NASA Earth Exchange
Global Daily Downscaled Projections (NEX-GDDP-CMIP6)

0. Update to Data License

As noted in the metadata of each file, the NEX-GDDP-CMIP6 archive was initially made available through a CC-BY-SA 4.0, as required by the standard data license of the original CMIP6 data. In June 2022, the CMIP6 community updated their underlying licenses (https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id_licenses.html), thereby allowing the NEX-GDDP-CMIP6 data license to also be updated. As of September 2022, all NEX-GDDP-CMIP6 outputs are made available under a blanket Creative Commons Zero license (CC0) license. All derived services and products are still required to follow the citation guidance in Section 5 below.

1. Intent of This Document and POC

This document provides a brief overview of the NASA Earth Exchange (NEX) Global Daily Downscaled Projections (GDDP) dataset and is intended for users who wish to apply the NEX-GDDP-CMIP6 dataset [https://doi.org/10.7917/OFSG3345; Thrasher et al. 2022] in studies of climate change impacts. This document summarizes essential information needed for accessing and using information contained within the NEX-GDDP-CMIP6 dataset. References and additional information are provided at the end of this document.

This NASA dataset is provided to assist the science community in conducting studies of climate change impacts at local to regional scales and to enhance public understanding of possible future climate patterns at the spatial scale of individual towns, cities, and watersheds. This dataset is intended for use in scientific research only, and use of this dataset for other purposes, such as commercial applications, and engineering or design studies is not recommended without consultation with a qualified expert. Community feedback to improve and validate the dataset for modeling usage is appreciated.

Email comments to bridget@climateanalyticsgroup.org with copy to weile.wang@nasa.gov

Dataset File Name: NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6)

Technical points of contact for this dataset:

Dr. Bridget Thrasher, bridget@climateanalyticsgroup.org

Weile Wang, weile.wang@nasa.gov

support@nccs.nasa.gov for data access

2. Data Field Descriptions

Dataset projection:	Geographic
Dataset datum:	WGS-84
Location of pixel *lat* and *lon*:	The pixel *lat* and *lon* fields in the metadata provide the location of the center of each pixel.
Coverage:	West Bounding Coordinate: 180° W East Bounding Coordinate: 180° E North Bounding Coordinate: 90° N South Bounding Coordinate: 60° S
Spatial resolution:	0.25 degrees x 0.25 degrees
Temporal resolution and extent:	Daily from 1950-01-01 to 2100-12-31 Units are in days since a reference date. The reference date varies by model and is based on the reference date used in the corresponding CMIP6 GCM experiment.

CF variable name and units:	*hurs* Near-Surface Relative Humidity percentage
	huss Near-Surface Specific Humidity dimensionless ratio (kg/kg)
	pr

	Precipitation (mean of the daily precipitation rate) kg m-2 s-1
	rlds Surface Downwelling Longwave Radiation W m-2
	rsds Surface Downwelling Shortwave Radiation W m-2
	sfcWind Daily-Mean Near-Surface Wind Speed m s-1
	tas Daily Near-Surface Air Temperature Degrees Kelvin
	tasmax Daily Maximum Near-Surface Air Temperature Degrees Kelvin
	tasmin Daily Minimum Near-Surface Air Temperature Degrees Kelvin

3. Data Origin and Methods

3.1. Introduction

The NEX-GDDP-CMIP6 dataset is comprised of global downscaled climate scenarios derived from the General Circulation Model (GCM) runs conducted under the Coupled Model Intercomparison Project Phase 6 (CMIP6) [Eyring et al. 2016] and across two of the four “Tier 1” greenhouse gas emissions scenarios known as Shared Socioeconomic Pathways (SSPs) [O’Neill et al. 2016; Meinshausen et al. 2020]. The CMIP6 GCM runs were developed in support of the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR6). This dataset includes downscaled projections from ScenarioMIP model runs [O’Neill et al. 2016; Tebaldi et al. 2021] for which daily scenarios were produced and distributed through the Earth

System Grid Federation. The purpose of this dataset is to provide a set of global, high resolution, bias-corrected climate change projections that can be used to evaluate climate change impacts on processes that are sensitive to finer-scale climate gradients and the effects of local topography on climate conditions.

