17 March 2017

The Plate Boundary Observatory (PBO) Data Management System Critical Design Review (CDR) document was one of the primary information sources describing the original design and implementation of PBO data collection, data flow and data products. The CDR in its original form was last updated on 3 April 2006, as version 1.2.

The CDR v1.2 continues to be an important reference today. However, many aspects of PBO data operations and products have changed since 2006. As 2017, updated documentation describing PBO data and products exists as a set of independent documents and web pages distributed throughout www.unavco.org rather than in a single document format like the original CDR. For example, the information in Chapter 5 and Appendix D of the CDR v1.2, describing GPS data and analysis, has been significantly revised and expanded to the point that an entire web section is dedicated to this topic: http://www.unavco.org/data/gps-gnss/derived-products/derived-products.html. As to be expected, significant changes to the membership and structure of oversight have occurred as well during the past decade.

Users of PBO data are therefore advised to consider information presented on the UNAVCO website as the most current information available, and to consider information in the CDR version 1.2 as historical reference. Questions regarding data products can also be addressed to Dr. David Phillips, Project Manager, Geodetic Data Services (phillips@unavco.org).

Plate Boundary Observatory Data Management System Critical Design Review Version 1.2

Greg Anderson
PBO Data Products Manager
anderson@unavco.org

Kathleen Hodgkinson
PBO Strainmeter Data Analyst
hodgkinson@unavco.org

Thomas Herring
PBO Analysis Center Coordinator
tah@mit.edu

Duncan Carr Agnew
Director, PBO Laser Strainmeter Analysis Center
dagnew@ucsd.edu

April 3, 2006

Contents

List of Tables

List of Figures

Acknowledgments

The PBO Data Management System could not be developed without significant efforts on the parts of many people; naming all of them is not possible, but all their contributions are greatly appreciated.

We particularly want to thank the PBO staff, and especially Steve Borenstein, Eliot Lee, Mike Hasting, James Matykiewicz, David Mencin, Bob Mueller, Erik Persson, Steve Smith, Damian Torrez, and Jim Wright. We also want to thank Duncan Agnew, Tim Ahern, John Beavan, Fran Boler, Lind Gee, Mick Gladwin, Tim Melbourne, Mark Murray, Doug Neuhauser, Chris Stolte, Frank Wyatt, the PBO Data Analysis Working Group, PBO Strainmeter Working Group, PBO Data Products Working Group, the EarthScope Operations Group, and the EarthScope Data Working Group (chaired by Christel Hennet) for their help in developing the plans laid out here. Finally, we want to thank those community members who participated in the review process for the PBO Data Management Plan, the predecessor to this document.

Part I

Introduction and Background

Chapter 1

Executive Summary

The Plate Boundary Observatory (PBO), part of the EarthScope project, will study active plate boundary deformation across the western United States using 875 continuous GPS (CGPS) stations, 103 borehole strainmeter (BSM) stations, and five laser strainmeters (LSMs), all installed over the next five years. In addition, there will be survey-mode GPS (SGPS) data collected using a 100-receiver PBO equipment pool, and starting in Year 2, 209 continuous GPS stations from current western United States geophysical networks are being absorbed i into the PBO network.

This network will give solid Earth scientists an unprecedented four-dimensional view of active plate boundary deformation across the western United States and Alaska and help the natural hazards community better understand earthquake and volcano behavior. As this network will also be the most accurate spatial reference system available in United States history, it will be actively used by surveyors, engineers, and others for land control surveys, infrastructure monitoring, and many other tasks. This document describes our plan for management of the considerable data and derived data products this large, broadly-useful network will generate.

PBO generates three levels of data and derived data products, all of which is governed by the PBO open data policy. Level 0 data are raw data collected at each instrument as well as site metadata. Continuous GPS stations collect data at a rate of 15 seconds/sample (15-sec) and 5 sample/sec (5-sps); all 15-sec data are routinely downloaded, while 5-sps data will only be downloaded following a large earthquake or similar event. As PBO evolves, a fraction of the network may be configured to also stream 1-sps continuous GPS data in real time. Borehole strainmeter stations collect 100-sps seismic data; 20-sps, 1-sps, and 600-sec strain data; and auxiliary channels at 1-sps and 300-sec. Laser strainmeter stations collect data at 1-sps and 300-sec.

