Global Precipitation Measurement (GPM) integration into i-Tree Eco

5/17/2022

Satoshi Hirabayashi[1]

[1] The Davey Institute, 5 Moon Library, State University of New York, Syracuse, New York 13210, United States

1. Introduction

Hourly weather data (i.e., temperature, dew point temperature, wind speed, pressure, cloud ceiling, sky cover, and precipitation) across a year are used in i-Tree Eco to calculate annual ecosystem services provided by trees. Such data are globally available from National Centers for Environmental Information (NCEI) (formerly known as National Climatic Data Center: NCDC) of National Oceanic and Atmospheric Administration (NOAA) [1]. The NCEI's worldwide weather data from 2005 through 2019 are hosted in i-Tree Eco's server computer, The same weather data are used by other i-Tree Research Suite applications, such as i-Tree Hydro, i-Tree CoolAir, and i-Tree Energy.

Precipitation data is an important variable in these i-Tree applications to calculate removals of particulate matter less than 2.5 micron ($PM_{2.5}$) from the atmosphere as well as other hydrologic estimates including rainfall interception and avoided runoff. Although the precipitation data are generally available in the measurement stations located in the United States, they tended to be missing or incomplete in other countries, which in turn results in incomplete estimates of the ecosystem services mentioned above in these countries.

To complement the missing or incomplete precipitation measurements, Global Precipitation Measurement (GPM) data for 2015 and 2020 from Goddard Earth Sciences Data and Information Services Center (GES DISC) of National Aeronautics and Space Administration (NASA)[2] were employed and incorporated into the NCEI weather data in the i-Tree Eco's server computer.

This document briefly describes the GPM data and its acquisition, manipulation, and validation. NECI weather stations to which the GPM data were incorporated were shown in maps.

2. NOAA National Centers for Environmental Information (NCEI) Weather Data

Fig. 1 presents worldwide weather station locations (19,775 stations) available at the NCEI web site[1]. Weather stations with no records for required variables across a year were excluded, resulted in 8,073 and 8,899 stations for the year 2015 and 2020 (Figs. 2 and 3), respectively, and hosted in the i-Tree Eco server computer. The weather stations are uniquely identified by a combination of 6-digits US Air Force Catalog Station Number (USAF) and 5-digits Weather Bureau Army Navy (WBAN) identifiers.

The quality of the data in the i-Tree Eco server are classified as “Good”, “Fair” or “Poor” in its completeness of the weather measurement availability[3].

