California State University, Monterey Bay

Digital Commons @ CSUMB

6-2-2022

NASA Global Daily Downscaled Projections, CMIP6

Bridget Thrasher
Climate Analytics Group

Weile Wang
NASA Ames Research Center

Andrew Michaelis
NASA Ames Research Center

Forrest Melton
California State University, Monterey Bay

Tsengdar Lee
NASA Headquarters

See next page for additional authors

Follow this and additional works at: https://digitalcommons.csumb.edu/aes_fac

Authors

Bridget Thrasher, Weile Wang, Andrew Michaelis, Forrest Melton, Tsengdar Lee, and Ramakrishna Nemani

scientific **data**

Check for updates

NASA Global Daily Downscaled Projections, CMIP6

Bridget Thrasher [1] ✉, Weile Wang[2], Andrew Michaelis[2], Forrest Melton[3], Tsengdar Lee[4] & Ramakrishna Nemani [2]

We describe the latest version of the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-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 1/4-degree horizontal resolution. Currently, eight variables from five CMIP6 experiments (historical, SSP126, SSP245, SSP370, and SSP585) are provided as procurable from thirty-five global climate models.

Background & Summary

The first version of the NASA Earth Exchange Global Daily Downscaled Projections[1] archive (NEX-GDDP) was originally released in 2015 with the intention of enhancing public understanding and assisting the scientific community in the study of climate change impacts over scales from regional to local. Since then, the projections have been used by a variety of groups across the globe, from government researchers to students, from non-governmental organizations to industry professionals. While this original archive was based on the global climate model (GCM) output from the Climate Model Intercomparison Project Phase 5 (CMIP5), results from the state-of-the-art CMIP6 experiments have recently become available. The release of these newer climate projections has resulted in a strong interest from the community in a new version of NEX-GDDP.

Here we present the latest version of the NEX-GDDP archive (NEX-GDDP-CMIP6). Not only are these updated projections based on the newer CMIP6 output, but they also include an expanded set of variables as shown in Table 1. Figures 1 and 2 show examples of the differences between corresponding model outputs from the two archives. The downscaling methodology, while essentially the same bias correction/spatial disaggregation (BCSD) method as that used for NEX-GDDP, has been slightly updated to reflect the changes in experimental design between the two incarnations of CMIP.

The NEX-GDDP-CMIP6 archive will be useful for assessing trends in projected changes in climate 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 might be available from other climate data products.

Methods

The statistical downscaling algorithm used to create the NEX-GDDP datasets is a daily variant of the monthly (BCSD) method described in Wood *et al.*[2,3]. This variant method[4] compares the original GCM output with corresponding climate observations over a common reference period and uses information derived from the comparison to adjust future climate projections so they are more consistent with the historical climate records and, presumably, more realistic for the spatial domain of interest. The algorithm also utilizes the spatial detail provided by observationally-derived datasets to interpolate the GCM outputs to higher-resolution grids.

The daily BCSD method described herein inherently assumes that the relative spatial patterns observed from the reference period will remain constant under future climate change. Other than the higher spatial resolution and bias correction, this dataset does not add information beyond what is contained in the original CMIP6 scenarios. While the frequency of periods of anomalously high and low values (i.e., extreme events) within each individual scenario are preserved, the downscaled values corresponding to extreme events might in some regions be underestimated due to diffusion over the spatial distribution resulting from the translation from low to high spatial resolution (e.g., Hwang and Graham[5] Gutmann *et al.*[6]).

[1]Climate Analytics Group, Boulder, CO, 80305, USA. [2]NASA Ames Research Center, Moffett Field, CA, 94035, USA. [3]NASA ARC-CREST, CSUMB, Seaside, CA, USA. [4]NASA Headquarters, Washington, DC, USA. ✉e-mail: bridget@climateanalyticsgroup.org

Variable	Description	Units
hurs	Near-surface relative humidity	percentage
huss	Near-surface specific humidity	kg/kg
pr	Precipitation (including both liquid and solid phases)	kg/m²/s
rlds	Surface downwelling longwave radiation	W/m²
rsds	Surface downwelling shortwave radiation	W/m²
sfcWind	Surface wind speed	m/s
tas*	Near-surface air temperature	degrees K
tasmax*	Maximum near-surface air temperature	degrees K
tasmin*	Minimum near-surface air temperature	degrees K

Table 1. Variables included in the NEX-GDDP-CMIP6 archive. *Downscaled tas (near-surface air temperature) is provided for models that did not provide tasmax and tasmin. Otherwise tas is derived from tasmax and tasmin.