The demand for downscaling of GCM outputs arises from two primary limitations inherent in current global simulation results. First, most GCMs use relatively coarse resolution grids (e.g., a few degrees or 10^2 km), which limit their ability to capture the spatial details in climate patterns that are often required or desired in regional or local analyses. Second, even the most advanced GCM may produce projections that are globally accurate but locally biased in their statistical characteristics (i.e., mean, variance, etc.) when compared with observations.

The Bias-Correction Spatial Disaggregation (BCSD) method used to generate the NEX-GDDP-CMIP6 dataset is a statistical downscaling algorithm specifically developed to address these limitations of global GCM outputs [Wood et al. 2002; Wood et al. 2004; Maurer et al. 2008; Thrasher et al. 2012]. The algorithm compares the GCM outputs with corresponding climate observations over a common period and uses information derived from the comparison to adjust future climate projections so that they are progressively 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.

With the help of the computational resources provided by NEX and the NASA Advanced Supercomputing (NAS) facility, we have applied the BCSD method to produce a global dataset of downscaled CMIP6 climate projections to facilitate the assessment of climate change impacts. The dataset compiles climate projections from thirty-five CMIP6 GCMs (Table 1) and four SSP scenarios (SSP2-4.5, SSP5-8.5, SSP1-2.6 and SSP3-7.0) for the period from 2015 to 2100, as well as the historical experiment for each model for the period 1950-2014. Each of these climate projections is downscaled to a spatial resolution of 0.25 degrees x 0.25 degrees, resulting in a data archive size of more than 18TB (1TB = 10^{12} Bytes).

This document provides a basic description of the implementation of the BCSD method as applied in the downscaling of the CMIP6 GCM data. Additional technical details for the algorithm may also be found in Wood et al. [2002, 2004], and Maurer et al. [2008]. The approach used to produce the NEX-GDDP-CMIP6 dataset was previously applied to data from the CMIP5 archive, and the approach used in production of both datasets is described in detail in Thrasher et al. [2012].

3.2 Methods

3.2.1 Datasets

Climate Model Data: We compiled output from thirty-five CMIP6 GCM models (Table 1), including the historical experiment and four SSP scenarios (SSP2-4.5, SSP5-8.5, SSP1-2.6 and SSP3-7.0). Each of the climate projections includes daily average variables for the periods from 1950 through 2014 ("retrospective simulation") and from 2015 to 2100 ("prospective simulation"). During the downscaling process, the retrospective simulations serve as the training data and are compared against the observational climate records (see below). The relationships derived from the comparison are then applied to downscale the prospective climate projections. Because all climate projections are downscaled through the same procedures, for simplicity we refer to them as "GCM data" without differentiating any individual models.

Observational Climate Data: We use a climate dataset from the Global Meteorological Forcing Dataset (GMFD) for Land Surface Modeling, available from the Terrestrial Hydrology Research Group at Princeton University [Sheffield et al. 2006]. This dataset blends reanalysis data with observations and is currently available at spatial resolutions of 0.25 degrees, 0.5 degrees and 1.0 degree, and temporal resolutions of 3-hourly, daily, and monthly timesteps. For development of the NEX-GDDP-CMIP6 dataset, we used the 0.25-degree historical daily-averaged data for maximum temperature, minimum temperature, precipitation, near-surface humidity, downwell shortwave/longwave radiation, and near-surface wind speed from 1960 to 2014.

3.2.2 Data Pre-processing

Because the BCSD method does not explicitly adjust the trends (the slopes, in particular) in climate variables produced by GCMs, we extract the monthly large-scale climate trends 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.

