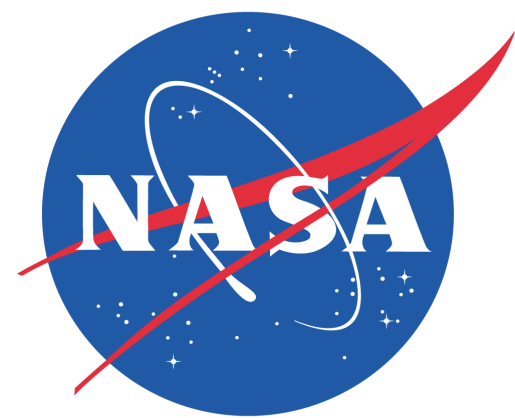




Quantitative Metrics from 20 Years of Terra Data Usage

Lalit Wanchoo^{1,2} (lalit.wanchoo@nasa.gov), Young-In Won^{1,2} (youngin.won@nasa.gov), Durga N. Kafle^{1,2} (durga.n.kafle@nasa.gov), Drew H. Kittel² (drew.h.kittel@nasa.gov)

¹ ADNET Systems, Inc., 7515 Mission Drive, Suite A100, Lanham, MD 20706, ² NASA GSFC, Code 423, Greenbelt, MD 20771

Introduction

NASA's Terra flagship satellite carries five Earth-observing instruments (Table 1) that have collected data for almost 20 years. NASA's Earth Science Data and Information System (ESDIS) Project makes these data, along with derived products, available to worldwide data users. Since the launch of Terra on December 18, 1999, more than 10,000 data products have been archived and distributed by NASA-funded Distributed Active Archive Centers (DAACs) that are part of NASA's Earth Observing System Data and Information System (EOSDIS). At the end of the 2019 Fiscal Year, about 3,000 Terra data products constituted almost 20% of the entire EOSDIS data archive volume (6 PB out of approximately 34 PB of total EOSDIS archive), and 6 PB of Terra data were distributed to over half-a-million public users worldwide. Raw data (level 0) is processed at full instrument resolution to other higher levels products to provide more useful parameters and data formats. A description of product levels is given in Table 2 and Table 3 lists the NASA's EOSDIS DAACs that are responsible for archiving and distributing of EOS Terra mission data.

Table 1: Terra Instruments Names

Instrument	Name
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
CERES	Clouds and Earth's Radiant Energy System
MISR	Multi-angle Imaging SpectroRadiometer
MODIS	Moderate-resolution Imaging Spectroradiometer
MOPITT	Measurements of Pollution in the Troposphere

Table 2: Data Product Level Definitions

Data Product Level	Description
Level 0	Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communications artifacts (e.g., synchronization frames, communications headers, duplicate data) removed. (In most cases, the EOS Data and Operations System (EOS) provides these data to the data centers as production data sets for processing by the Science Data Processing Segment (SDPS) or by a SIPS to produce higher-level products.)
Level 1A	Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended but not applied to Level 0 data.
Level 1B	Level 1A data that have been processed to sensor units (not all instruments have Level 1B source data).
Level 2	Derived geophysical variables at the same resolution and location as Level 1 source data.
Level 3	Variables mapped on uniform space-time grid scales, usually with some completeness and consistency.
Level 4	Model output or results from analyses of lower-level data (e.g., variables derived from multiple measurements).

Reference: EarthData Website, <https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy/data-levels>

Table 3: EOSDIS Distributed Active Archive Centers for Terra Data

DAAC	Description
ASDC	Atmospheric Science Data Center
GES DISC	Goddard Earth Sciences Data and Information Services Center
GHRC	Global Hydrology Resource Center
LAADS	Level 1 and Atmosphere Archive and Distribution System DAAC
LPDAAC	Land Processes DAAC
NSIDC	National Snow and Ice Data Center DAAC
OB.DAAC	Ocean Biology DAAC
ORNL	Oak Ridge National Laboratory DAAC
PO.DAAC	Physical Oceanography DAAC