GPS Level 1 data products are produced when Level 0 GPS data are quality-checked, but otherwise are indistinguishable from Level 0 data. GPS Level 2a data products are produced by two GPS Analysis Centers, one formerly at the University of California, Berkeley, and now at the New Mexico Institute of Mining and Technology, and the other at Central Washington University. These products include GPS station position estimates in SINEX format and input and auxiliary output files from GPS processing codes. “Rapid” Level 2a products are produced within 18–24 hours of the end of a given UTC day, with “final” products produced approximately two weeks later, and “rerun” solutions generated on average once per year. GPS Level 2 products, produced by the GPS Analysis Center Coordinator, include combined GPS station position estimates in SINEX or similar format; simplified position time series in

ASCII format; periodic estimates of long-term GPS station velocity in SINEX for similar format; regular estimates of simplified velocity fields, time series noise properties, and periodic time series components, all in ASCII format; GPS baseline time series generated automatically upon user request from the position SINEX files; and coseismic offsets estimated following significant earthquakes in or near the PBO network.

Borehole and laser strainmeter Level 1 data products are identical to Level 0 data products, except for scaling to natural strain units. They also include automated conversion to areal, shear, and linear strain, which will be available almost immediately but will have no human intervention. Level 2 derived products include fully corrected and scaled tensor and linear strain time series and ancillary series, with Level 2a products available with latency of less than 14 days. Level 2b products involve reexamining the entire dataset from the beginning to find long-term problems, and are produced every four months. Level 2 borehole strainmeter products are generated by the Borehole Strainmeter Analysis Center in Socorro, New Mexico. Level 2 laser strainmeter products are generated by the Laser Strainmeter Analysis Center in San Diego.

The Level 0 data are collected at remote stations and sent through a variety of telecommunications paths to a unified data center in Boulder, Colorado. All metadata related to PBO stations and data, including those gathered during the reconnaissance, installation, and operational stages as well as those derived from the generation of higher-level products, are collected in the PBO Operational Database (POD) at the Boulder Data Center. Once in Boulder, the data will be quality checked using the POD, and then distributed to the PBO Archives and Analysis Centers. There will also be a second data center in Socorro, New Mexico, which will be kept in sync with the Boulder Data Center on a regular basis, and will be capable of taking over for the Boulder facility in the event of system failure in Boulder. This will help ensure stable, reliable data flow despite problems at any one data center.

All PBO GPS data products are be archived in two GPS Archives operated by the UNAVCO Facility with support from the IRIS Data Management Center. One Archive is located at the UNAVCO Facility in Boulder, while the other is located at the IRIS DMC. All PBO Strainmeter data products are archived at two Strain Archives, one at the Northern California Earthquake Data Center at UC Berkeley and the other at the IRIS DMC. These Archives store all PBO data products and will distribute them to the community through their current distribution systems, the PBO web pages, and ultimately the EarthScope Integrated Data Access System (still to be developed). Having two geographically distributed archives for each kind of PBO data will greatly reduce the risk of loss of archive function due to catastrophic failure at any one center, and will help guarantee easy, free access to all PBO data products.

Chapter 2

Policies

2.1 Document Scope

This is the primary document defining PBO data management policies, strategies, and PBO data products. Where this document summarizes any matters (e.g., station installations, monumentation, etc.) that are officially and fully documented in another primary PBO document, such as the Project Execution Plan, other management plans, or a contract statement of work, that other document takes precedence.