3.2.3 Bias Correction (BC)

The Bias-Correction step "corrects" the bias of the GCM data through comparisons performed against the GMFD historical data. For each climate variable in a given day, the algorithm generates the cumulative distribution function (CDF) for the GMFD data and for the retrospective GCM simulations, respectively, by pooling and sorting the corresponding source values (day of year +/- 15 days) over the period from 1960 through 2014. It then compares the two CDFs at various probability thresholds to establish a quantile map between the GCM data and the historical climate data. Based on this map, GCM values in any CDF quantile (e.g., p=90%) can be translated to corresponding GMFD values in the same CDF quantile. Assuming that the CDF of the GCM simulations is stable across the retrospective and the prospective periods, to "correct" the projected future climate variations the algorithm simply looks up the probability quantile associated with the predicted climate values from the estimated GCM CDF, identifies the corresponding observed climate values at the same probability quantile in the GMFD CDF, and then accepts the latter as the adjusted climate predictions. The climate projections adjusted in this way have the same CDF as the GMFD data; therefore, the possible biases in the statistical structure (the variance, in particular) of the original GCM outputs are removed by this procedure. At the end of the Bias-Correction step, the previously extracted temperature climate trends are added back to the adjusted GCM climate fields

3.2.4 Spatial Disaggregation (SD)

The Spatial-Disaggregation step spatially interpolates the Adjusted GCM data to the finer resolution grid of the 0.25-degree GMFD data. Rather than simple linear spatial interpolation, multiple steps are adopted in the SD algorithm to preserve spatial details of the observational data. First, a smoothed daily climatology is generated from the gridded observations over the reference period 1960-2014 using a Fast Fourier Transform retaining three harmonics at both the native and GCM resolutions. Second, for each time step, the algorithm compares the bias-corrected GCM variables with the corresponding GMFD climatology to calculate "scaling factors". In particular, the scaling factors are calculated as the differences between the bias-corrected GCM and the GMFD data for temperature, and as the quotients (between the two datasets) for the remaining variables to avoid negative values for the latter. Third, the coarse-resolution scaling factors are bilinearly interpolated to the fine-resolution GMFD grid. Finally, the scaling factors are applied, by addition or "shifting" for temperatures and by multiplication for the remaining variables, on the fine-resolution GMFD climatologies to obtain the desired downscaled climate fields. As such, the algorithm essentially merges the observed historical spatial climatology with the relative changes at each time step simulated by the GCMs to produce the final results.

4. Considerations and Recommended Use

4.1 Recommended Use

This dataset has been generated and is being distributed to assist the science community in conducting studies of climate change impacts at local to regional scales, and to enhance public understanding of possible future climate patterns and climate impacts at the scale of individual cities, communities, and watersheds. This dataset is intended for use in scientific research only, and use of this dataset for other purposes, such as commercial applications, and engineering or design studies is not recommended without consultation with a qualified expert.

4.2 Assumptions and Limitations

The BCSD approach used in generating this downscaled dataset inherently assumes that the relative spatial patterns observed from 1960 through 2014 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, and preserves the frequency of periods of anomalously high and low values (i.e., extreme events) within each individual scenario.

4.3 Trend Adjustment to Individual Models

As described in Section 2.1, the BCSD algorithm does not adjust the *slope* of the trends in the GCM projections. In the case of temperature, for instance, if the GCM predicts a mean temperature increase of 2°C between 2015 and 2100, the same temperature change (i.e., a trend of 2°C over 86 years) will be observed in the downscaled temperature field. However, the BCSD algorithm does adjust the *offset* of the climate trends by shifting the retrospectively simulated climate variables (1960 through 2014) to match the GMFD data. In the previous example, if the simulated mean temperature from the GCM over the period 1996-2005 is 14°C, while the observed mean temperature is 15°C, the BCSD algorithm will correct the “bias” by shifting the GCM retrospective and prospective projections upward by 1°C. The adjusted mean temperature projected for the end of the 21st century will then be raised from 16°C to 17°C, though its relative change over the period 2015-2100 is preserved as 2°C. Though such adjustments of future climate projections are qualitatively justifiable, quantitatively the linear shifting itself may not be realistic because the climate system is nonlinear in nature. Users of this dataset should be aware of this limitation of the downscaled data, particularly when using downscaled scenarios from individual GCMs.

