

# Spatial and Temporal Subsets of Daymet Data Using the THREDDS NetCDF Subset Service (NCSS) for Grids

This document describes how to extract and download a spatial and/or temporal subset of individual variables of Daymet daily gridded data within a bounding box of longitude and latitude. This same methodology can be applied to the Daymet climatological summary files also available via THREDDS.

## Introduction to THREDDS and NetCDF Subset Service

Access to the Daymet collection of CF compliant netCDF files is available through a Thematic Real-time Environmental Distributed Data Services (THREDDS) Data server. A THREDDS data server allows users to find, access, and download data from a simple, hierarchical catalog within a Web browser or compatible client software. THREDDS instances are available that point to the Daymet mosaic data set, Daymet 2-degree x 2-degree tile data, or pre-derived annual and monthly climatologies. In the mosaic daily data set, each Daymet variable is a continuous grid over the entire North American study region, for each year of data. Due to file sizes, it may be cumbersome for a user to download the entire North American extent of the data. Subsetting the data through THREDDS is one option available for users who don't need access to entire data set.

Through the THREDDS data server, there is an integrated NetCDF Subset Service (NCSS) using a REST API that allows subsetting of netCDF datasets in coordinate space. From this, gridded data subsets are returned in CF-compliant netCDF-3 or netCDF-4 formats. The THREDDS data server provides an interactive GUI web form which allows users to perform NCSS-based spatio-temporal subsetting and download of one Daymet variable\_year per submission. These interactive requests can be automated through programmatic, machine-to-machine requests which involve the construction and submission of an HTTP request to the THREDDS data server through an extended URL with defined parameters. Each of these methods is described below.

## NetCDF Subset Service Access GUI

Link to the ORNL DAAC [THREDDS Data Server for Daymet data](#) or through the THREDDS access points available on the [Daymet](#) Project web site.

Several file sets for Daymet data are available. Daily mosaic files for North America are found [here](#). Browse through the catalog of data to a year and then variable selection. Once there, a number of options for data Access will be presented including NetcdfSubset, or NCSS.

### Access:

1. **OPENDAP:** [/thredds/dodsC/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
2. **HTTP Server:** [/thredds/fileServer/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
3. **WCS:** [/thredds/wcs/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
4. **WMS:** [/thredds/wms/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
5. **NCML:** [/thredds/ncml/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
6. **NetcdfSubset:** [/thredds/ncss/grid/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
7. **ISO:** [/thredds/iso/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)
8. **UDDC:** [/thredds/uddc/ornl/daac/1328/1988/daymet\\_v3\\_tmin\\_1988\\_na.nc4](#)

Clicking the NetcdfSubset link leads to the GUI web form for the NetCDF Subset Service.

ORNL DAAC THREDDS NetCDF Subset

NCSS for Grids ( Grid as Point Dataset )

Dataset: /thredds/ncss/ornl/daac/1328/1988/daymet\_v3\_tmin\_1988\_na.nc4 ( Dataset Description )

Base Time: 1988-01-01T12:00:00Z

Select Variable(s):

- lat = latitude coordinate
- lon = longitude coordinate

Variables with Time coordinate time

- tmin = daily minimum temperature

Choose Spatial Subset:

Lat/lon subset  
Coordinate subset  
Bounding box, in decimal degrees (initial extents are approximate):

north: 36.61  
west: -85.37 east: -81.29  
south: 33.57

Disable horizontal subsetting  
[reset to full extension](#)

Choose Time Subset:

Time range Single time  
Start: 1988-01-01T12:00:00Z  
End: 1988-12-30T12:00:00Z  
Stride: 1  
[reset to full extension](#)

Add 2D Lat/Lon to file (if needed for CF compliance)  
 Add Lat/Lon variables

Choose Output Format:  
Format: netcdf4

NCSS Request URL:

```
http://thredds.daac.ornl.gov/thredds/ncss/ornl/daac/1328/1988/daymet_v3_tmin_1988_na.nc4?var=lat&var=lon&var=tmin&north=36.61&west=-85.37&east=-81.29&south=33.57&disableProjSubset=on&horizStride=1&time_start=1988-01-01T12:00:00Z&time_end=1988-12-30T12:00:00Z&timeStride=1&accept=netcdf4
```