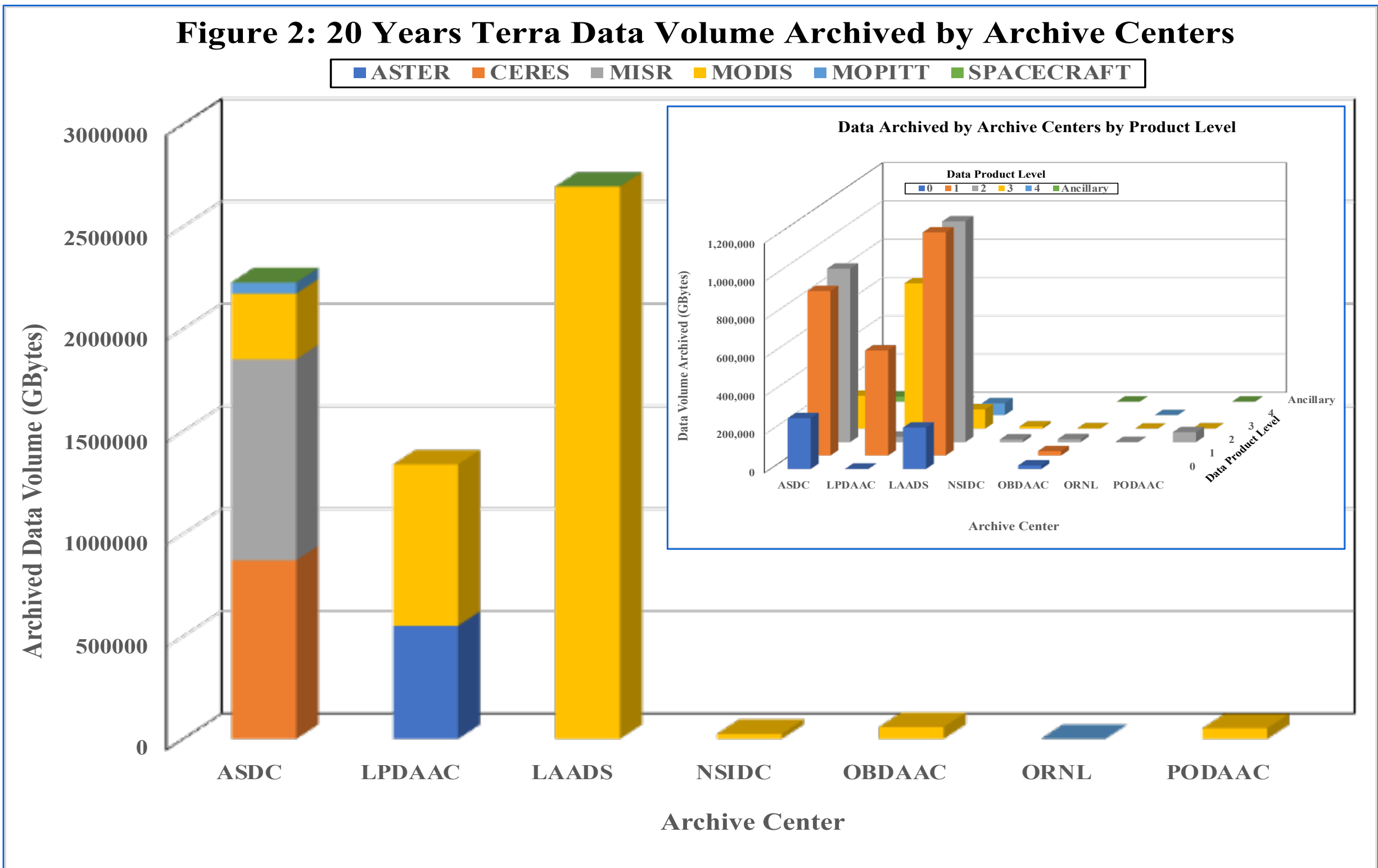
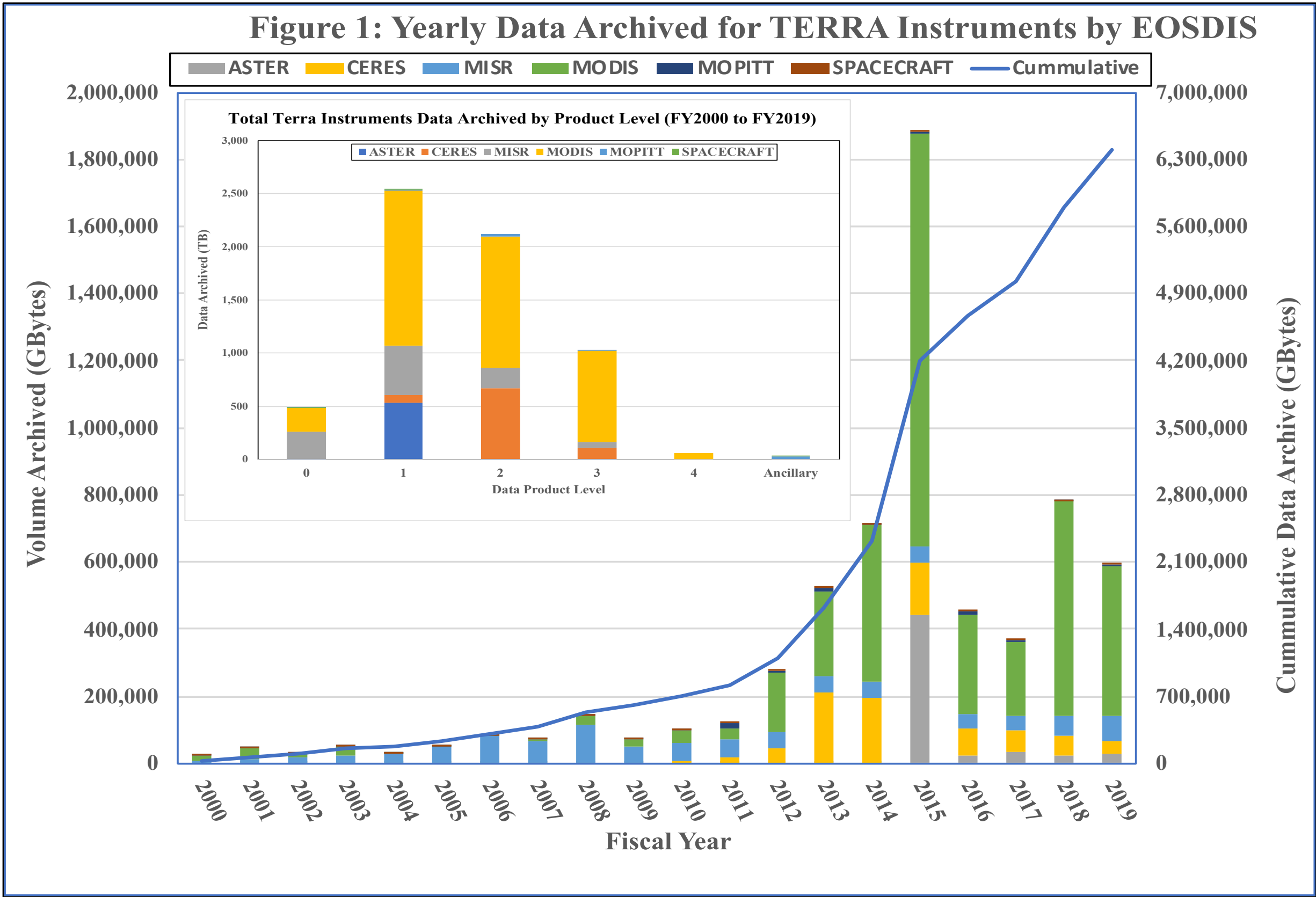
Purpose