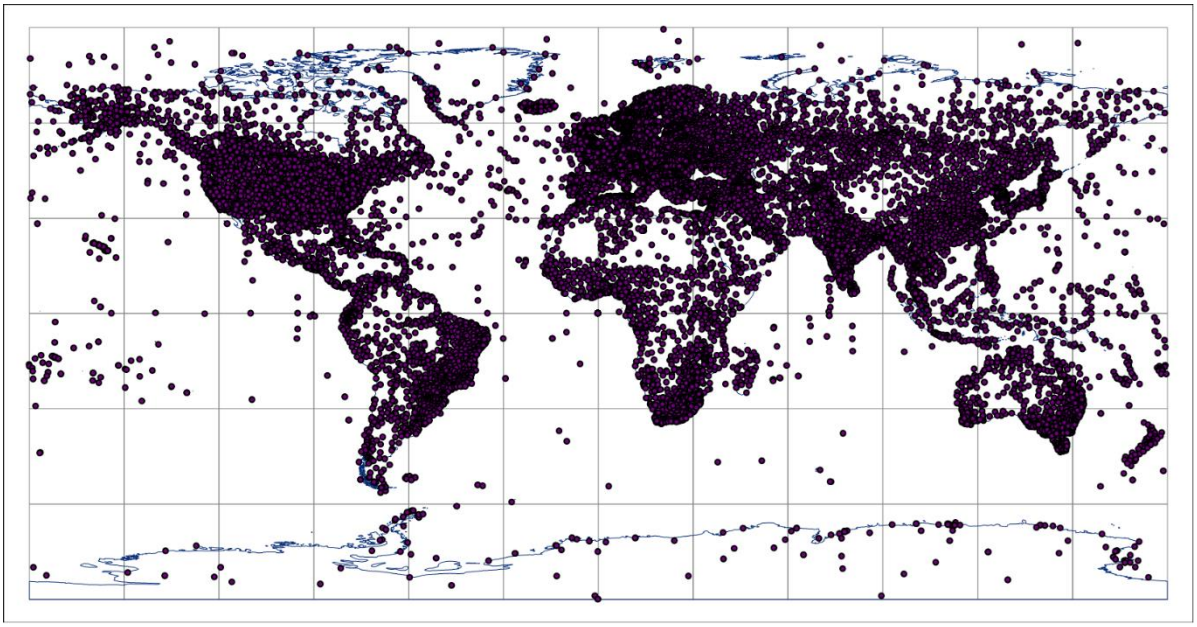


Figure 1 Worldwide weather station (19,775 stations) available from NCEI

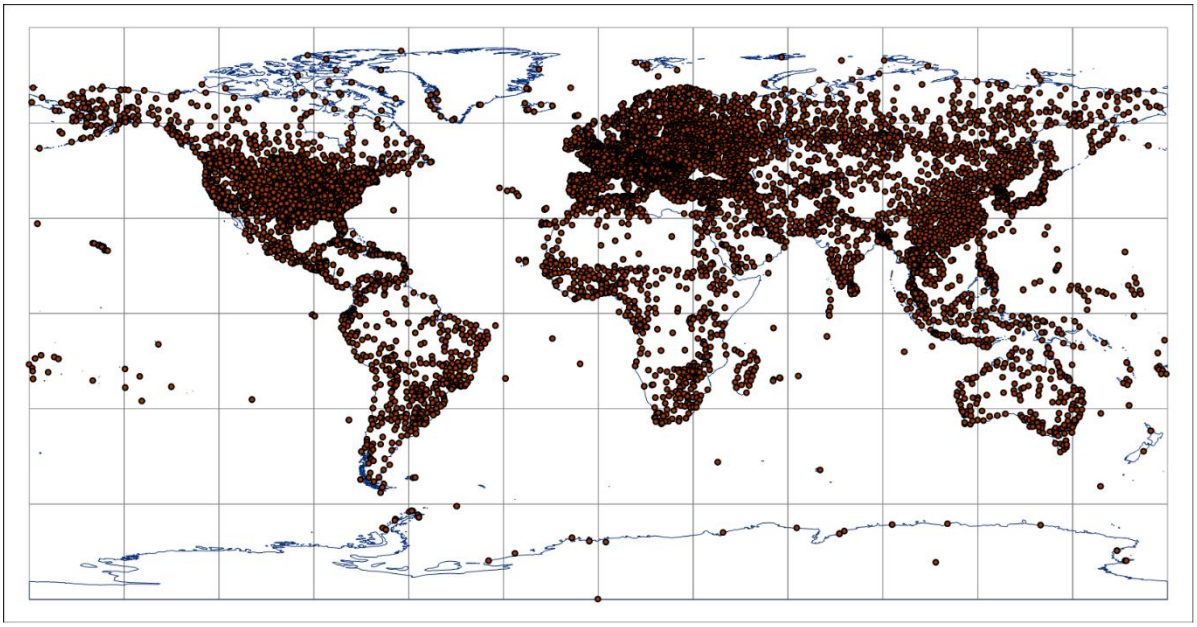


Figure 2 Worldwide NCEI weather stations for 2015 (8,073 stations) hosted in i-Tree Eco server computer

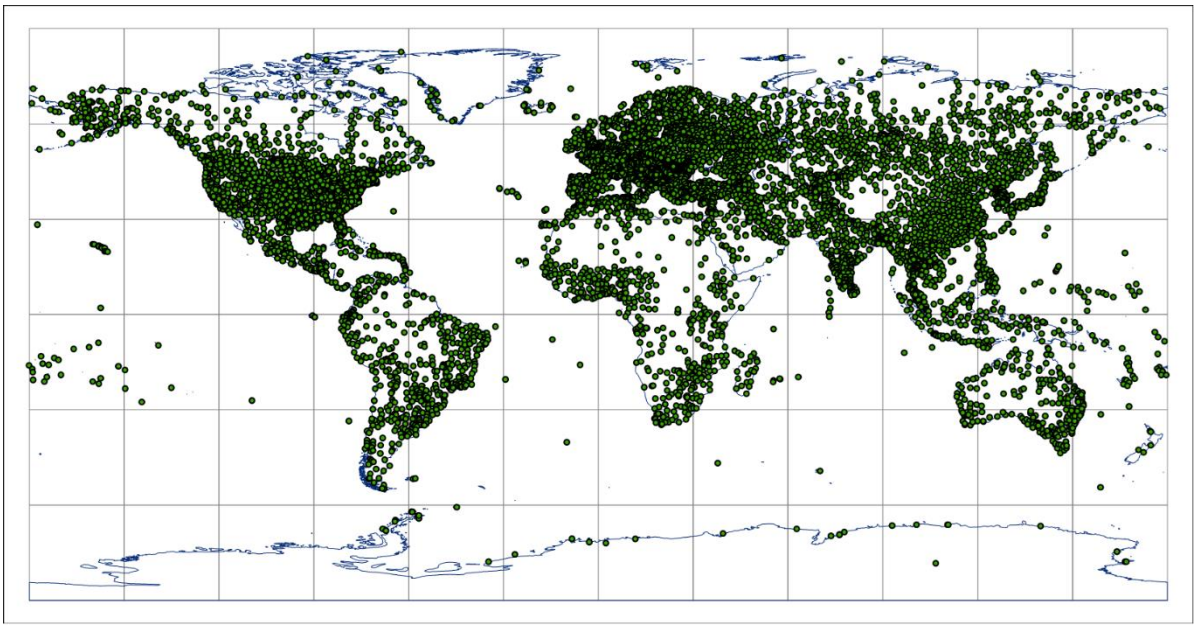


Figure 3 Worldwide NCEI weather stations for 2020 (8,899 stations) hosted in i-Tree Eco server computer

