

CDDIS Global Data Center Technical Report 2012

C. Noll

NASA Goddard Space Flight Center, Code 690.1
Greenbelt, MD 20771 USA
Carey.Noll@nasa.gov

1 Introduction

The CDDIS is NASA's data archive and information service supporting the international space geodesy community. For over 30 years, the CDDIS has provided continuous, long term, public access to the data (mainly GNSS-Global Navigation Satellite System, SLR-Satellite Laser Ranging, VLBI-Very Long Baseline Interferometry, and DORIS-Doppler Orbitography and Radiopositioning Integrated by Satellite) and products derived from these data required for a variety of science observations, including the determination of a global terrestrial reference frame and geodetic studies in plate tectonics, earthquake displacements, volcano monitoring, Earth orientation, and atmospheric angular momentum, among others. The specialized nature of the CDDIS lends itself well to enhancement to accommodate diverse data sets and user requirements. The CDDIS is one of NASA's Earth Observing System Data and Information System (EOSDIS) distributed data archive centers; EOSDIS data centers serve a diverse user community and are tasked to provide facilities to search and access science data and products.

The CDDIS serves as one of the primary data centers and core components for the geometric services established under the International Association of Geodesy (IAG), an organization that promotes scientific cooperation and research in geodesy on a global scale. The system has supported the International GNSS Service (IGS) as a global data center since 1992. The CDDIS activities within the IGS during 2012 are summarized below; this report also includes any recent changes or enhancements made to the CDDIS.

2 System Description

The CDDIS archive of IGS data and products are accessible worldwide through anonymous ftp. The CDDIS is located at NASA's Goddard Space Flight Center (GSFC) and is available to users 24 hours per day, seven days per week.

The CDDIS computer system consists mainly of incoming, outgoing, and processing servers. All ftp and web access is performed on the outgoing server. Data centers, stations, and analysis centers push files to the CDDIS incoming server. Processing of incoming files for the on-line archive is performed in a separate environment that also includes a database server for managing metadata extracted from incoming data.

In 2011, the CDDIS staff procured new server hardware to further enhance the capabilities of the system and ensure a robust archive environment. The new system is fully redundant with the primary and secondary/failover system located in different buildings on the GSFC campus. Each system utilizes a distributed functionality (incoming, outgoing, processing, database, and map servers) and is configured with a local backup system as well as a full backup system located in a third building at GSFC. The archive is equipped with a 32 Tbyte RAID storage system and is scaled to accommodate future growth. The new server configuration became operational in May 2012.

3 Archive Content

As a global data center for the IGS, the CDDIS is responsible for archiving and providing access to GNSS data from the global IGS network as well as the products derived from the analyses of these data in

support of both operational and working group/pilot project activities. Approximately 7 Tbytes of the CDDIS archive are devoted to GNSS data (6.5 Tbytes), products (250 Gbytes), and ancillary information. All data and products are accessible through subdirectories of <ftp://cddis.gsfc.nasa.gov/gnss> (a symbolic link to <ftp://cddis.gsfc.nasa.gov/gps>).

3.1 GNSS Tracking Data

The user community has access to the on-line archive of GNSS data available through the global data center archives of the IGS. Over 50 operational and regional IGS data centers and station operators make data (observation, navigation, and meteorological) available in RINEX format to the CDDIS from selected receivers on a daily, hourly, and sub-hourly basis. The CDDIS also accesses the archives of the other three IGS global data centers, Scripps Institution of Oceanography (SIO) in California, the Institut Géographique National (IGN) in France, and the Korea Astronomy and Space Science Institute (KASI) to retrieve (or receive) data holdings not routinely transmitted to the CDDIS by an operational or regional data center. Table 1 below summarizes the types of GNSS data archived at the CDDIS.

Table 1a: GNSS Data Type Summary.

Data Type	Sample Rate	Data Format	Available On-line
Daily GNSS	30 sec.	RINEX and compact RINEX	Since 1992
Hourly GNSS	30 sec.	Compact RINEX	2+ years
High-rate GNSS	1 sec.	Compact RINEX	Since May 2001
Satellite GPS	10 sec.	Compact RINEX	Since 2002

Table 1b: GNSS Data Archive Summary for 2012.