By categorizing the Terra data products and their distribution, we can get a quantitative assessment of Terra data usage. NASA's EOSDIS Project has collected archive, distribution, and user information from EOSDIS data users since February 2000 when public distribution began right after on-orbit commissioning. These metrics are available through the EOSDIS Metrics System (EMS). EMS information is stored in a relational database from which quantitative metrics of Terra data were retrieved and analyzed for this study.

The purposes of this study are to: 1) perform a comprehensive investigation of the 20-year trend in the archive and distribution of Terra data products; 2) identify and characterize data product usage over the last 20 years; and 3) identify and characterize the global user community for these data.

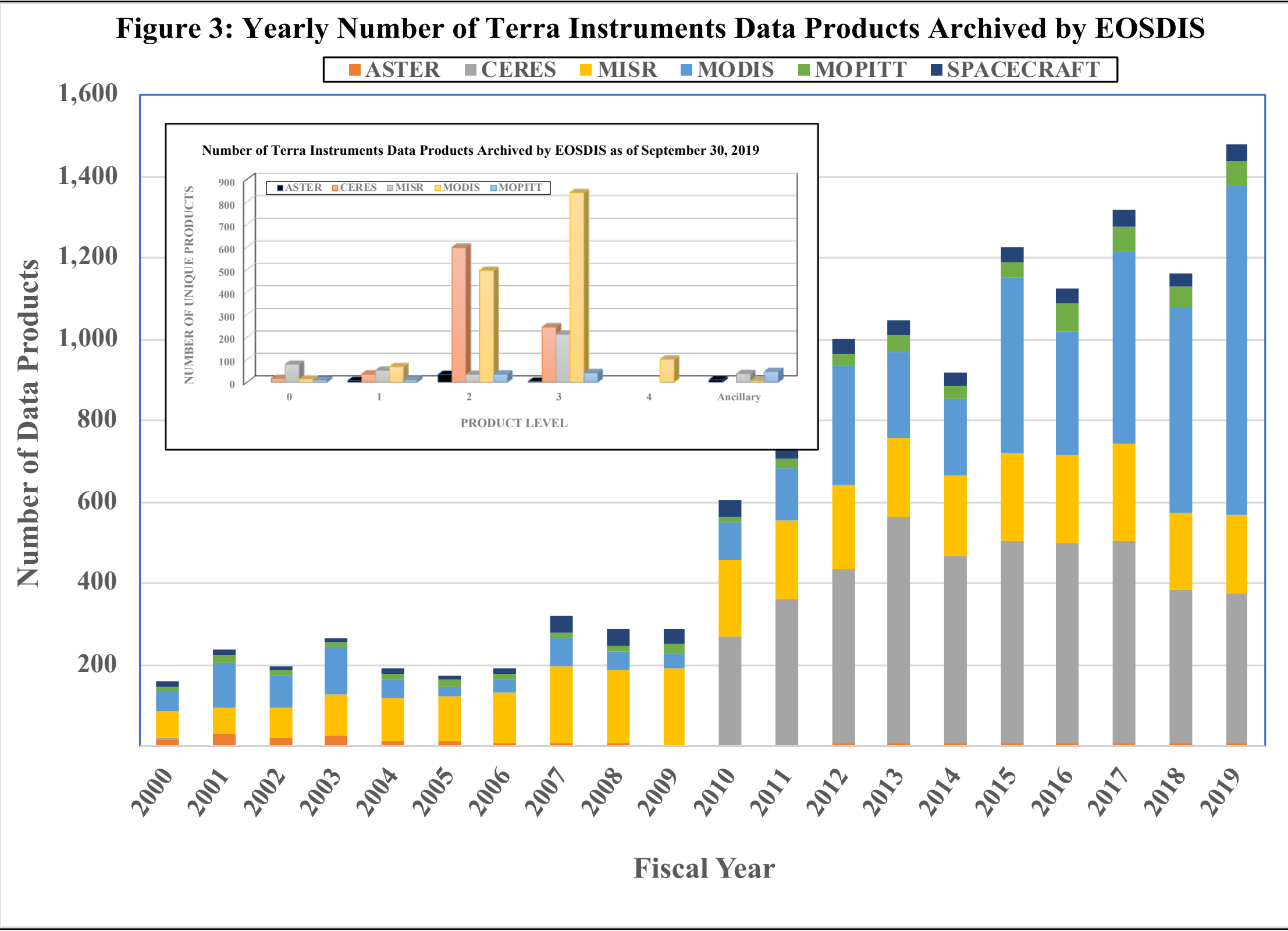
Archive Terra Data

NASA EOSDIS Distributed Active Archive Centers (DAACs) store data products that they produce or that they receive, and facilitate the distribution of these data products to end users. This means that the total EOSDIS data archive volume is distributed over multiple DAACs. On a daily basis, EOSDIS DAACs process an average of more than 70 Gigabytes (GB) of Level 0 Terra data from all five instruments to produce more than 800 GBs of Level 1 through Level 4 data products. The 20 years of the Terra mission have seen the addition of new products and multiple reprocessing of all Terra data, which further increases the Terra data archived by the DAACs. Figures 1 and 2 show how the DAACs have been achieving Terra data over 20 years, with a majority of these data being Level 1 and Level 2 data from the CERES and MODIS instruments. Over 98% of all Terra data is archived at three DAACs: 1) ASDC (CERES, MISR, MOPITT), 2) LP DAAC (ASTER, MODIS), and 3) LAADS DAAC (MODIS). Terra data are archived and made available to users through various distribution services at their respective DAAC.



