Data Management Plan Boulder Atmospheric Observatory

1. **General Description of Data to be Managed**
    1. Name of the Dataset or data collection project [→ gmd:title].
        Boulder Atmospheric Observatory (BAO)

    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]
        BAO, research tower, wind profiles, temperature profiles, climate monitoring, BSRN, Tall Tower Network, planetary boundary layer

    3. Summary description of the data to be generated [→ gmd:abstract].
        The Boulder Atmospheric Observatory is a 300m tall instrumented meteorological tower located on the eastern plains of Colorado (http://www.esrl.noaa.gov/psd/technology/bao/). The BAO provides continuous boundary layer measurements of wind, temperature and relatively humidity at 3 different levels, surface pressure and precipitaion, sodar backscatter, ceilometer backscatter and cloud base height (http://www.esrl.noaa.gov/psd/technology/bao/browser/).

System or Sensor Performance Attribute Values							
Environmental Parameters this System Measures	GCMD Variable measured	Geographic Coverage (e.g. Global, hemispheric, coastal, EEZ, ship or aircraft track, etc.)	Sensor/Instrument Name	Horizontal Resolution (e.g. km, meters)	Measurement Accuracy (e.g. mm, degrees C, etc.)	Update Frequency (e.g. days, hours, minutes, seconds)	Vertical Resolution (e.g. km, meters, hPA ...)
Air Temperature	Air Temperature: 10,100,300m	Point Source	Vaisala HMP45		.5 K	1 min	na
Atmospheric Pressure	Atmospheric Pressure: Surface	Point Source	Vaisala 110	na	.1 hPa	1 min	1 m
Precipitation	Precipitation Rate sfc	Point Source	Texas Inst. TR-525 heated tipping bucket		1 mm/hr	1 min	na
Relative Humidity	Water Vapor: 10,100,300m	Point Source	Vaisala HMP45		2 %	1 min	na
Wind Direction	Winds: Direction 10,100,300m	Point Source	RM Young 05103 Prop-vane	na	3 deg	1 min	na
Wind Speed	Winds: Speed 10,100,300m	Point Source	RM Young 05103 Prop-vane	na	.3 m/sec	1 min	na

The BAO is currently part of the CO2 Tall Tower Network (http://www.esrl.noaa.gov/gmd/ccgg/towers/) and World Climate Research Program (WCRP) Baseline Surface Radiation Network (BSRN) (http://cmdl1.cmdl.noaa.gov:8000/www/all/rad/index.html#BAO8).

4.   Anticipated temporal coverage of the data [→ gmd:EX_TemporalExtent].
     June 14, 2007- continuous operations into the foreseeable future

5.   Anticipated geographic coverage of the data [→ gmd:EX_Extent]
     Latitude: 40 03 00.10028(N)

     Longitude: 105 00 13.80781(W)

     Elevation: 1584 m

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.)
     digital numeric data, data plots, photographs from web camera

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.)

Data from meteorological instruments are collected via data loggers located on the tower

8. Where will this plan be stored electronically besides in the NOAA DMP Repository?
Google Docs

9. What volume of data is anticipated to be collected in the Project Time Frame?
MB/Day

Tower	0.27	0.23	
Ceilo	23.14	1.22	
Sodar	0.07	0.07	
Ozone	0.24	0.51	
Lights	0.02	0.01	
CU	117.60		
Total	141.34	2.04	143.3MB/ day or 4.301 GB/month

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

2. **Points of Contact** (Give name, title, location, e-mail address, phone number and mailing address, as appropriate.) [→ gmd:CI_ResponsibleParty]
   1. Who can, or could, represent this data collection project on NOAA's Data Management Integration Team (DMIT)? Current members of DMIT are listed at https://geo-ide.noaa.gov/wiki/index.php?title=DMIT_Membership.
   Steve Hankin from PMEL and/or Chris MacDermaid from ESRL/GSD