Data Type	Avg. No. Sites/Day	Avg. Volume/Day	Total Volume/Year	Directory Location	Latency of Majority of Data
Daily GNSS	475	985 Mb	340 Gb	<i>/gnss/data/daily</i>	Within 1 hour
Hourly GNSS	290	310 Mb	105 Gb	<i>/gnss/data/hourly</i>	Within 10 minutes
High-rate GNSS	155	1800 Mb	565 Gb	<i>/gnss/data/highrate</i>	Within 10 minutes
LEO GPS	1	0.5 Mb	200 Mb	<i>/gnss/data/satellite</i>	Within 10 days

Data, in RINEX V2.10 or V2.11 format, from GPS and GPS+GLONASS receivers are archived within the GNSS directory structure */gnss/data*.

The CDDIS archives four major types/formats of GNSS data, all in RINEX format, as described in Table 1a. Daily RINEX data are quality-checked, summarized, and archived to public disk areas in subdirectories by year, day, and file type; the summary and inventory information are also loaded into an on-line database. Nearly 170K daily station days from 525 distinct GNSS receivers were archived at the CDDIS during 2012. A complete list of daily, hourly, and high-rate sites archived in the CDDIS can be found in the yearly summary reports at URL <ftp://cddis.gsfc.nasa.gov/reports/gnss/>.

Within minutes of receipt, the hourly GNSS files are archived to subdirectories by year, day, and hour. These data are retained on-line indefinitely; the daily files delivered at the end of the UTC day contain all data from these hourly files and thus can be used in lieu of the individual hourly files. A total of 320 hourly sites (2.5M files) were archived during 2012.

High-rate (typically 1-second sampling) GNSS data are archived in files containing fifteen minutes of data and in subdirectories by year, day, file type, and hour. Many of these data files are created from real-time streams. Data from 170 high-rate sites (5M files) were also archived in the CDDIS in 2012

The CDDIS generates a global broadcast ephemeris file on an hourly basis. This file is derived from the site-specific ephemeris data files for each day/hour. These files are appended to a single file that contains the orbit information for all GNSS satellites for the day up through that hour. This merged ephemeris data file is then copied to the day's subdirectory within the hourly data file system. Within 1-2 hours after the end of the UTC day, after sufficient station-specific navigation files have been submitted, this concatenation procedure is repeated to create the daily broadcast ephemeris file, using daily site-specific

navigation files as input. The daily file is copied to the corresponding subdirectory under the daily file system. Users can thus download this single, daily (or hourly) file to obtain the unique navigation messages rather than downloading multiple broadcast ephemeris files from the individual stations.

The CDDIS successfully submitted a proposal to the IGS Multi-GNSS Experiment (MGEX) call for proposals for archive and distribution of data and products. During 2012 the CDDIS expanded its data archive and distribution service to include data from participating multi-GNSS receivers, products derived from the analysis of these data, and required metadata for the experiment. These data include newly available signals (e.g., Galileo, QZS, SBAS, and Compass). The CDDIS data ingest procedures were modified to accommodate these new data sets, the majority of which are archived in RINEX V3; some sites have delivered data in RINEX V2 in support of MGEX. The CDDIS developed new software to extract metadata from incoming data files since the software package currently used for summarization and metadata extraction on RINEX V2 data, *teqc*, will not process data in RINEX V3 format. The summary of the MGEX data holdings at the CDDIS are shown in Table 2 below.

Table 2: GNSS MGEX Data Archive Summary for 2012.

Data Type	Avg. No. Sites/Day RINEX 3/2	Avg. Volume/Day	Directory Location
Daily GNSS	45/5	200 Mb	<i>/gnss/data/campaign/mgex/daily</i>
Hourly GNSS	23/6	50 Mb	<i>/gnss/campaign/mgex /data/hourly</i>
High-rate GNSS	17/5	370 Mb	<i>/gnss/campaign/mgex /data/highrate</i>

The CDDIS archived data from space-borne GPS receiver data from selected missions (e.g., SAC-C, CHAMP, Jason-1, ICESat). In 2012 additional ICESat GPS receiver data and attitude files were archived.

3.2 IGS Products

The CDDIS routinely archives IGS operational products (daily, rapid, and ultra-rapid orbits and clocks, and weekly ERP and station positions) as well as products generated by IGS working groups and pilot projects (ionosphere, troposphere, real-time clocks). The CDDIS currently provides on-line access through anonymous ftp or the web to all IGS products generated since the start of the IGS Test Campaign in June 1992 in the file system */gnss/products*; products from GPS+GLONASS products are available through this filesystem. Products derived from GLONASS data only continued to be archived at the CDDIS in a directory structure within the file system */glonass/products*.

The CDDIS also continues to archive combined troposphere estimates in directories by GPS week. Global ionosphere maps of total electron content (TEC) from the IONEX AACs are archived in subdirectories by year and day of year. Real-time clock comparison products have been archived at the CDDIS in support of the IGS Real-Time Pilot Project since 2009. Table 3 below summarizes the GNSS products available through the CDDIS.