Temperature Pre-processing. Because the BCSD method does not explicitly adjust the trends (the slopes, in particular) in climate variables produced by GCMs, the monthly large-scale climate trends are extracted from the GCM temperature data. This is calculated as a 9-year running average for each individual month (e.g. the trend for all Januaries taken together). These trends are preserved and added back to the adjusted data after the bias-correction step.

Bias correction. While recent studies[7,8] have shown that the choice of observational data can impact the final downscaled output, to maintain consistency between dataset versions, the latest version of the same Global Meteorological Forcing Dataset (GMFD) for Land Surface Modeling[9] that was used in creation of the original NEX-GDDP archive was used here. GMFD blends reanalysis data with observations and is available at spatial resolutions of 0.25 degrees, 0.5 degrees and 1.0 degrees, and at temporal resolutions of 3-hourly, daily, and monthly timesteps. Development of NEX-GDDP-CMIP6 utilized the 0.25-degree daily-averaged data for each variable of interest from 1960 to 2014.

These gridded observations were first interpolated to the resolution of the particular CMIP6 model being processed. A reference period of 1960–2014 was used as the basis for the cumulative distribution function (CDF), and each daily GCM value was evaluated using a +/− 15-day window over all years in the reference period to capture a full representation of candidate values for a particular variable. This resulted in 1,705 values (31 days over each of 55 years) to define each of the two CDFs for each value being corrected, one from the gridded observations and one from the GCM output from the historical experiment. The quantile associated with each original GCM value was calculated from that day's GCM-based CDF distribution, then that quantile was used to calculate the corresponding value from the observation-based CDF distribution. This final value is the bias-corrected version of the GCM output. Figure 3 provides a visual example of this procedure. The one exception is the creation of bias-corrected minimum daily temperature, which is derived from the difference between the bias-corrected maximum daily temperature and the bias-corrected diurnal temperature range fields, as in Thrasher, *et al.*[4].

Spatial disaggregation. After all GCM values for a particular variable were bias-corrected, a spatial disaggregation was performed. This step essentially merges the observed historical spatial climatology with the relative changes at each time step simulated by the GCMs to produce the final results. A smoothed daily climatology over the reference period was created from the GMFD fields using a Fast Fourier Transform retaining three harmonics. This climatology was interpolated to the native grid of the GCM being processed and factored out of the bias-corrected output either through subtraction from the temperature variables or through division from the other variables. The residual fields were then bilinearly interpolated to the original 1/4-degree horizontal grid of the GMFD. At this point, the 1/4-degree climatology was factored back in either through addition to the temperature variables or through multiplication by the other variables. This results in the final downscaled product.

Data Records

The NEX-GDDP-CMIP6 archive[10] is made available through a CC-BY-SA 4.0 Creative Commons license and can be found at the NASA Center for Climate Simulation. Table 2 lists the current contents, including the variables available for each model, and notes on any inventory discrepancies between models. A single variant from each of the 35 models listed was used for downscaling. The historical experiment was downscaled over the years 1950–2014, while the SSP experiments were downscaled over the years 2015–2100. The downscaled output, on a standard 1/4-degree horizontal grid, is contained in yearly files in netCDF-4 classic format, with file sizes ranging from ~200–290MB. The current total data volume of the entire archive is approximately 30TB.

Technical Validation

A quality control process was applied to all downscaled output to check that values fell within a realistic range. Table 3 shows the bounding values used to evaluate each variable. While values that fall outside these ranges are not necessarily erroneous, they should not occur very often. Indeed, these extraneous values occurred in less than 0.1% of values in the entire archive. Any additional validation required for a particular application of these data is left as an exercise for the end user.