3. Global Precipitation Measurement (GPM)

3.1. Data description

Global Precipitation Measurement (GPM) data employed is GPM IMERG (The Integrated Multi-satellitE Retrievals for GPM) Final Precipitation L3 Half Hourly 0.1 degree × 0.1 degree V06 (GPM_3IMERGHH) constructed based on precipitation estimates from the various precipitation-relevant satellite passive microwave (PMW) sensors. Spatial resolution is 0.1°×0.1° (roughly 10x10 km), while the temporal resolution is half hour. "Final" version of IMERG is satellite-gauge product ~3.5 months after the observation month, using both forward and backward morphing and including monthly gauge analyses[2]. Fig. 4 presents an example of the GPM data.

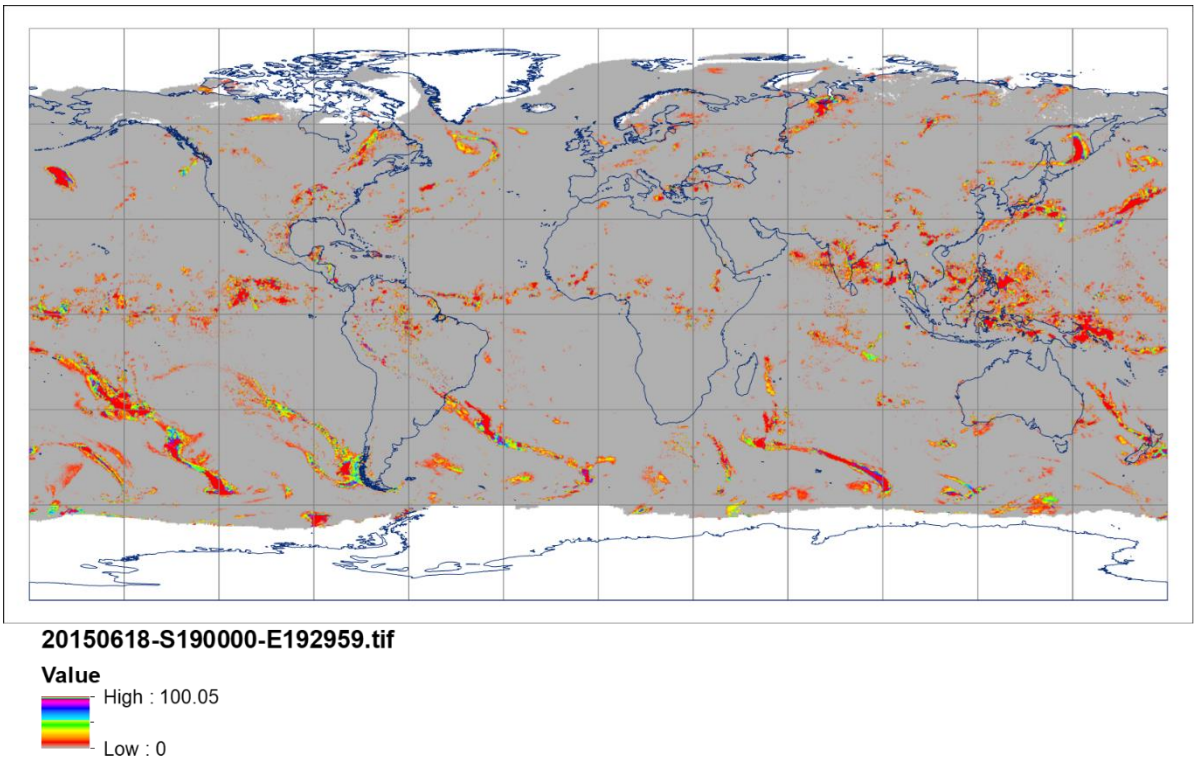


Figure 4 GPM data for June 18th, 2015 from 7 pm to 7:30 pm

3.2. Data Acquisition

GPM IMERG Final Precipitation L3 Half Hourly data were obtained from NASA's web site[2]. Fig. 5 presents the setting for retrieving data from 1/1/2015 through 12/31/2015 for the globe.