4.4 Known Issues

The model TaiESM1 exhibits a significant deviation in temperatures from all other models for all SSPs. Users may want to take note that the temperature outputs from TaiESM1 may be considered an outlier.

Due to the lack of validation of the GMFD over oceans, GDDP values over smaller island areas might not be realistic. If this affects your area of interest, please contact us for alternative output.

5. Credits and Acknowledgements

Please use the references below as the primary citations for the dataset described herein:

Thrasher, B., Wang, W., Michaelis, A. et al. NASA Global Daily Downscaled Projections, CMIP6. Sci Data 9, 262 (2022). https://doi.org/10.1038/s41597-022-01393-4

Thrasher, B., Wang, W., Michaelis, A. Nemani, R. (2021). NEX-GDDP-CMIP6. NASA Center for Climate Simulation. https://doi.org/10.7917/OFSG3345

Please use the reference below as the primary citation for the methods used to produce this dataset:

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B., 2012: Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, 16(9), 3309-3314, doi:10.5194/hess-16-3309-2012.

Please add the following acknowledgement to any publications that result from use of this dataset:

Climate scenarios used were from the NEX-GDDP-CMIP6 dataset, prepared by the Climate Analytics Group and NASA Ames Research Center using the NASA Earth Exchange and distributed by the NASA Center for Climate Simulation (NCCS).

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

6. Terms of Use

All CMIP6 GCM model inputs, as noted in Table 1, and any derivatives work, such as this dataset, are governed by the original Terms of use (https://pcmdi.llnl.gov/CMIP6/TermsOfUse/TermsOfUse6-1.html) and may have some restrictions on usage. See the individual netcdf files for the data licensing, the global attribute “cmip6_license” notes the specific license the data may fall under.

7. References

Daly, C., R.P. Neilson, and D.L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**, 140-158.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E., 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, **9**, 1937–1958.

Maurer, E. P. and Hidalgo, H. G., 2008: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods, *Hydrology and Earth System Sciences*, **12**, 551-563.

Meinshausen, M. S.J. Smith, K. Calvin, J.S. Daniel, M.L.T. Kainuma, and et al., 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213-241.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J., 2020: The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, *Geosci. Model Dev.*, **13**, 3571–3605.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M., 2016: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geosci. Model Dev.*, **9**, 3461–3482.

Sheffield, J., G. Goteti, and E. F. Wood, 2006: Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, *J. Climate*, **19** (13), 3088-3111.

Shepard, D.S., 1984: Computer mapping: The SYMAP interpolation algorithm, in Spatial Statistics and Models, edited by G.L. Gaile and C.J. Willmott. D. Reidel Publishing Co., Norwell, Dordrecht, Holland, pp. 133-145.

Taylor, Karl E., Ronald J. Stouffer, Gerald A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498.

Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., Knutti, R., Lowe, J., O'Neill, B., Sanderson, B., van Vuuren, D., Riahi, K., Meinshausen, M., Nicholls, Z., Tokarska, K. B., Hurtt, G., Kriegler, E., Lamarque, J.-F., Meehl, G., Moss, R., Bauer, S. E., Boucher, O., Brovkin, V., Byun, Y.-H., Dix, M., Gualdi, S., Guo, H., John, J. G., Kharin, S., Kim, Y., Koshiro, T., Ma, L., Olivié, D., Panickal, S., Qiao, F., Rong, X., Rosenbloom, N., Schupfner, M., Séférian, R., Sellar, A., Semmler, T., Shi, X., Song, Z., Steger, C., Stouffer, R., Swart, N., Tachiiri, K., Tang, Q., Tatebe, H., Voldoire, A., Volodin, E., Wyser, K., Xin, X., Yang, S., Yu, Y., and Ziehn, T., 2021: Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6, *Earth Syst. Dynam.*, **12**, 253–293.

Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., & Nemani, R., 2022: NASA Global Daily Downscaled Projections, CMIP6. *Nature Scientific Data*, **9** (254), https://doi.org/10.1038/s41597-022-01393-4.

Thrasher, B., Xiong, J., Wang, W., Melton, F., Michaelis, A., & Nemani, R., 2013: Downscaled climate projections suitable for resource management. *Eos, Transactions American Geophysical Union*, **94** (37), 321-323, https://doi.org/10.1002/2013EO370002.

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B., 2012: Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, **16** (9), 3309-3314, https://doi.org/10.5194/hess-16-3309-2012.

Wood, A.W., E.P. Maurer, A. Kumar, and D.P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophysical Research-Atmospheres*, **107**, 4429, https://doi.org/10.1029/2001JD000659.

Wood, A.W., L.R. Leung, V. Sridhar, and D.P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **15**, 189-216.

7. Dataset and Document Revision History

Rev 1 – 24 September 2021 – Document created. This is a new document/dataset.

Rev 2 – 18 March 2022 – Updates to reflect addition of SSP126 and SSP370.

Rev 3 – 04 August 2022 – Update to data license, updates to Table 1 legend and reference list.

Table 1. CMIP6 models included in downscaled archive

Model	**Variant**	**hurs**	**huss**	**pr**	**rlds**	**rsds**	**sfcWind**	**tas**	**tasmax**	**tasmin**
ACCESS-CM2	r1i1p1f1									
ACCESS-ESM1-5	r1i1p1f1									
BCC-CSM2-MR	r1i1p1f1									
CanESM5	r1i1p1f1									
CESM2	r4i1p1f1									
CESM2-WACCM	r3i1p1f1									
CMCC-CM2-SR5	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 (gr1)	r1i1p1f1									
GFDL-CM4 (gr2)	r1i1p1f1									
GFDL-ESM4	r1i1p1f1									
GISS-E2-1-G	r1i1p1f2									
HadGEM3-GC31-LL	r1i1p1f3									
HadGEM3-GC31-MM	r1i1p1f3									
IITM-ESM**	r1i1p1f1									
INM-CM4-8	r1i1p1f1									
INM-CM5-0	r1i1p1f1									
IPSL-CM6A-LR	r1i1p1f1									
KACE-1-0-G	r1i1p1f1									
KIOST-ESM	r1i1p1f1	***								
MIROC-ES2L	r1i1p1f2									
MIROC6	r1i1p1f1									
MPI-ESM1-2-HR	r1i1p1f1									
MPI-ESM1-2-LR	r1i1p1f1									

MRI-ESM2-0	r1i1p1f1									
NESM3	r1i1p1f1									
NorESM2-LM	r1i1p1f1									
NorESM2-MM	r1i1p1f1									
TaiESM1	r1i1p1f1									
UKESM1-0-LL	r1i1p1f2									

Key: Green = historical & all SSPs available; yellow = historical & some SSPs available; red = no data available

* Original GCM tasmax & tasmin output retracted by CMCC (https://errata.es-doc.org/static/view.html?uid=33496f30-9e86-c0ff-874c-61f78df0509a)

** Original GCM output missing year 2100 & SSP370 missing year 2099 for all variables

*** Original GCM output for hurs SSP245 missing year 2058, hurs SSP126 missing year 2023

Note, all CMIP6 source_id and institutional_id's as listed in the CMIP6 controlled vocabulary are noted in each individual netcdf files as a global attribute cmip6_source_id and cmip6_institution_id. The controlled vocabulary is defined here:
https://github.com/WCRP-CMIP/CMIP6_CVs/blob/master/CMIP6_source_id.json
and here:
https://github.com/WCRP-CMIP/CMIP6_CVs/blob/master/CMIP6_institution_id.json.

Appendix I – Working with the NetCDF files

The data provided is in NetCDF (https://www.unidata.ucar.edu/software/netcdf/) file format. All files are written with the CF-1.7 metadata conventions (https://cfconventions.org/), as noted by the "Conventions" global attribute within the files.