   2. Who is the overall point of contact for the data collection?
   Daniel Wolfe daniel.wolfe@noaa.gov 303-497-6204

   3. Who is responsible for verifying the quality of the data?
   Daniel Wolfe daniel.wolfe@noaa.gov 303-497-6204

   4. Who is responsible for answering questions about the data collection?
   David Welsh david.c.welsh@noaa.gov 303-497-6510

   5. Who is responsible for data documentation and metadata activities?
   Daniel Wolfe daniel.wolfe@noaa.gov 303-497-6204

   6. Who is responsible for the data storage and data disaster recovery activities?
   David Welsh david.c.welsh@noaa.gov 303-497-6510

   7. Who is responsible for ensuring adherence to this data management plan, including ensuring that appropriate resources are available to implement the data management plan?
   TBD

3. **Data Stewardship**
   1. What quality control procedures will be employed?
   Visual inspection of the data with bad data flagged

   2. What is the overall lifecycle of the data from collection or acquisition to making it available to customer?
   Non-QC'd data are available in near real-time

QC'd Data are available 1 day after being collected

4. **Data Documentation**
   1. Which metadata repository will be used to document this data collection?
      NNMR http://www.ngdc.noaa.gov/metadata/

      GCMD http://gcmd.nasa.gov/

      In addition to discovery-level metadata, what additional metadata or other documentation 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?

      No requirement to document this data in another collect. No additional necessary documentation is anticipated at this time.

   2. What standards will be used to represent data and metadata elements in this data collection. *Note:* The EDMC Data Documentation Procedural Directive calls for the use of ISO 19115 and related standards for data documentation.
      ISO 19115

5. **Data Sharing**
   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.]
      Data are available now to the public

   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.
      NA

   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?
      We request we are acknowledged by anyone using the data.

      We also reserve the right to **deny access to any individual or organization that we determine is abusing this service**. Examples of abuse include automated transfers resulting in excessive data requests (because it hinders others from accessing the service) and attempting to gain access to documents and host machines not intended for public use.

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).
    FTP

5. In what catalogs will these services or data be made registered to enable discovery by users and other Catalogs?
    NNMR and GCMD

Initial Data Storage and Protection

6. Where and how will the data be stored initially (i.e., prior to being sent to a long-term archive facility)?
    The data are stored on a iSCSI RAID 6 RAID array. The data storage is located in a secure server room that is temperature controlled. The facility is the David Skaggs Research Center in Boulder, Colorado.

7. 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 are protected by using RAID (redundant array of independant disks). The disk array can withstand two drive failures and still operate at reduced performance. A hotspare drive is ready to replace any drives that fail. The system is also backed up weekly with differential backups and a full backup monthly. A set of tapes is taken offsite and then that set is rotated every six months. The backup system is tested monthly to verify that data is being stored properly and it can be recovered.

8. 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?
    -PSD protects systems from unauthorized access via a combination of network-based and system-based controls. Access to all PSD systems is restricted via a single network firewall. Firewall access to/from a given system and service is granted based on business need, with a 'default deny' policy. On each system, data is protected from unauthorized access via Linux file system permissions, local iptables firewall policies, and TCP wrappers.

    -Access permissions are managed via Active Directory, and Linux systems are connected to the Active Directory with LDAP and Kerberos. Users are placed into Active Directory groups based on collaborative needs, and file system permissions are managed on both a per-user and per-group basis depending on specific access needs.

    -File system integrity checking software is implemented on critical systems in the PSD infrastructure to detect unauthorized access. Intrusion Detection Systems (IDS) are also in place to detect unauthorized access and notify relevant IT staff.

6. **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.

    1. In what NOAA Data Center (NODC, NCDC, NGDC) will the data be archived and preserved?
        No, but if we do it would be with NCDC

2. Have you begun discussions with that Data Center regarding your intended submission?
   No

3. If you have not identified a NOAA Data Center, what is your long-term strategy for maintaining, curating, and archiving the data?

4. How will the costs of long-term data archiving be provided and maintained?
   PSD IT funding, as directed by the director of PSD.

5. 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.]) .

Identify the Record Schedule applicable to these data and provide the retention time for these data.

   Data are to be retained indefinitely