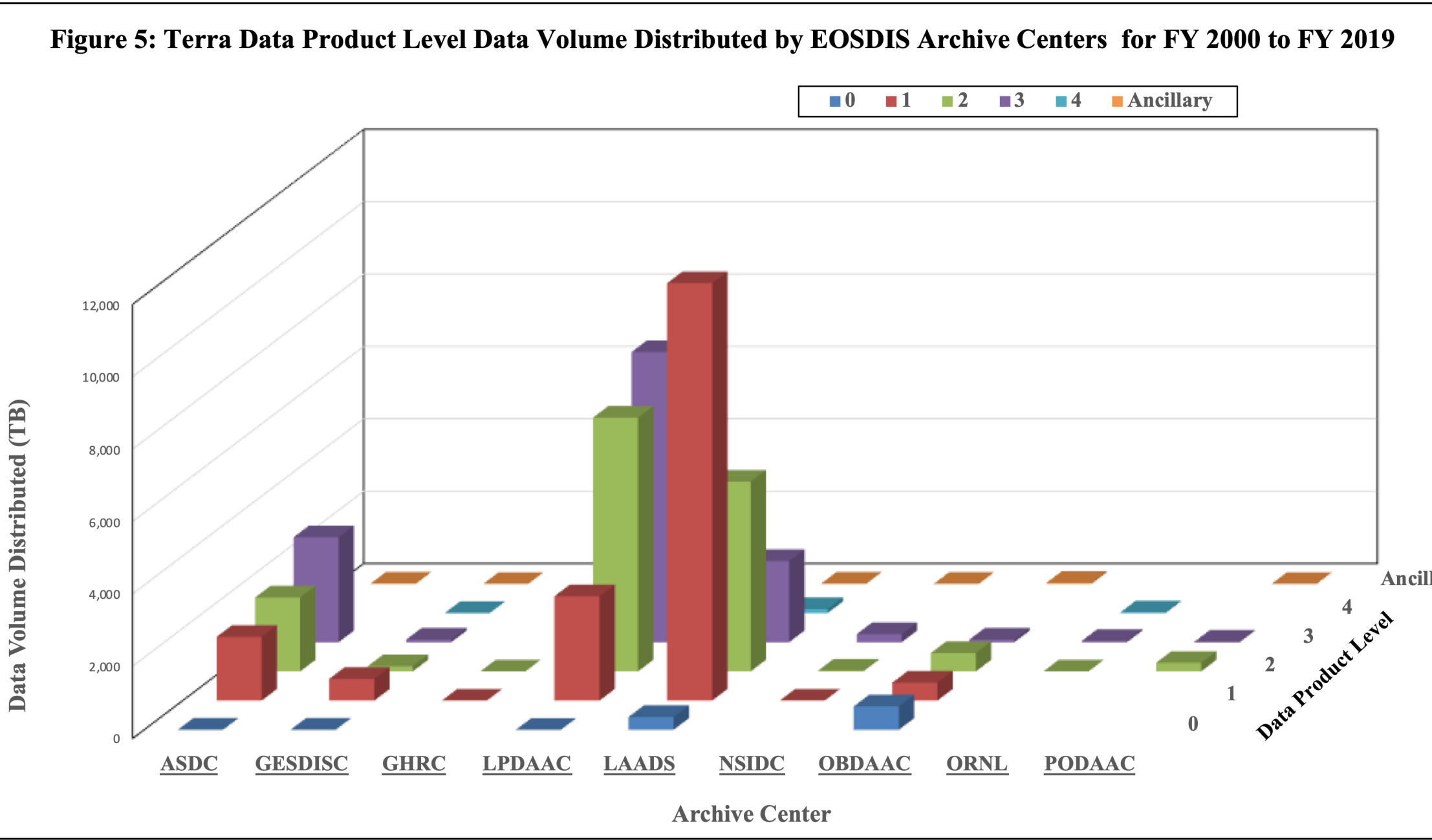
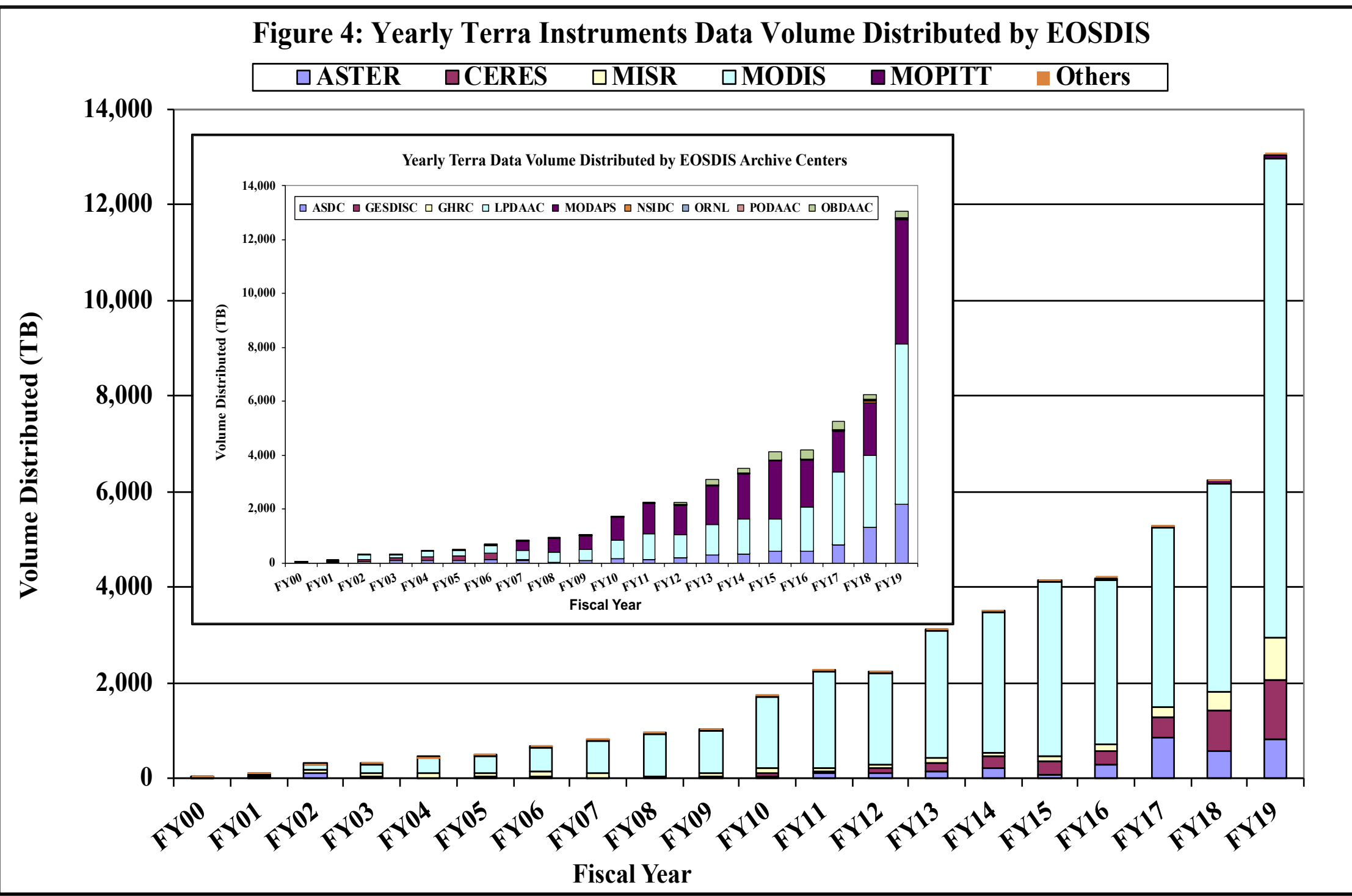
Number of Data Products Archive