Model	Variant	hurs	huss	pr	rlds	rsds	sfcWind	tas	tasmax	tasmin
ACCESS-CM2	r1i1p1f1	green	green	green	green	green	green	green	green	green
ACCESS-ESM1-5	r1i1p1f1	green	green	green	green	green	green	green	green	green
BCC-CSM2-MR	r1i1p1f1	red	green	green	green	green	green	green	green	green
CanESM5	r1i1p1f1	green	green	green	green	green	green	green	green	green
CESM2	r4i1p1f1	green	green	green	green	green	green	green	red	red
CESM2-WACCM	r3i1p1f1	yellow	yellow	yellow	yellow	yellow	yellow	yellow	red	red
CMCC-CM2-SR5	r1i1p1f1	green	green	green	green	green	green	yellow	red	red
CMCC-ESM2	r1i1p1f1	green	green	green	green	green	green	green	green	green
CNRM-CM6-1	r1i1p1f2	green	green	green	green	green	green	green	green	green
CNRM-ESM2-1	r1i1p1f2	green	green	green	green	green	green	green	green	green
EC-Earth3	r1i1p1f1	green	green	green	green	green	green	green	green	green
EC-Earth3-Veg-LR	r1i1p1f1	green	green	green	green	green	green	green	green	green
FGOALS-g3	r3i1p1f1	green	green	green	green	green	green	green	green	green
GFDL-CM4 (gr1)	r1i1p1f1	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
GFDL-CM4 (gr2)	r1i1p1f1	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
GFDL-ESM4	r1i1p1f1	green	green	green	green	green	green	green	green	green
GISS-E2-1-G	r1i1p1f2	green	green	green	green	green	green	green	green	green
HadGEM3-GC31-LL	r1i1p1f3	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
HadGEM3-GC31-MM	r1i1p1f3	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
IITM-ESM	r1i1p1f1	green	green	green	green	green	green	green	red	red
INM-CM4-8	r1i1p1f1	green	green	green	green	green	green	green	green	green
INM-CM5-0	r1i1p1f1	green	green	green	green	green	green	green	green	green
IPSL-CM6A-LR	r1i1p1f1	green	red	green	green	green	green	green	green	green
KACE-1-0-G	r1i1p1f1	green	green	green	green	green	green	green	green	green
KIOST-ESM	r1i1p1f1	yellow *	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
MIROC-ES2L	r1i1p1f2	green	green	green	green	green	green	green	green	green
MIROC6	r1i1p1f1	green	yellow	green	green	green	green	green	green	green
MPI-ESM1-2-HR	r1i1p1f1	green	green	green	green	green	green	green	green	green
MPI-ESM1-2-LR	r1i1p1f1	green	green	green	green	green	green	green	green	green
MRI-ESM2-0	r1i1p1f1	green	green	green	green	green	green	green	green	green
NESM3	r1i1p1f1	red	red	yellow	yellow	yellow	yellow	yellow	yellow	yellow
NorESM2-LM	r1i1p1f1	green	green	green	green	green	green	green	green	green
NorESM2-MM	r1i1p1f1	green	green	green	green	green	green	green	green	green
TaiESM1	r1i1p1f1	green	green	green	green	green	green	green	green	green
UKESM1-0-LL	r1i1p1f2	green	green	green	green	green	green	green	green	green

Table 2. CMIP6 models included in downscaled archive . Key: Green = all experiments available; yellow = historical & some SSP(s) available; red = no data available. *Original GCM output for hurs SSP245 missing year 2058