There are many software options available for reading and writing NetCDF files. The most basic option is to use the ncdump command line utility, which the standard netCDF-c software package (https://github.com/Unidata/netcdf-c) provides. Note, many Linux distributions provide pre-built packages that can be installed, e.g. Ubuntu, there are pre-built Windows and MacOS packages also available.

An example of viewing a file's structure, or header, in a terminal one can execute the following command line:

```
$ ncdump -h hurs_day_ACCESS-CM2_historical_r1i1p1f1_gn_1971.nc
netcdf hurs_day_ACCESS-CM2_historical_r1i1p1f1_gn_1971 {
dimensions:
        time = UNLIMITED ; // (365 currently)
        lat = 600 ;
        lon = 1440 ;
variables:
        double time(time) ;
                time:axis = "T" ;
.
.
.
}
```

Python (https://www.python.org/) is a popular scripting language that can be used to view and analyze the data. You will need to install the Python package netCDF4 (https://unidata.github.io/netcdf4-python/) via pip, conda or build the package from source. With the netCDF4 package installed on your system, in the Python interpreter:

```
$ python
Python 3.7.3 (default, Mar 27 2019, 22:11:17)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import netCDF4
```

```
>>> data = netCDF4.Dataset('tasmax_day_ACCESS-CM2_historical_r1i1p1f1_gn_2014.nc')
>>> # print out the file structure
>>> print(data)
.
.
.
>>> # print out the variables
>>> print(data.variables)
>>> # get a variable
>>> var_tasmax = data.variables['tasmax']
>>> # print the variable structure and attributes
>>> print(var_tasmax)
.
.
.
>>> # print the variable units and fill value
>>> print(var_tasmax.units)
K
>>> print(var_tasmax._FillValue)
1e+20
>>> # actually load the data into an numpy array
>>> dat = var_tasmax[:]
dat = var_tasmax[:]
>>> # compute the max and min
>>> dat.max()
329.3559
>>> dat.min()
211.7591
>>> # subset or slice the data
>>> sub_dat = dat[10:50,20:30]
>>> # close for clean up.
>>> data.close()
```

Appendix II – Bulk downloading NetCDF files

Users often wish to download a large number of files from the data service and then do a checksum of the downloads, therefore we have provided an "index file" containing URLs and file checksums for convenience. From a terminal, one can execute the following command line to download the index file:

```
$ curl -O \
https://ds.nccs.nasa.gov/thredds2/fileServer/listing/gddp-cmip6-thredds-fileserver.csv
```

The example python script stub below illustrates the use of the index file to download the netcdf files directly from the server and perform checksums:

```python
— example.py —-
# This is a simple script stub that performs downloads from the
# NCCS thredds service at https://ds.nccs.nasa.gov/
# The script assumes the curl program is available.
import csv
import os
import hashlib
import logging
from subprocess import run as srun
from urllib.parse import urlparse

logging.basicConfig(level=logging.INFO)

def download(uri, ofile, md5):
    srun(['curl', '-s', '-o', ofile, uri], capture_output=True, check=True)
    md5dld = str(hashlib.md5(open(ofile, 'rb').read()).hexdigest())
    if md5 != md5dld:
        logging.warning("%s != %s", md5, md5dld)
    logging.info("uri %s (%s == %s) : %s", uri, md5, md5dld, ofile)

with open('gddp-cmip6-thredds-fileserver.csv') as index:
    fobjects = csv.reader(index)
    next(fobjects)
    for objs in fobjects:
        md5, uri = [o.strip() for o in objs]
        prsout = urlparse(uri)
        ofile = os.path.split(prsout.path)[1]
        download(uri, ofile, md5)
```

Appendix III – Amazon Web Services Data Access

As of September 2022, a cache copy of the NEX-GDDP-CMIP6 data is also available on Amazon Web Services (AWS) Simple Storage Service, S3. The S3 bucket name is s3://nex-gddp-cmip6/. See the AWS open data registry, https://registry.opendata.aws/ (2022), for more information.