The number of Terra data products has increased every year over the last 20 years (Figure 3). More than 3,000 data products are currently archived by all EOSDIS DAACs. The increase in data products during 2011 is due to addition of CERES data to the EOSDIS Metrics System (EMS) along with the reprocessing of MODIS data that led to the creation of new MODIS data products. Another reason for the increase in data products in 2011 is that Level 2 data products were produced routinely during the early years after Terra's launch while Level 3 and Level 4 products were not produced until the algorithms were more robust and the Terra instruments were stable.



Terra Data Distribution

EOSDIS DAACs provide free access to all data in the EOSDIS collection to worldwide data users and constantly evolve services to make discovering and accessing these data as simple and efficient as possible. The volume of Terra data distributed has steadily increased over the 20 years of the mission, as shown in Figure 4. Since the launch of Terra, more than 50 Petabytes (PB) of Terra instrument data have been distributed to worldwide data users. The majority of these data are from the MODIS instrument, which are distributed by NASA's LAADS DAAC and LP DAAC (Figure 4). A significant increase in the volume of data distributed by NASA's LP DAAC for one MODIS Level 4 data product during FY 2019 is seen in Figure 4. Figure 5 shows that most distributed data products are higher level products (Levels 1-3), indicating that there is more usage of derived geophysical parameters than of the lower level radiance parameters. Also, these figures show the increased usage of higher-level products over time, which indicates enhanced data quality.