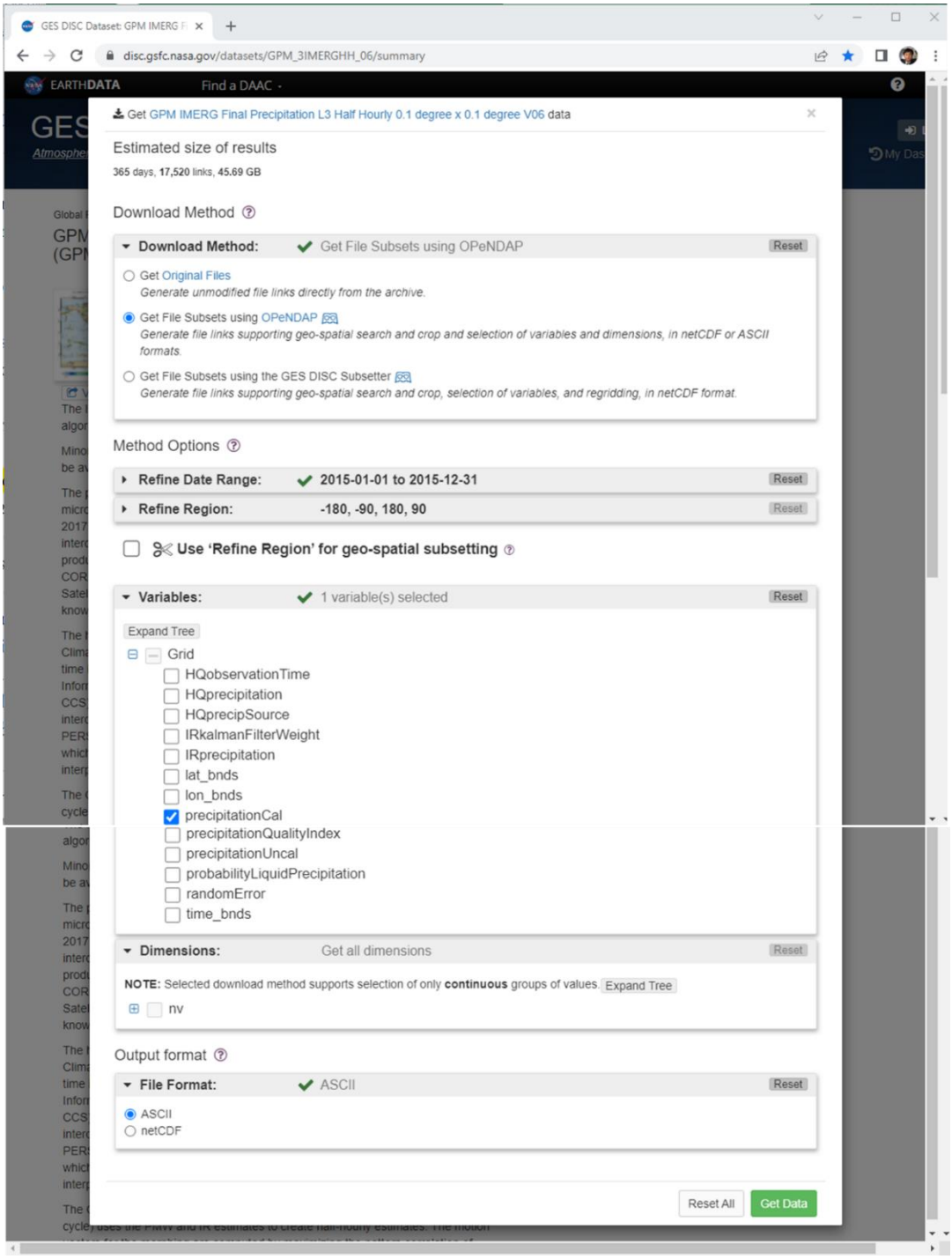


Figure 5 User interface for GPM data retrieval setting

Once the data request was accepted by the NASA's site the download link list is created as shown in Fig. 6. The downloaded list is shown in Fig.7, in which URLs for the annual half-hourly GPM data (totally 48 data/day × 365 = 17,520 files) are listed. Using wget utility, the data were retrieved from the site.