Submit Reset

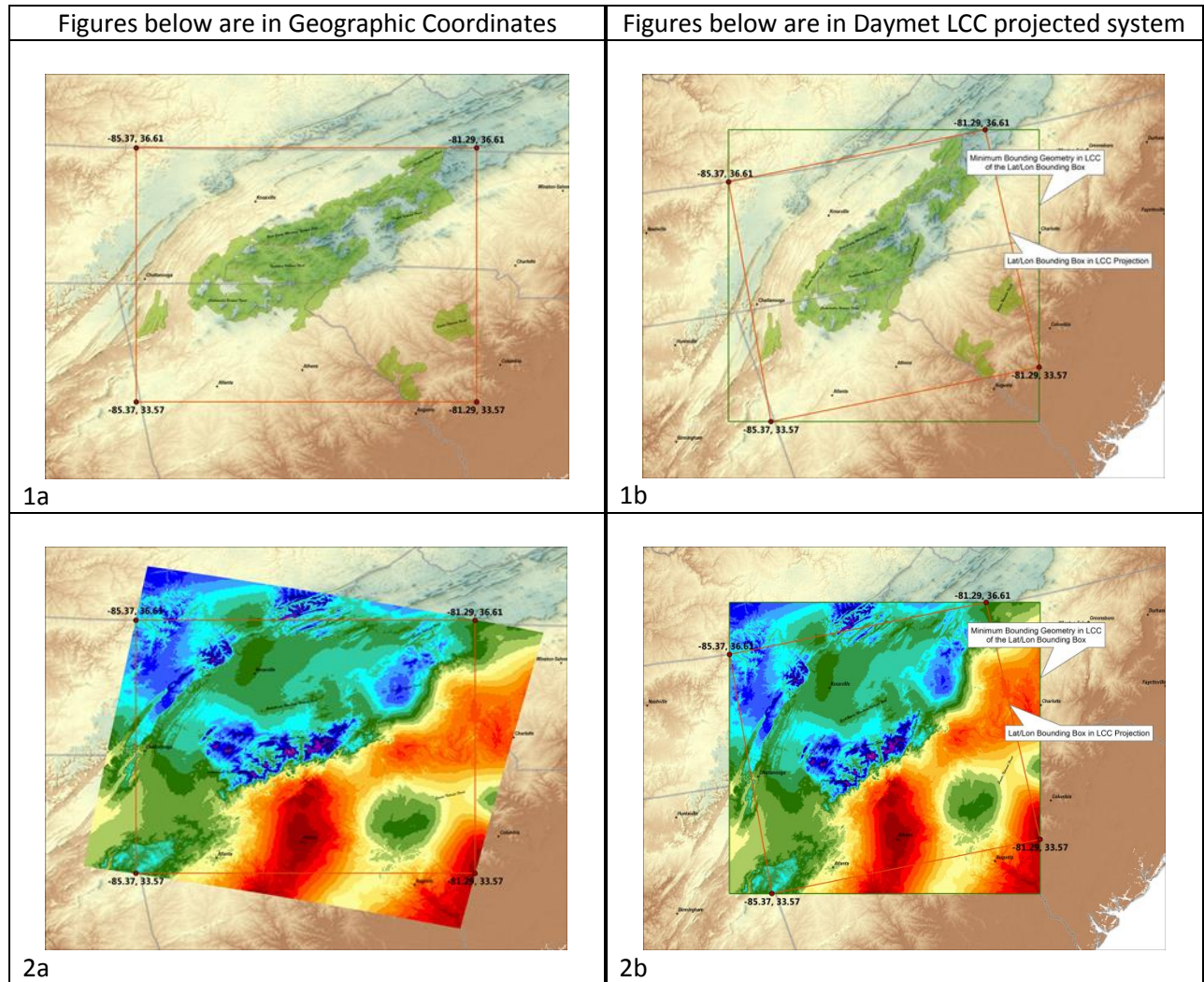
On the GUI form:

- Under Select Variable(s): Check lat, lon and Daymet data variable (eg. tmin).  
***It's very important to include the lat, lon variables from the Daymet file.***
- Enter the Coordinate and Time subset.
- ***DO NOT select*** the “Add 2D Lat/Lon to file” option. All Daymet netCDF data files were created with CF-compatibility and Lat/Lon variables are already contained in the data files. Selecting the “Add 2D Lat/Lon to file” option instead of including the existing “lat” and “lon” variables in the subset will cause “lat” and “lon” variables to be recreated on-the-fly, may be incorrect due to transformation errors, and significantly slow down the subset process.
- Choose the Output Format. netcdf-4 is recommended (see the accepted data formats section below)
- Click Submit

A popup response will appear in your window asking you to save or open the file.

**A note about the lat/lon subset and Daymet projection system**

The Daymet data set is defined in a Lambert Conformal Conic (LCC) Projection system. It's therefore necessary for the NCSS subset tool to find a minimum bounding area that encompasses the input geographic lat/lon subset Bounding Box within the corresponding LCC projected system. This minimum bounding area is used to subset the Daymet data. The resulting output file will be square in the LCC projection, with the corners estimated by the minimum bounding area of the input lat/lon coordinates. An example is provided below. A bounding box that includes the Smoky Mountain National Park and surrounding National Forests is input in latitude/longitude. In a geographic coordinate system (fig 1a) the bounding box is square. In LCC projected coordinates (fig 1b) the geographic bounding box is shown in the LCC projection. Also shown in fig 1b is the minimum bounding area that includes the lat/lon subset coordinates (green). One time slice of the resulting subset Daymet minimum temperature netCDF data file is displayed in a geographic coordinate system (fig 2a) and in the Daymet LCC projection (fig 2b).