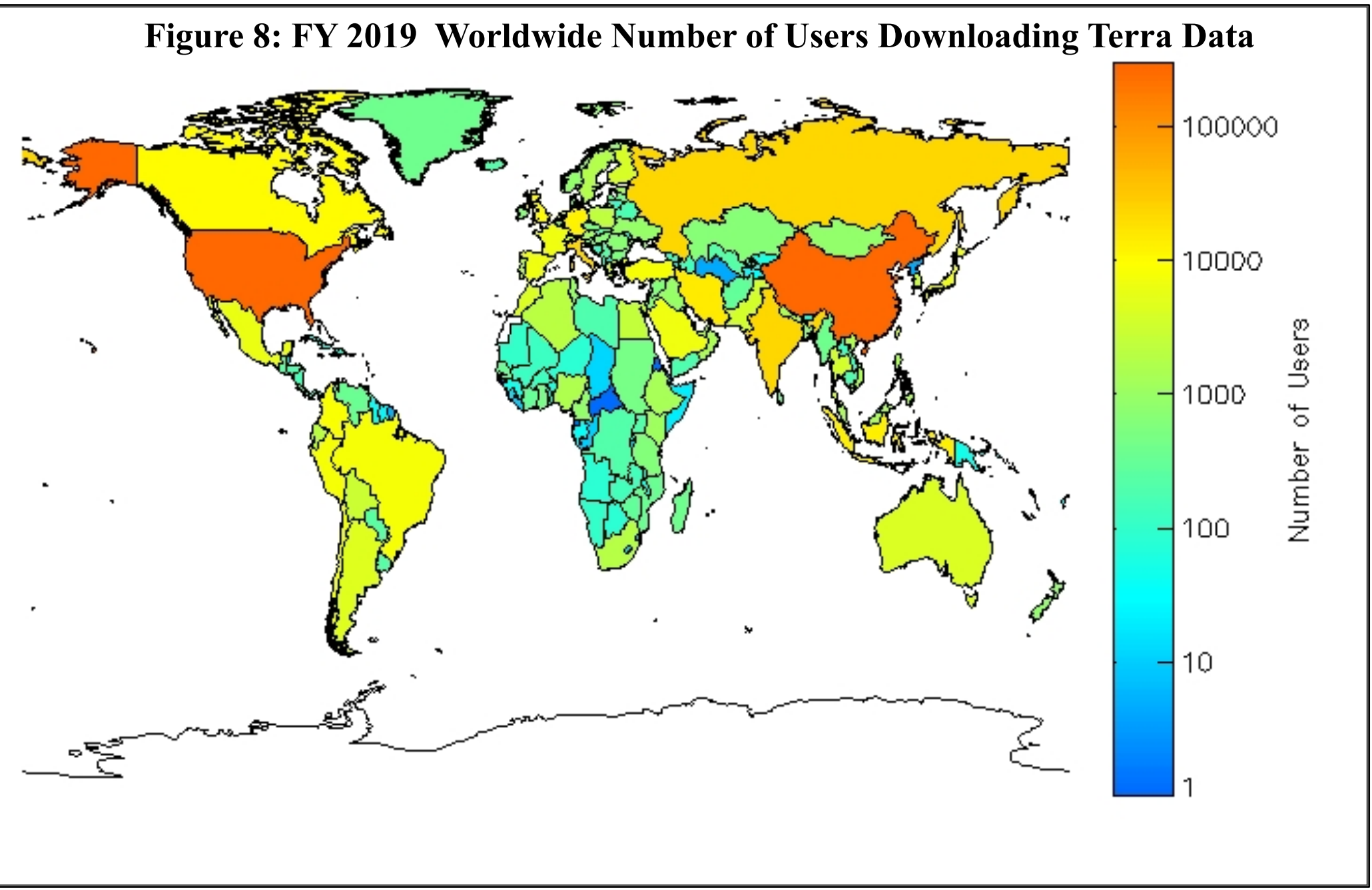
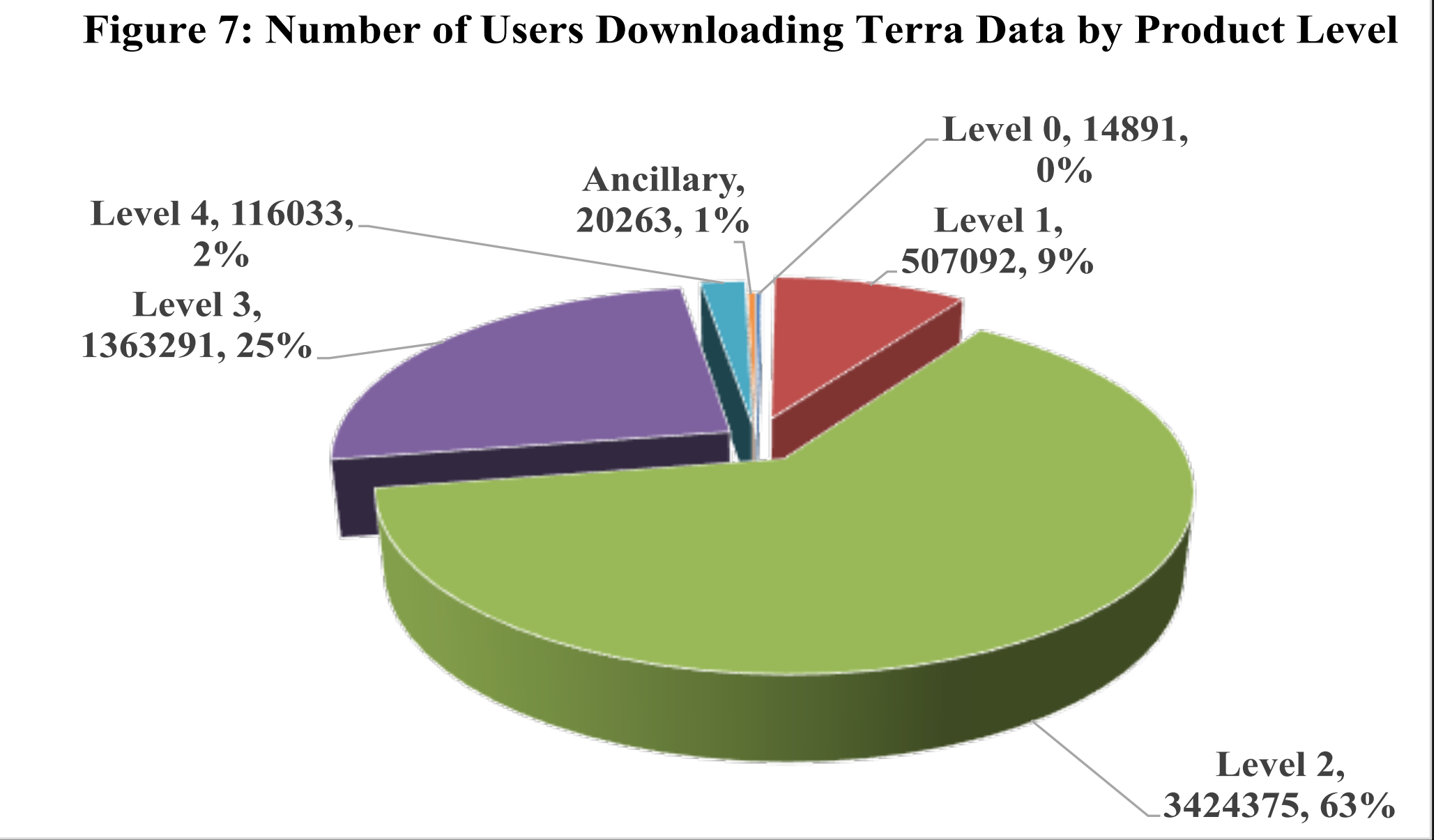