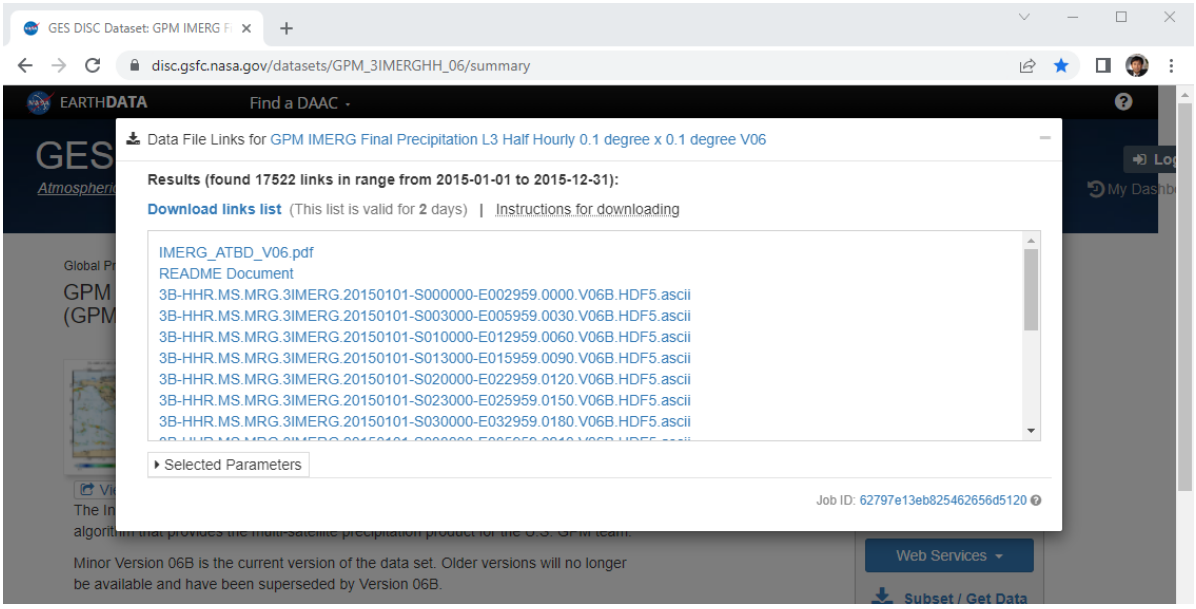


Figure 6 User interface for GPM data download list

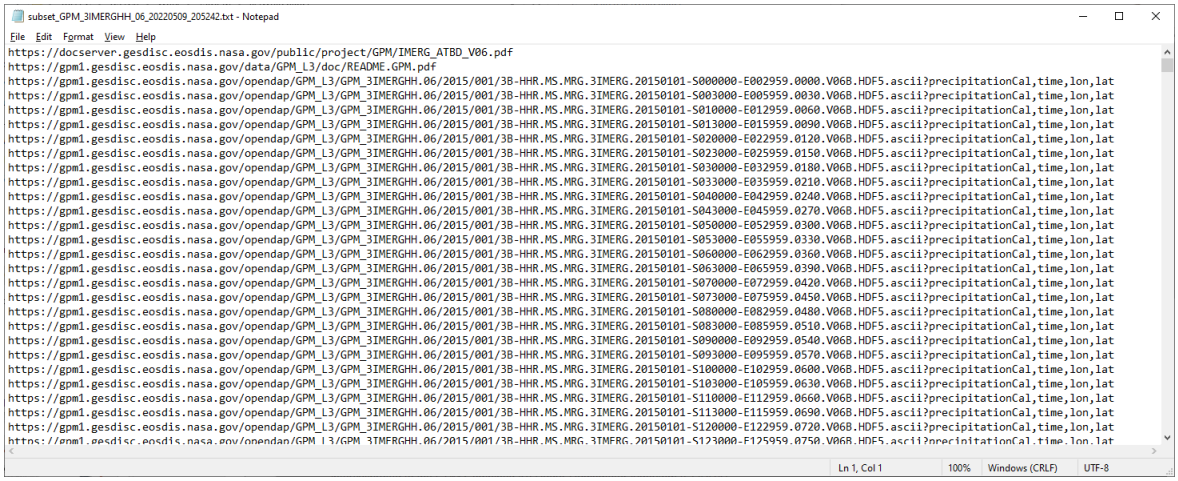


Figure 7 GPM data download URLs

3.3. Data Manipulation

Fig. 8 presents an example of the downloaded raw GPM data. There are 3600 data rows (i.e., precipitationCal[0][0] to precipitationCal[0][3599]) where each row contains 1800 values. Given the horizontal resolution of 0.1 arc degrees, the rows represent longitude from -180 to 180 degrees, whereas the columns represent latitude from -90 to 90 degrees. The unit for the precipitation is mm/hr. The raw data is converted to an ACSII-based raster presented in Fig. 9 by adding header lines defining the number of columns, number of rows, coordinates at the lower left corner, cell size (in arc degrees), and NODATA value. Columns and rows in the raw data were swapped to fit into the coordinates defined in the header.

