GPS-Met Data Management Plan

1. General Description of Data to be Managed

1.1. Name of the Dataset or data collection project [→ gmd:title].

Ground-based GPS Meteorology

1.2. Keywords that could be used to characterize the data, and vocabulary from which those keywords were obtained (e.g., GCMD, CF Conventions, etc.) [→ gmd:MD_Keywords]
   - **GPS**
   - **Remote sensing**
   - **Water vapor**
   - **Refractivity**
   - **Signal delay**

1.3. Summary description of the data to be generated [→ gmd:abstract].
   Integrated (total column) precipitable water (IPW) in the lower atmosphere are retrieved from measured delays in the Global Positioning System (GPS) radio signals. Because of the characteristics of GPS signal, these measurements can be made anywhere on Earth under virtually all conditions including the presence of clouds and heavy precipitation. The accuracy of the measurement is linked to the accuracy of the atomic clocks in the space and ground segments, and improves (rather than degrade) with time.

1.4. Anticipated temporal coverage of the data [→ gmd:EX_TemporalExtent].
   Continuous record of PW at sites over CONUS maintained at ESRL GSD since 1998. Using GPS observations archived at NGS and other locations, it is possible to reprocess the data and extend the record back to about 1994. The number of sites has grown from 4 in 1994 to more than 400 today (June, 1012).

1.5. Anticipated geographic coverage of the data [→ gmd:EX_Extent]
   Mostly the continental U.S., some sites in Alaska, Hawaii, Puerto Rico, and OCONUS including Western Pacific, Canada, Mexico, South Africa, Taiwan, and other locations.

1.6. What data types will you be creating or capturing? (e.g., digital numeric data, photographs, video, acoustic records, database tables, spreadsheets, paper records, physical samples, etc.)

   Records contain estimated water vapor, signal delay (refractivity), quality information, and observed surface meteorological parameters as a function of time for individual sites available in ASCII, NetCDF, and BUFR format, and complete station metadata including station name, location, time established and removed, and equipment installed.

1.7. How will you capture or create the data? (e.g., satellite, airplane, unmanned aerial system, radar, weather station, moored buoy, research vessel, autonomous underwater vehicle, animal tagging, manual surveys, enforcement activities, etc.)
Surface stations at fixed locations.

1.8. Where will this plan be stored electronically besides in the NOAA DMP Repository?
ESRL data repository in Boulder, CO.

1.9. What volume of data is anticipated to be collected in the Project Time Frame?
7.5 Tb

1.10. Will the data contain Personally Identifiable Information or any information whose distribution may be restricted by law or national security?
No.

2. **Points of Contact** (Give name, title, location, e-mail address, phone number and mailing address, as appropriate.) [→ gmd:CI_ResponsibleParty]

2.1. Who can, or could, represent this data collection project on NOAA's Data Management Integration Team (DMIT)?
Kirk L. Holub, Kirk.L.Holub@noaa.gov
Seth I. Gutman, Seth.I.Gutman@noaa.gov

2.2. Who is the overall point of contact for the data collection?
Kirk L. Holub, Kirk.L.Holub@noaa.gov

2.3. Who is responsible for verifying the quality of the data?
Seth.I.Gutman@noaa.gov

2.4. Who is responsible for answering questions about the data collection?
Kirk L. Holub, Kirk.L.Holub@noaa.gov and/or Seth I. Gutman,
Seth.I.Gutman@noaa.gov

2.5. Who is responsible for data documentation and metadata activities?
Kirk L. Holub, Kirk.L.Holub@noaa.gov and/or Seth I. Gutman,
Seth.I.Gutman@noaa.gov
Site metadata is available online at: http://gpsmet.noaa.gov/cgi-bin/get_site_info.cgi

2.6. Who is responsible for the data storage and data disaster recovery activities?
Kirk L. Holub, Kirk.L.Holub@noaa.gov

2.7. Who is responsible for ensuring adherence to this data management plan, including en-suring that appropriate resources are available to implement the data management plan?
Seth.I.Gutman@noaa.gov

3. Data Stewardship

3.1. What quality control procedures will be employed?
Guidelines for GPS data processing provided by International GNSS Service (http://igscb.jpl.nasa.gov/), implemented through documented, repeatable proce-dures using standard data sources, and compared with independent observations of comparable data and products and operational weather models.

3.2. What is the overall lifecycle of the data from collection or acquisition to making it avail-able to customer?
GPS pseudorange and carrier phase observations are acquired from all sites incorpo-rated into the network as soon as they are available. This involves querying each site at regular (currently 30-min) intervals. Observations are downloaded to the GPS Network Hub in Boulder, Colorado and processed in sub-networks as soon as a obser-vations from a critical number of sites in each sub-network is available. This currently takes less than 15 minutes from the top and bottom of each hour. Data are transmit-ted to NCEP Central Operations (monitored at http://www.nco.ncep.noaa.gov/pmb/nwprod/realtime/), and are available in ASCII, NetCDF and BUFR format via FTP at ftp://aftp.fsl.noaa.gov/gpsmet/ipw/.