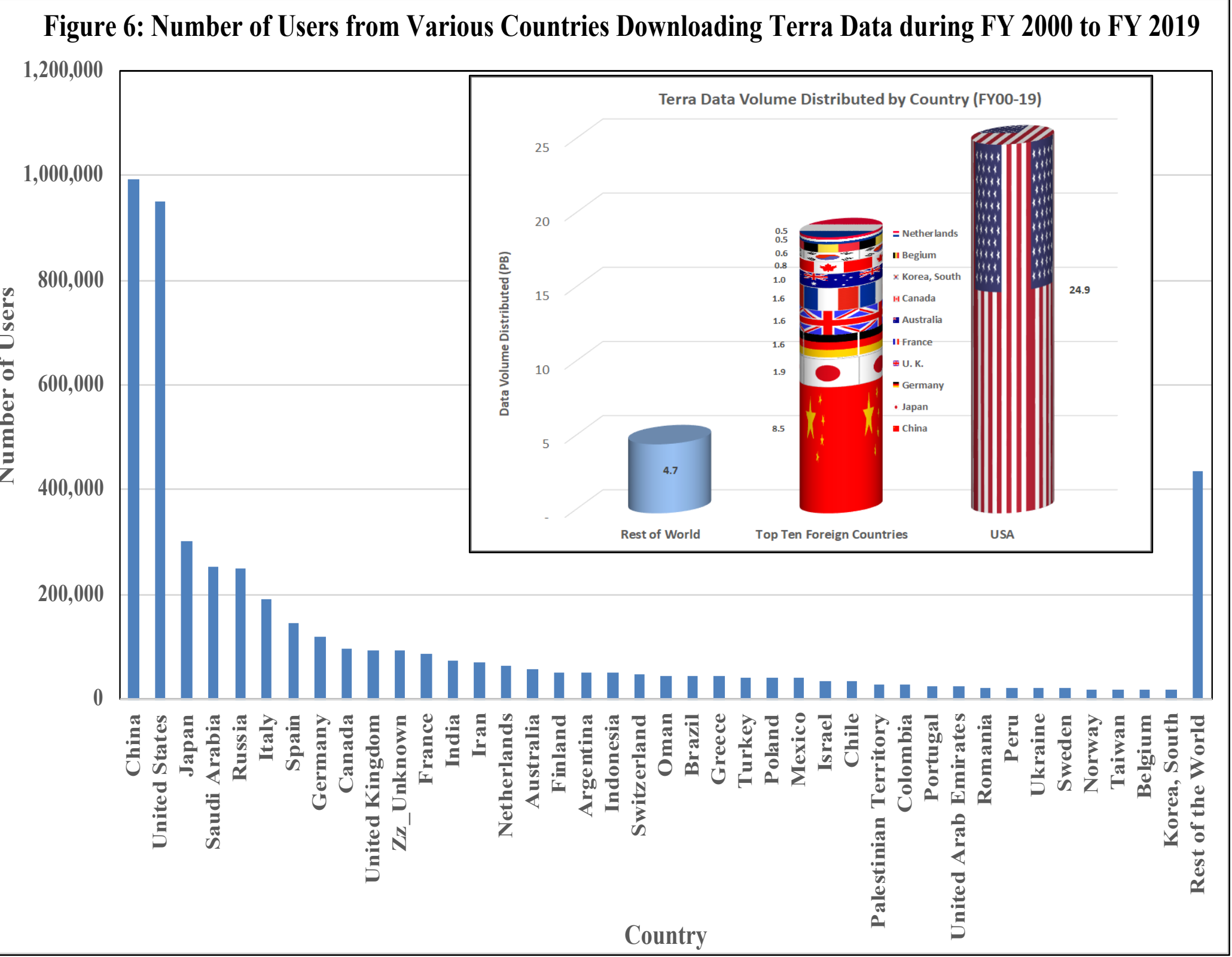