Figure 8 An example of raw GPM data

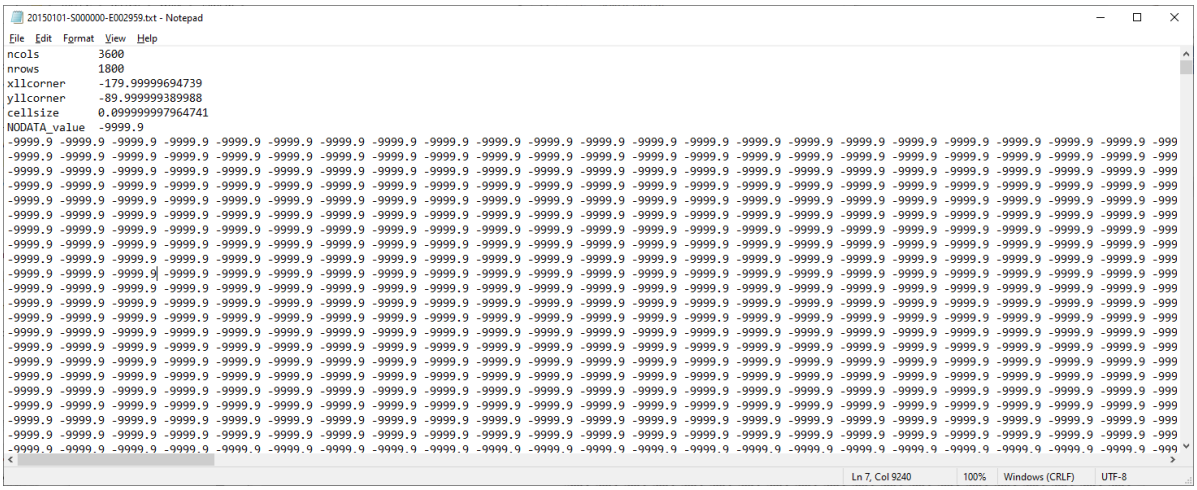


Figure 9 An example of ASCII-based raster data

The half-hourly ASCII-based raster data were then converted to a GEOTIFF raster as exemplified in Fig. 4. By overlaying weather station locations shown in Figs. 2 and 3 onto these GEOTIFF rasters, the precipitation for each half-hour were extracted across one year of 2015 and 2020. Some of the GPM raster data had no values at some weather station locations. These stations were excluded and hourly precipitation records were created as a csv file for the remaining stations where all of 17,520 half-hourly precipitation data were available (Figs. 10 and 11). As can be seen in these figures, mostly weather stations located higher than 60 degrees North were excluded. Fig.12 presents an example of the created csv file named with UUUUUU-WWWWW-YYYY.csv, where UUUUUU represents 6-digits USAF, WWWWW represents 5-digits WBAN, and YYYY represents 4-digits year. Table 1 summarizes the number of weather stations.

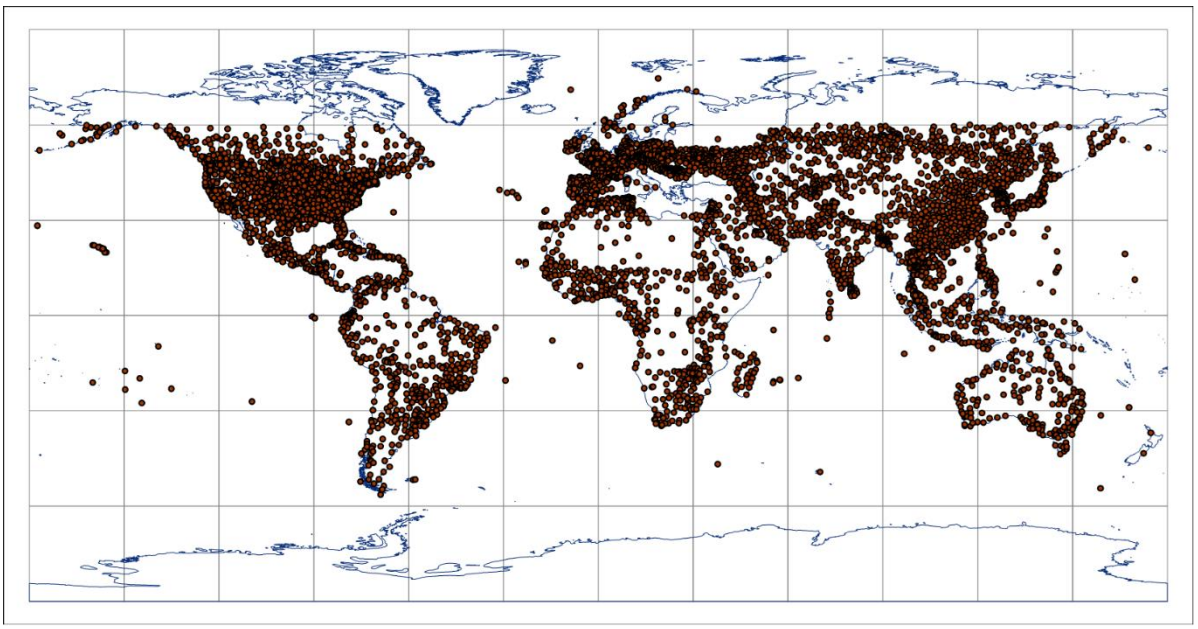


Figure 10 2015 NCEI weather stations (6,328 stations) to which GPM data were extracted