Table 3: GNSS Product Summary.

Product Type	Number of ACs/AACs	Volume	Directory
Orbits, clocks, ERP, positions	13+Combinations	830 Mb/week	<i>/gnss/products/WWWW (GPS, GPS+GLONASS)</i> <i>/glonass/products/WWWW (GLONASS only)</i>
Troposphere	Combination	2.5 Mb/day, 860 Mb/year	<i>/gnss/products/troposphere/YYYY</i>
Ionosphere	4+Combination	4 Mb/day, 1.5 Gb/year	<i>/gnss/products/ionex/YYYY</i>
Real-time clocks	Combination	6.0 Mb/week	<i>/gnss/products/rtp/YYYY</i>

3.3 Supporting Information

Daily status files of GNSS data holdings, reflecting timeliness of the data delivered as well as statistics on number of data points, cycle slips, and multipath, continue to be generated by the CDDIS. By accessing these files, the user community can receive a quick look at a day's data availability and quality by viewing a single file. The daily status files are available through the web at URL

<ftp://cddis.gsfc.nasa.gov/reports/gps/status>. The daily status files are also archived in the daily GNSS data directories.

In preparation for the analysis center’s second reprocessing campaign, the CDDIS has developed site-specific reports detailing missing data. Station operators and operational data centers can consult these lists (<ftp://cddis.gsfc.nasa.gov/gnss/data/daily/reports>) and if available, supply missing files to the CDDIS for inclusion in the global data center archives.

Ancillary information to aid in the use of GNSS data and products are also accessible through the CDDIS. Weekly and yearly summaries of IGS tracking data (daily, hourly, and high-rate) archived at the CDDIS are generated on a routine basis. These summaries are accessible through the web at URL <ftp://cddis.gsfc.nasa.gov/reports/gps>. The CDDIS also maintains an archive of and indices to IGS Mail, Report, Station, and other IGS-related messages.

4 System Usage

Figure 1 summarizes the usage of the CDDIS for the retrieval of GNSS data and products in 2012. This figure illustrates the number and volume of GNSS files retrieved by the user community during 2012, categorized by type (daily, hourly, high-rate, products). Over 420 million files (40 Tbytes), excluding robot downloads, were transferred in 2012, with an average of over 35 million files per month. Figure 2 illustrates the profile of users accessing the CDDIS IGS archive during 2012. The majority of CDDIS users are from hosts in Europe and North America.

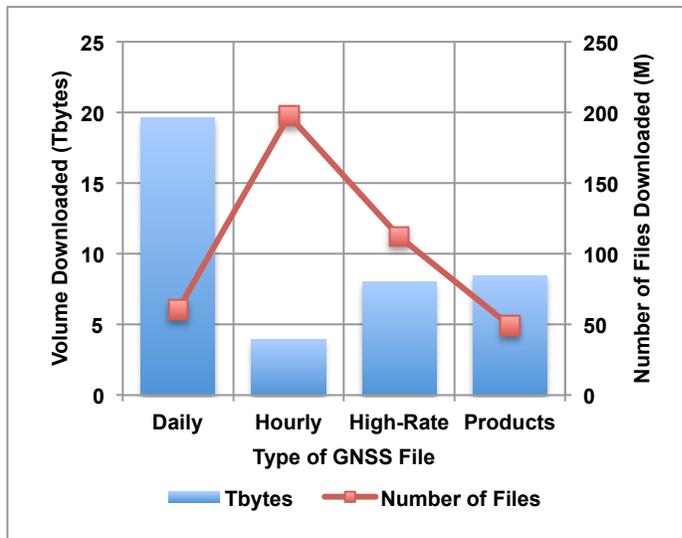


Figure 1: Number and volume of GNSS files transferred from the CDDIS in 2012.

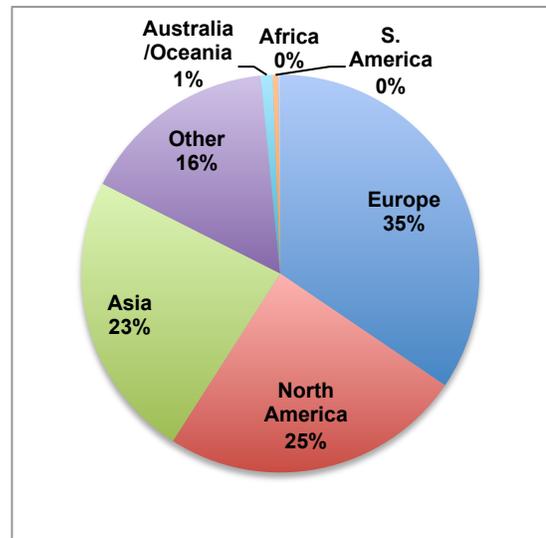


Figure 2: Geographic distribution of IGS users of the CDDIS in 2012.