Acknowledgments

This study was funded through the science and Exploration Data Analysis (SESDA IV) GSFC NASA contract No: 80GSFC17C0003

Terra Data User Characterization

The Terra data user community has evolved significantly since the start of the mission, and the number of users downloading Terra data has grown from approximately 500 in FY 2000 to close to 500,000 in FY 2019. Figure 6 illustrates the distribution of Terra data to more than five million users from the United States and foreign countries over the last 20 years. A majority of Terra data was downloaded by U.S. users, 86% of whom downloaded Level 2 and Level 3 data products (Figure 7). Figure 8 shows that Terra data were downloaded by world-wide users.



Summary

Figure 9 shows how the archiving and distribution of Terra data have evolved over last 20 years. A comparison of FY 2000 and FY 2019 data metrics shows the following: 1) From only a few hundred data products, the Terra archive has grown to more than 3,000 data products, 2) The distribution of Terra instrument data has grown from 18 Terabytes (TB) to 13 PB per FY, and 3) The number of users downloading data has grown from about 500 to 500,000, with data being distributed globally. This global distribution of Terra data is possible due to NASA's open data policy, which ensures that these data are freely distributed. These metrics for Terra data and data distribution provide insight to the innovation and work of EOSDIS DAACs in managing the growth of the Terra data archive and facilitating the efficient distribution of these data. EOSDIS DAACs continue to provide guidance for the requirements of future data systems that will be needed to effectively and efficiently handle increasing amounts of data produced by Terra as well as on-going and future Earth science missions.