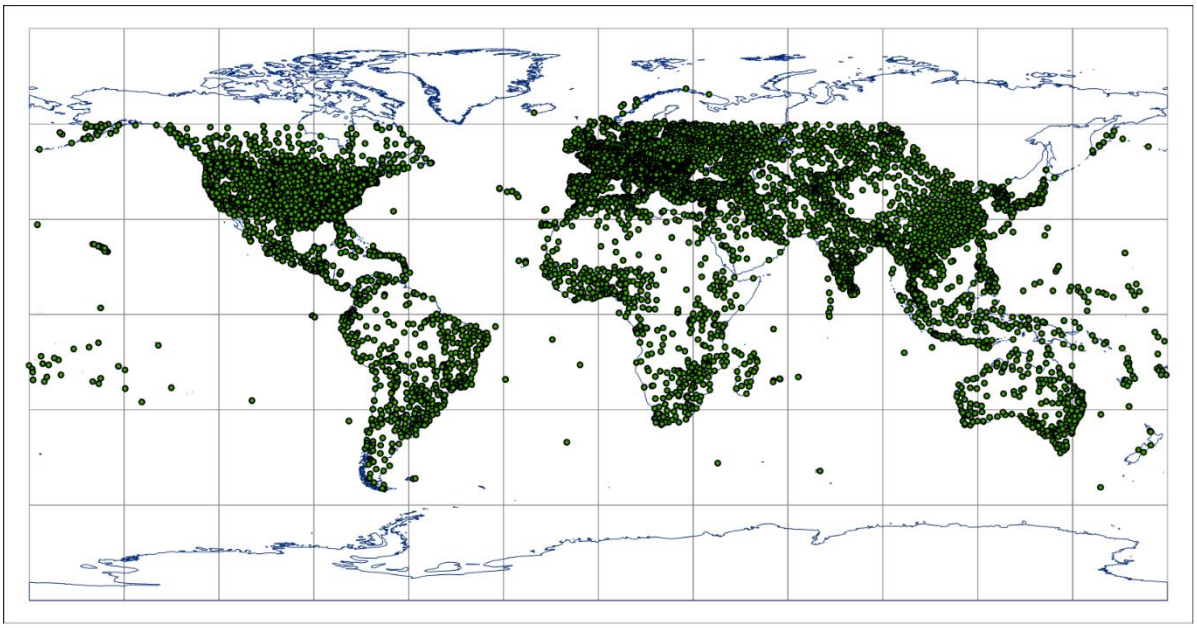
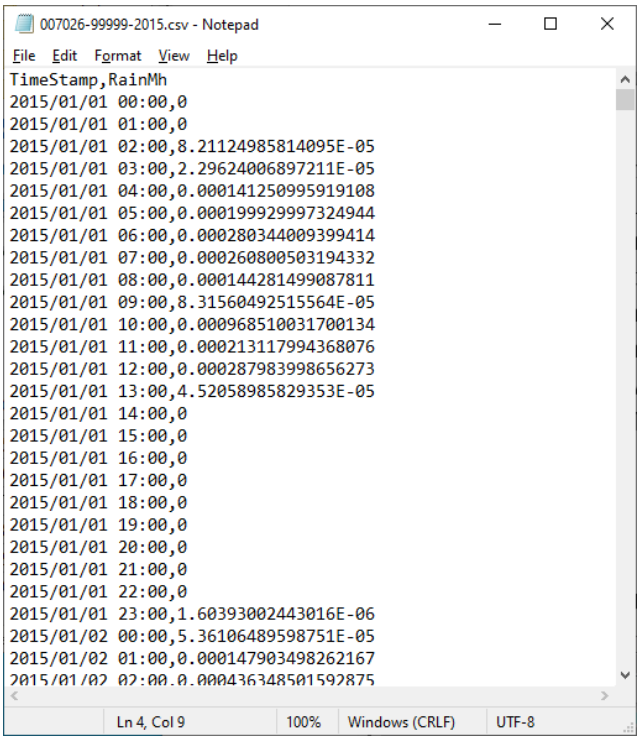


Figure 11 2020 NCEI weather stations (7,868 stations) to which GPM data were extracted



007026-99999-2015.csv - Notepad

File Edit Format View Help

```
TimeStamp,RainMh
2015/01/01 00:00,0
2015/01/01 01:00,0
2015/01/01 02:00,8.21124985814095E-05
2015/01/01 03:00,2.29624006897211E-05
2015/01/01 04:00,0.000141250995919108
2015/01/01 05:00,0.000199929997324944
2015/01/01 06:00,0.000280344009399414
2015/01/01 07:00,0.000260800503194332
2015/01/01 08:00,0.000144281499087811
2015/01/01 09:00,8.31560492515564E-05
2015/01/01 10:00,0.000968510031700134
2015/01/01 11:00,0.000213117994368076
2015/01/01 12:00,0.000287983998656273
2015/01/01 13:00,4.52058985829353E-05
2015/01/01 14:00,0
2015/01/01 15:00,0
2015/01/01 16:00,0
2015/01/01 17:00,0
2015/01/01 18:00,0
2015/01/01 19:00,0
2015/01/01 20:00,0
2015/01/01 21:00,0
2015/01/01 22:00,0
2015/01/01 23:00,1.60393002443016E-06
2015/01/02 00:00,5.36106489598751E-05
2015/01/02 01:00,0.000147903498262167
2015/01/02 02:00,0.000436348501592875
```