4. Data Documentation

4.1. Which metadata repository will be used to document this data collection?
Metadata are stored on the GPSMet web site, its FTP site, its database server, and an administrative server.

4.2. In addition to discovery-level metadata, what additional metadata or other documenta-tion is necessary to fully describe the data and ensure its long-term usefulness? How will that metadata be collected and updated?" Is there a requirement to document this data collection in other metadata repositories?
Meatadata are collected and updated per project operating procedures. There is no requirement that project metadata be stored in other repositories.

4.3. What standards will be used to represent data and metadata elements in this data col-lection. *Note:* The EDMC Data Documentation Procedural Directive calls for the use of ISO 19115 and related standards for data documentation Internal standards are used to represent metadata.
There is no current requirement OAR for this, nor has OAR allocated the resources needed to produced documentation in ISO 19115 format.

5. Data Sharing

    5.1. Will the data be made available to the public? If so, what is the expected date of first availability? Is this a one-time data collection, or an ongoing series of measurements? Will there be a Principal Investigator hold or other delay between data collection and publication, and if so for how long? [Note: the Data Sharing for NOAA Grants Procedural Directive provides useful guidance for sharing data in a timely manner.]

- **All current and historical GPS observations used in this project are archived at and available from the following public GPS data archives:**
  NOAA CORS: http://www.ngs.noaa.gov/CORS/
  UCSD Scripps SOPAC: http://sopac.ucsd.edu/
  UNAVCO: http://www.unavco.org/

- **GPS-Met data products are available in near real-time (usually within 30 minutes of the time-stamp) at: http://gpsmet.noaa.gov/cgi-bin/gnuplots/rti.cgi**

- **There are no intentional holds or delays imposed on the availability of these data or products by NOAA or its affiliates, Scripps and UNAVCO.**

    5.2. If the data are not to be made available to the public, explain why and under what authority distribution may be restricted. NOAA Administrative Order 212-15, "Management of Environmental Data and Information" (2010) states that Environmental data will be visible, accessible and independently understandable to users, except where limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements.
N/A

    5.3. Will users be subject to any access conditions or restrictions, such as submission of non-disclosure statements, special authorization, or acceptance of a licensing agreement?
No

    5.4. What data access protocols will be used to enable data sharing? The use of open-standard, interoperable, non-proprietary web services is recommended (for example, OPeNDAP, or Open Geospatial Consortium (OGC) web services).
Open standards are used for data sharing. The data may be accessed via web pages, ftp sites, and web services.

    5.5 In what catalogs will these services or data be made registered to enable discovery by users and other Catalogs
There are currently no plans or funds available to add GPSMet data to any discovery services.

6. Initial Data Storage and Protection

6.1. Where and how will the data be stored initially (i.e., prior to being sent to a long-term archive facility)?
Data is stored at ESRL.

6.2. How will the data be protected from accidental or malicious modification or deletion? Discuss data back-up, disaster recovery/contingency planning, and off-site storage relevant to the data collection.
- **The data is protected using standard NOAA security practices**
- **The data is backed up on a regular schedule**
- **Copies of the data are held off site per disaster recovery requirements.**

6.3. If there will be limitations to data access, how will these data be protected from unauthorized access? How will access permissions be managed? What process is to be followed in the event of unauthorized access?
N/A

7. Long-Term Archiving and Preservation

Note: NOAA's [Procedure for Scientific Records Appraisal and Archive Approval](#) describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.

7.1. In what NOAA Data Center (NODC, NCDC, NGDC) will the data be archived and preserved? Have you begun discussions with that Data Center regarding your intended submission?
NCDC. Informal discussions with John Bates at NCDC were conducted, and formal discussions will be initiated when sufficient program funds to accomplish this have been identified.

7.2. If you have not identified a NOAA Data Center, what is your long-term strategy for maintaining, curating, and archiving the data?
Continue to maintain data at ESRL and engage NCDC when sufficient funds have been identified.

7.3. How will the costs of long-term data archiving be provided and maintained?
This will be determined as part of the planned transition of GPS Met from research to operations.

7.4. What transformations or procedures will be necessary to prepare data for preservation or sharing? (e.g., quality control, format conversion, anonymization of personally-identifiable information, etc.). What related information will be submitted to the archive to enable future use and understanding of the data [e.g., metadata, references,

reports, research papers, algorithms, audio or video codecs, special character sets or fonts, etc.]) .

- **No additional requirements for preservation have been identified.**
- **Data for archiving will be submitted to NCDC with all existing documentation including metadata**

7.5. Identify the Record Schedule applicable to these data and provide the retention time for these data.
To be specified by NCDC.