Usage Notes

The NASA Center for Climate Simulation (NCCS) provides open access to the data via a Thematic Real-time Environmental Distributed Data Services (THREDDS) Data Server (TDS). TDS is a web server that provides metadata and data 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 (see the NCCS TDS endpoint), which may be a more tenable option for users that cannot download the full files to their local systems due to the dataset size.
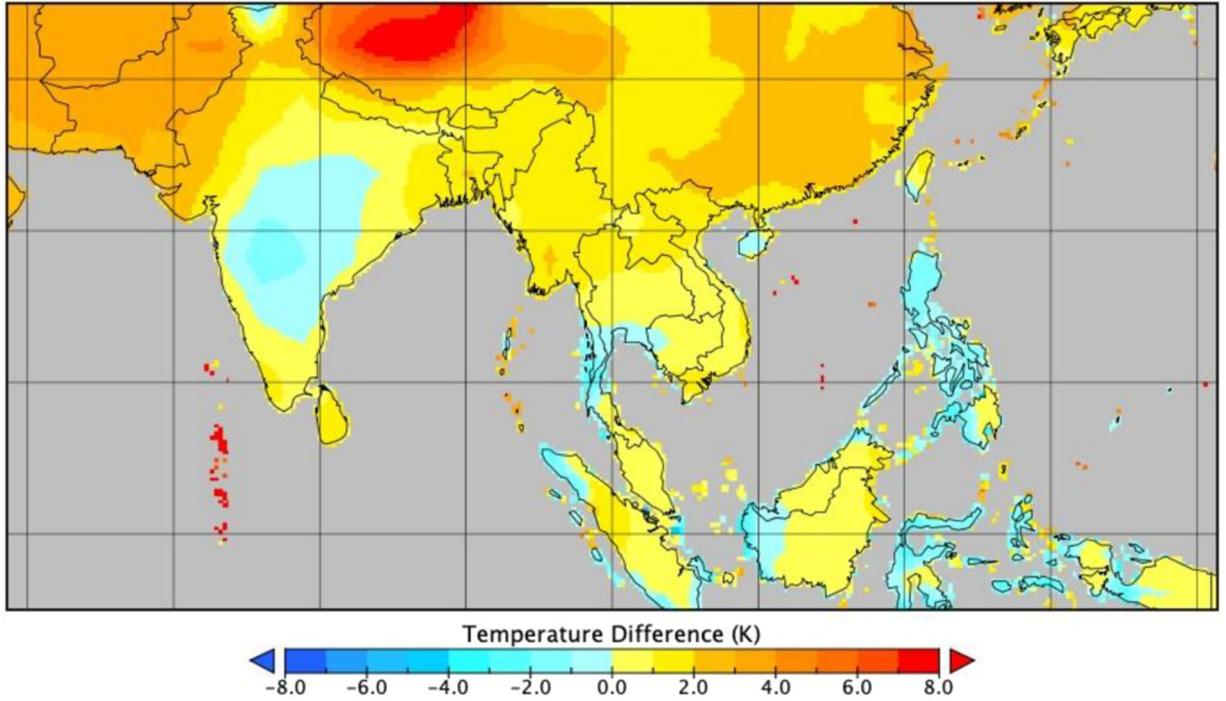
Fig. 1 End of century daily average maximum near-surface air temperature difference between the downscaled CanESM5 model output (CMIP6) and the corresponding downscaled CanESM2 model output (CMIP5).
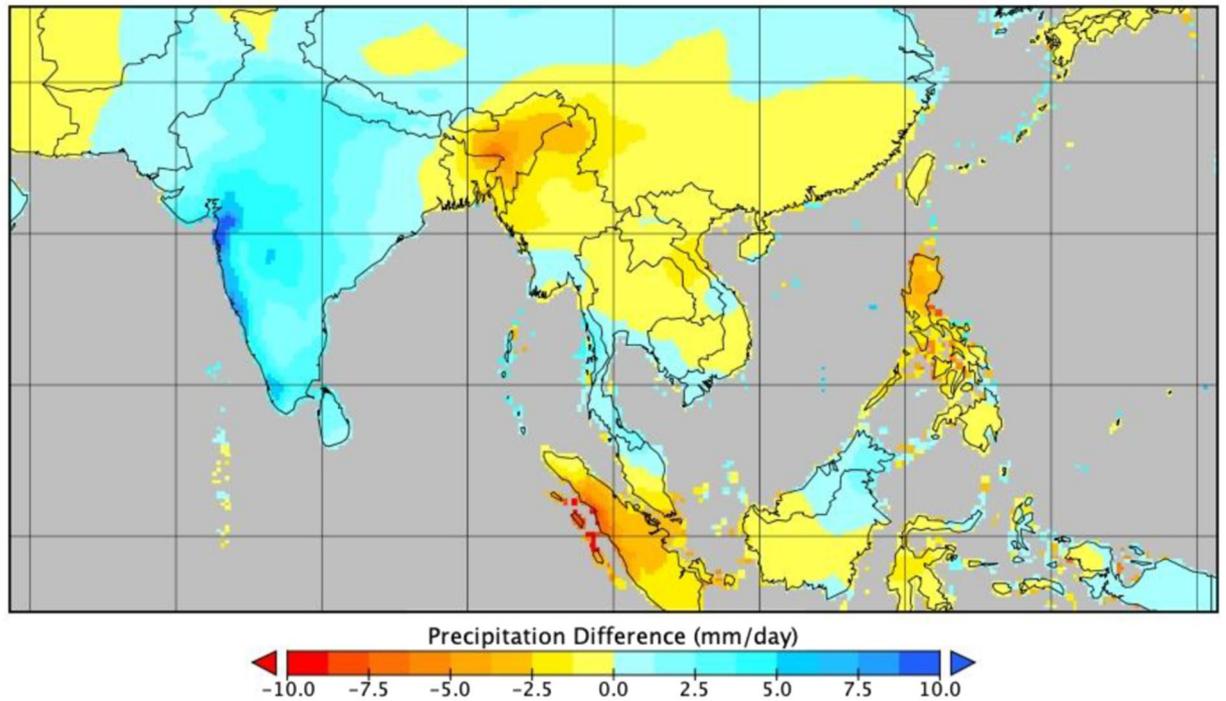