Ln 4, Col 9 100% Windows (CRLF) UTF-8

Figure 12 An example of created hourly precipitation records

Table 1 The number of weather stations

Year	2015	2020
Worldwide NCEI weather stations	19,775	19,775
Valid NCEI weather stations hosted in i-Tree Eco server	8,073	8,899
NCEI weather stations with GPM data extracted	6,328	7,868

3.4. Data Validation

Annual precipitation (mm) was calculated for each of the weather station locations based on the hourly precipitation records created as csv files for 2015 and 2020. To make comparisons, Global Precipitation Climatology Center (GPCC)'s products for 2015[4] (2020 is not available) as well as the precipitation measurements (PCP01) recorded in the NCEI hourly weather data for 2015 and 2020 were used. Of 6,328 weather stations for 2015, 6,310 stations had the GPCC's annual precipitation. GPM and GPCC data were aligned very well with the Pearson's correlation coefficient (R^2) of over 0.94 (Fig. 13a). For the comparisons between the GPM and NCEI data, the weather stations located in the United States were used as the precipitation data are often missing or incomplete in other countries. Also, to eliminate outliers in the NCEI records, the stations with erroneous precipitation measurement (i.e., less than 100 mm/yr and greater than 3000 mm/yr) were excluded, resulted in 1,614 and 1,708 stations for 2015 and 2020, respectively. The GPM data tended to overestimate the precipitation, but they are relatively in a good fit with the NCEI's annual precipitation. R^2 were around 0.80 for both years (Figs. 13b and 13c).

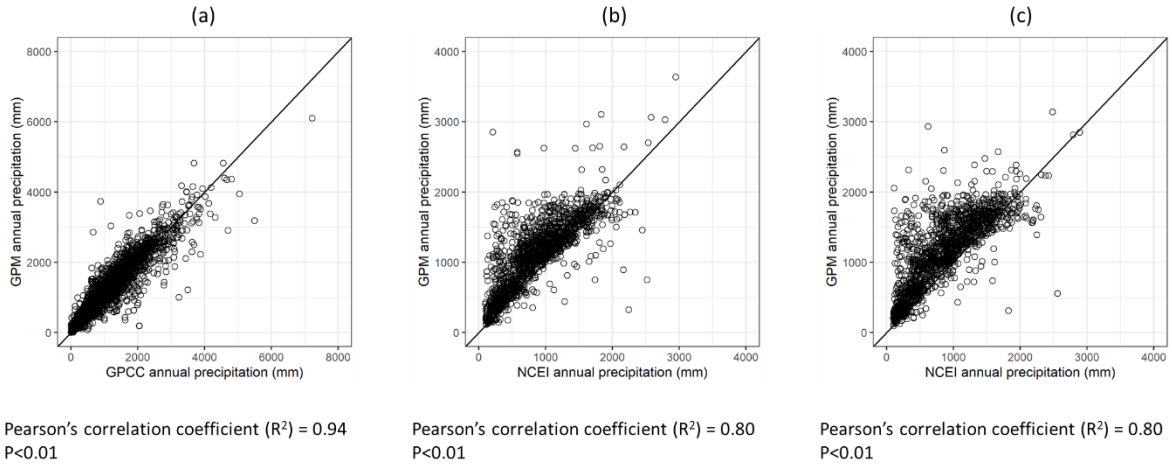


Figure 13 Scatterplots of precipitation data a) GPM vs. GPCC for 2015, b) GPM vs. NCEI for 2015 only for the US stations, c) GPM vs. NCEI for 2020 only for the US stations

References

1. National Centers for Environmental Information (NCEI) NO and AA (NOAA). Integrated Surface Dataset (Global). Published 2022. Accessed May 9, 2022. https://www.ncei.noaa.gov/access/search/data-search/global-hourly

2. Goddard Earth Sciences Data and Information Services Center (GES DISC) NA and SA (NASA). GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06 (GPM_3IMERGHH). Published 2022. Accessed May 9, 2022. https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_06/summary

3. Satoshi Hirabayashi. Weather Station Data Quality Assessment . Published 2017. Accessed May 9, 2022. http://www.itreetools.org/eco/resources/v6/Eco_v6_weather_station_validation_description.pdf

4. Global Precipitation Climatology Center (GPCC) NC for AR. Climate Data. Published 2022. Accessed May 9, 2022. https://climatedataguide.ucar.edu/climate-data/gpcc-global-precipitation-climatology-centre