Accepted: 28 November 2024;
Published online: 18 December 2024

References

1. Thrasher, B. *et al*. Downscaled climate projections suitable for resource management. *Eos, Transactions American Geophysical Union* **94**(37), 321–323, https://doi.org/10.1002/2013EO370002 (2013).
2. *NASA Earth Exchange (NEX) Downscaled Climate Projections (NEX-DCP30)* https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-dcp30 (2013).
3. *NEX-DCP30-CMIP6* https://doi.org/10.25966/427f-8370 (2024).
4. *NEX-GDDP-CMIP6* https://doi.org/10.7917/OFSG3345 (2021).
5. Thrasher, B. *et al*. NASA Global Daily Downscaled Projections, CMIP6. *Sci Data* **9**, 262, https://doi.org/10.1038/s41597-022-01393-4 (2022).
6. Kim, J. B., Jiang, Y., Hawkins, L. R. & Still, C. J. A comparison of multiple statistically downscaled climate change datasets for the conterminous USA. *Environmental Research Communications* **4**(12), 125005, https://doi.org/10.1088/2515-7620/aca3ee (2022).
7. Wood A. W., Maurer E. P., Kumar A. & Lettenmaier D. P. Long-range experimental hydrologic forecasting for the eastern United States. *J. of Geophysical Research-Atmospheres* **107**, https://doi.org/10.1029/2001JD000659 (2002).
8. Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change* **62**, 189–216, https://doi.org/10.1023/B:CLIM.0000013685.99609.9e (2004).
9. Thrasher, B., Maurer, E. P., McKellar, C. & Duffy, P. B. Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences* **16**, 3309–3314, https://doi.org/10.5194/hess-16-3309-2012 (2012).
10. Hwang, S. & Graham, W. D. Assessment of Alternative Methods for Statistically Downscaling Daily GCM Precipitation Outputs to Simulate Regional Streamflow. *J. of the American Water Resources Assn* **50**(4), 1010–1032, https://doi.org/10.1111/jawr.12154 (2014).
11. Gutmann, E. *et al*. An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research* **50**(9), 7167–7186, https://doi.org/10.1002/2014WR015559 (2014).
12. Daly, C. *et al*. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology: a Journal of the Royal Meteorological Society* **28**(15), 2031–2064, https://doi.org/10.1002/joc.1688 (2008).
13. Drapek, R. J., Thrasher, B. L. & Kim, J. B. Dataset of monthly downscaled future vapor pressure projections for the conterminous USA for RCP 4.5 and RCP 8.5 compatible with NEX-DCP30. *Data in Brief* **48**, 109169, https://doi.org/10.1016/j.dib.2023.109169 (2023).
14. Drapek, R. J,; Kim, J. B., & Thrasher, B. L. Vapor pressure data for the conterminous United States at a 30 arcsecond resolution for 28 CMIP5 Global Climate Models under RCP 4.5 and RCP 8.5 scenarios. *Forest Service Research Data Archive*. https://doi.org/10.2737/RDS-2023-0001 (2023).
15. Ritchie, J. & Dowlatabadi, H. Why do climate change scenarios return to coal? *Energy* **140**, 1276–1291, https://doi.org/10.1016/j.energy.2017.08.083 (2017).
16. Shiogama, H. *et al*. Important distinctiveness of SSP3–7.0 for use in impact assessments. *Nature Climate Change* **13**(12), 1276–1278, https://doi.org/10.1038/s41558-023-01883-2 (2023).
17. Simpson, I. R. *et al*. Observed humidity trends in dry regions contradict climate models. *Proceedings of the National Academy of Sciences* **121**(1), e2302480120, https://doi.org/10.1073/pnas.2302480120 (2024).

Acknowledgements

This work received funding from the USDA Forest Service (contract 1240BG22P0087) and NASA (under the NASA Earth eXchange (NEX) workplans 2018–2022 and 2023–2027). NEX, a collaborative computing platform for the Earth Science Community, provides a sustained assessment capability for the U.S. National Climate Assessment (NCA).

Author contributions

All authors contributed to the development of the manuscript. B.T. performed data collection, wrangling, management, and processing, applied the downscaling algorithm, and performed the technical evaluation. S.K. performed extensive data analysis for technical validation and managed data distribution. J.K. provided funding, performed data collection, and defined the vapour pressure methodology. I.B. provided funding and contributed to the technical evaluation.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to B.T.

Reprints and permissions information is available at www.nature.com/reprints.

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