## Evaluating the URL Request

The query above generates the HTTP GET NCSS Request URL (note that this request can be obtained **directly** from the NCSS GUI page and can act as a standard for other requests) :

```
https://thredds.daac.ornl.gov/thredds/ncss/ornl/daac/1328/1988/daymet_v3_tmin_1988_na.nc4?var=lat&var=lon&var=tmin&north=36.61&west=-85.37&east=-81.29&south=33.57&disableProjSubset=on&horizStride=1&time_start=1988-01-01T12%3A00%3A00Z&time_end=1988-12-30T12%3A00%3A00Z&timeStride=1&accept=netcdf4
```

It is possible to simply copy and paste this URL into a web browser to issue the NCSS subset request. Knowing the structure of this http request allows simple modifications to be made for different queries. For example, a different netCDF file can be queried or changes can be made to the http request parameters to download different subsets.

### Parameters

In addition to the NCSS URL, parameters after the “?” provide the subsetting information. The parameters accepted by the NCSS are described below:

**var:** This parameter allows users to specify the netCDF variables, including both coordinate variables and Daymet data variables, to subset. Multiple “var=VARNAME” pairs can appear in the HTTP GET request URL if a subset includes multiple netCDF variables.

**north:** The northern extent of the bounding box (latitude in decimal degrees) of the subset

**west:** The western extent of the bounding box (longitude in decimal degrees) of the subset

**east:** The eastern extent of the bounding box (longitude in decimal degrees) of the subset

**south:** The southern extent of the bounding box (latitude in decimal degrees) of the subset

**horizStride:** Will take every nth point (in both x and Y) of the gridded dataset. The default, “1”, will take every point

**time\_start:** The beginning of the time range. Specify a time range subset in the form:

```
YYYY '-' mm '-' dd 'T' hh ':' mm ':' ss Z
```

**time\_end:** The end of the time range. Specify a time range subset in the form:

```
YYYY '-' mm '-' dd 'T' hh ':' mm ':' ss Z
```

**timeStride:** Will take only every nth time in the available series on gridded datasets. The default, “1”, will take every time step.

**accept:** The format of the subset data returned by the NCS. Two options are available: “netcdf” for netCDF v3 format and “netcdf4” for netCDF v4 format. The netCDF v4 data format provided by the Daymet NCS has internal compression enabled, so the same subset will create smaller data in netCDF v4 format compared with netCDF v3 format. It’s highly recommended to choose netCDF v4 data format, especially for large subsets, given that Daymet NCS has a size limit on any single subset request.

## Automating Downloads

In general, all Daymet data granules available through the NCS always follow this URL pattern:

```
https://thredds.daac.ornl.gov/thredds/ncss/grid/ornl daac/1328/[YEAR]/daymet_v3_[DAYMETVAR]_[YEAR]_[region].nc4
```

Further automation is possible with a downloading agent such as `wget`. The output filename can be specified after the “-O”, as shown below.

```
wget -O tmin_subset.nc4 "https://thredds.daac.ornl.gov/thredds/ncss/ornl daac/1328/1988/daymet_v3_tmin_1988_na.nc4?var=lat&var=lon&var=tmin&north=36.61&west=-85.37&east=-81.29&south=33.57&disableProjSubset=on&horizStride=1&time_start=1988-01-01T12%3A00%3A00Z&time_end=1988-12-30T12%3A00%3A00Z&timeStride=1&accept=netcdf4"
```

Users can programmatically change [YEAR] and [DAYMETVAR] and [region] to automatically construct the base URL for each single Daymet granule. Multiple URL submissions can be programmatically issued by dynamically updating parameters including the Daymet file, longitude and latitude, and start and end times. Stride should be kept equal to 1 to maintain the daily temporal resolution.

Below is an example bash script to automate the download of years 1980 and 1983 of Daymet data for three variables (tmin, tmax, and prcp) of the example Great Smoky Mountains National Park spatial subset from above.

```

#!/bin/bash
for year in {1980..1983..1}; do
    for par in tmin tmax prcp; do
        if [ $(( $year % 4 )) -eq 0 ]; then
            wget -O ${par}_${year}GSMNPsubset.nc4
"https://thredds.daac.ornl.gov/thredds/ncss/grid/ornl daac/1328/${year}/daymet_v3_${par}
}_${year}_na.nc4?var=lat&var=lon&var=${par}&north=36.61&west=-85.37&east=
81.29&south=33.57&horizStride=1&time_start=${year}-01-01T12:00:00Z&time_end=${year}-
12-30T12:00:00Z&timeStride=1&accept=netcdf4"

        else
            wget -O ${par}_${year}GSMNPsubset.nc4
"https://thredds.daac.ornl.gov/thredds/ncss/grid/ornl daac/1328/${year}/daymet_v3_${par}
}_${year}_na.nc4?var=lat&var=lon&var=${par}&north=36.61&west=-85.37&east=
81.29&south=33.57&horizStride=1&time_start=${year}-01-01T12:00:00Z&time_end=${year}-
12-31T12:00:00Z&timeStride=1&accept=netcdf4"

        fi
    done;
done

```

## Download Size Limitations

The current Daymet NCSS has a size limit of 6GB for each single subset request. NCSS subset requests that result in data exceeding this size limit will return an error message, instead of the actual subset data. It's highly recommended to choose netCDF v4 data format, especially for large subsets, since the same subset data result will have a smaller size in netCDF v4 format.