5 Recent Developments

The CDDIS is cooperating in the development of Geodetic Seamless Archive Centers (GSAC) with colleagues at UNAVCO, SIO, and the University of Nevada at Reno. The activity provides web services to facilitate data discovery within and across participating archives. A prototype implementation of these GSAC web services at the CDDIS has been developed and should be operational in early 2013. In addition, the CDDIS is currently implementing modifications to the metadata extracted from incoming data and product files pushed to its archive. These enhancements have facilitated cross discipline data discovery by providing information about CDDIS archive holdings to other data portals such as Earth

Observing System (EOS) Clearinghouse (ECHO) and future integration into the Global Geodetic Observing System (GGOS) portal.

6 Publications

The CDDIS staff attended several conferences during 2012 and presented papers on or conducted demos of their activities within the IGS, including:

C. Noll, P. Michael, M. Dube, N. Pollack. An Update on the CDDIS, 2012 IGS Analysis Workshop, Olsztyn, Poland, July 2012.

C. Noll, M. Dube, P. Michael, N. Pollack, L. Tyahla. "Improvements to the Crustal Dynamics Data Information System, Abstract IN43B-1515 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 03-07 Dec.

Electronic versions of this poster and other publications can be accessed through the CDDIS on-line documentation page on the web at URL <http://cddis.gsfc.nasa.gov/reports.html>.

7 Future Plans

The CDDIS will continue to support the IGS MGEX. The experiment is an excellent opportunity to prepare the data centers for archive of data in RINEX V3.

The CDDIS is supporting the IGS Real-Time Pilot Project as a data center. During 2012, the CDDIS purchased hardware to implement an NTRIP Castor for the transmission of real-time data streams from stations to users. CDDIS has set up a dedicated server for this task. This service will be tested during early 2013. Possible future activities include capturing the streams for generation of 15-minute high-rate files for archive at the CDDIS.

In 2013, the IGS analysis centers will begin work on the second IGS reprocessing campaign (repro2). The CDDIS will provide support through upload of files from the ACs and online archive of these products ([/gnss/products/WWW/repro2](http://gnss/products/WWW/repro2)).

The CDDIS generates and updates status files, ([/gnss/data/daily/YYYY/DDD/YYDD.status](http://gnss/data/daily/YYYY/DDD/YYDD.status)) that summarize the holdings of daily GNSS data. These files include a list of stations. The archive status files of CDDIS GNSS data holdings reflect timeliness of the data delivered as well as statistics on number of data points, cycle slips, and multipath. The user community can receive a snapshot of data availability and quality by viewing the contents of such a summary file. The CDDIS will enhance this information by providing additional files that list stations not providing data for the particular day.

8 Contact Information

To obtain more information about the CDDIS IGS archive of data and products, contact:

Ms. Carey E. Noll	Phone:	(301) 614-6542
Manager, CDDIS	Fax:	(301) 614-6015
Code 690.1	E-mail:	Carey.Noll@nasa.gov
NASA GSFC	WWW:	http://cddis.gsfc.nasa.gov
Greenbelt, MD 20771		http://cddis.nasa.gov

9 Acknowledgments

The author would like to thank the CDDIS contractor staff, Maurice Dube and Nathan Pollack (Science Systems and Applications, Inc./SSAI), Patrick Michael (Catholic University of America), Lori Tyahla and Lisa Lee (Stinger Ghaffarian Technologies/SGT), and James Roark (ADNET Systems). The recognition and success of the CDDIS in many international programs can be directly attributed to the continued dedicated, consistent, professional, and timely support of its staff.

References

- C. Noll, The Crustal Dynamics Data Information System: A resource to support scientific analysis using space geodesy, *Advances in Space Research*, Volume 45, Issue 12, 15 June 2010, Pages 1421-1440, ISSN 0273-1177, DOI: 10.1016/j.asr.2010.01.018.
- C. Noll, Y. Bock, H. Habrich and A. Moore, “Development of data infrastructure to support scientific analysis for the International GNSS Service”, *Journal of Geodesy*, Feb 2009, pages 309-325, DOI 10.1007/s00190-008-0245-6.