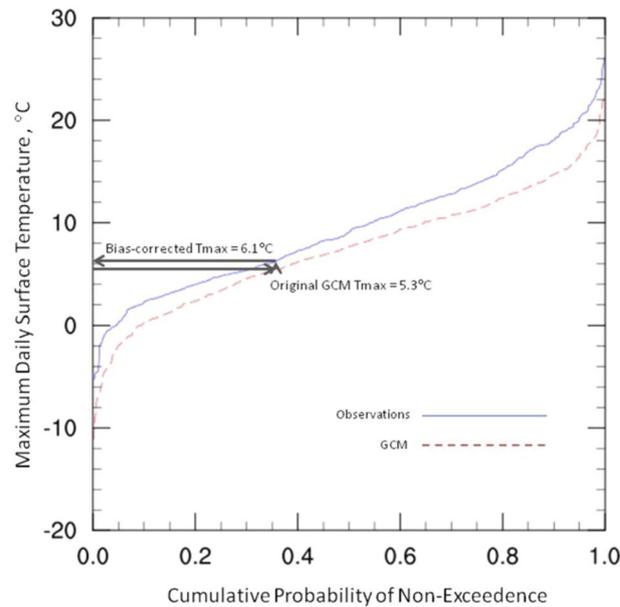**Fig. 2** Same as Fig. 1, but for precipitation.

Fig. 3 Example CDF for maximum daily near-surface temperature. Arrows illustrate the quantile mapping process used for bias correction. Adapted from Thrasher, *et al.*, 2012.

Variable	Lower bound	Upper bound
hurs	0%	102%
huss	0 kg/kg	0.04 kg/kg
pr	0 kg/m²/s	0.012 kg/m²/s
rlds	0 W/m²	500 W/m²
rsds	0 W/m²	500 W/m²
sfcWind	0 m/s	50 m/s
tas	200 K	340 K
tasmax	200 K	340 K
tasmin	200 K	340 K

Table 3. Bounds for validation of data values.

A simple option for data access is to download a set of files of interest from the TDS and then use some 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 (see the NCCS TDS endpoint). There are numerous high-level scripting languages, such as Python, and higher-level open source tools (e.g. GrADS, NCO, CDO, NCL, and Ncview) readily available to meet most user's needs for data exploration and analytics.

The packaged NetCDF files are written with the CF-1.7 metadata conventions, which may provide an opportunity to read and work with the data in GIS oriented applications, such as the open source QGiS application. Users are encouraged to explore many of the tools and applications available.

Code availability
The NCL code used to generate the downscaled products can be found at https://github.com/bthrasher/daily_BCSD.

References
1. *NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP)* https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp (2015).
2. 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** (2002).
3. 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 (2004).
4. 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 (2012).

5. 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 (2014).
6. 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.
7. Risser, M. D., Feldman, D. R., Wehner, M. F., Pierce, D. W. & Arnold, J. R. Identifying and correcting biases in localized downscaling estimates of daily precipitation return values. *Climatic Change* **169**(3), 1–20 (2021).
8. Wootten, A. M., Dixon, K. W., Adams-Smith, D. J. & McPherson, R. A. Statistically downscaled precipitation sensitivity to gridded observation data and downscaling technique. *Intl. J. of Climatology* **41**(2), 980–1001 (2021).
9. Global Meteorological Forcing Dataset for land surface modeling. *Princeton Global Forcings* http://hydrology.princeton.edu/data.pgf.php (2018).
10. *NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6)* https://doi.org/10.7917/OFSG3345 (2021).

Acknowledgements

Author contributions

B.T. performed data collection, wrangling, and processing and applied the downscaling algorithm to the CMIP6 GCM output. R.N. and T.L. conceived the idea and helped with the logistics of computing, storage and data distribution. A.M., W.W. and F.M. worked on the technical aspects of downscaling (computing, storage and distribution) and community engagement. All authors contributed to the development of the manuscript.

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.