2.2 PBO Data Policy

The following policy will govern all PBO data and data products, which are defined as any data and/or data product generated with PBO funding:

- Data from all PBO stations will be available online as soon as they can be moved from the site to the PBO Archives.
- All raw data and metadata from PBO continuously operating stations and survey-mode observations made by PBO personnel or with core EarthScope funding, and Level 1 and 2 data products derived from those data, will be made freely available through the PBO Data Portal as rapidly as possible, except where release of such data could harm PBO operations; examples of data that will not be released include receiver/datalogger IP addresses, lockbox key numbers, and landowner name and contact information.
- Raw data collected with independent PI-driven research funding but using PBO survey-mode GPS equipment, and metadata related to those data, will be archived in the PBO GPS Archives and PBO Operational Database as rapidly as possible after the end of a given survey, and within six months of data collection at the latest. PIs may request these raw data and metadata be withheld from public access for a period of no more than two years after initial data collection; otherwise, these data will be publicly accessible immediately.

If a project PI desires the PBO Analysis Centers to analyze raw data he/she collected using the PBO survey-mode receiver pool, he/she must supply to PBO sufficient funding to cover this analysis; analogous conditions hold for analysis by the PBO Analysis Center Coordinator. If Analysis Center (and/or Analysis Center Coordinator) processing is requested, and sufficient funding is made available by the PI, the Analysis Centers (and/or Analysis Center Coordinator) will process those data to create Level 2a (and/or 2b) data products. These products will be sent to the project PI and to the GPS Archives; the results will be made freely available immediately after their generation.

- All software developed by PBO or with PBO funding will be governed by the PBO Software License, which will be based on existing open-source licenses such as the Gnu General Public License. The primary features of the license are that all source code will be freely available to the United States academic and non-profit communities, and that no PBO software may be used for commercial purposes by any entity (person, group, organization, or institution) receiving PBO funding.
- Any PBO-funded entity that modifies software licensed from any other entity must submit the modified source code to the license holder for further distribution to the community under the external entity's license.
- No entity that receives funding to generate, archive, or analyze PBO data, metadata, derived data products, or software may derive commercial profit from his/her/their/its PBO-funded work.
- Users of PBO data must acknowledge EarthScope and the NSF as the source of those data. We request users include text similar to the following in reports, papers, and other products using PBO data: "We acknowledge EarthScope and its sponsor, the National Science Foundation, for providing data [or derived data products, as appropriate] used in this study."

2.3 Conflicts of Interest Policy

PBO will impose the following main restrictions on the PBO Archives, Analysis Centers, Analysis Center Coordinator, and advisory group membership:

1. No PBO GPS Analysis Center may include any member who receives, is supervised by someone who receives, or supervises someone who receives, funding to act as a PBO GPS Archive or the PBO Analysis Center Coordinator. Similarly, no PBO GPS Archive may include any member who receives, is supervised by someone who receives, or supervises someone who receives, funding to act as a PBO GPS Analysis Center or the PBO Analysis Center Coordinator.
2. Entities (individuals, groups, organizations, or institutions) that receive funding to act as PBO GPS Archives may bid for a position as a PBO GPS Analysis Center or the Analysis Center Coordinator. However, if such an entity's bid is successful, the entity will be required to give up its role and all funding as a PBO GPS Archive when the entity takes on its new role as a PBO GPS Analysis Center or the Analysis Center Coordinator.
3. No individual that receives, is supervised by someone who receives, or supervises someone who receives, funding to act as a PBO GPS Archive or PBO GPS Analysis Center may receive funding to act as the PBO Analysis Center Coordinator. Any individual may bid to become the PBO

Analysis Center Coordinator; however, if the successful bidder falls into any category listed in the previous sentence, he or she will be required to resign from any and all Archiving or Analysis Center activities and forfeit all Archiving or Analysis Center funding when they begin their new role as PBO Analysis Center Coordinator.

4. No PBO Strainmeter Analysis Center may include any member who receives, is supervised by someone who receives, or supervises someone who receives, funding to act as a PBO Strainmeter Archive. Similarly, no PBO Strainmeter Archive may include any member who receives, is supervised by someone who receives, or supervises someone who receives, funding to act as a PBO Strainmeter Analysis Center.

5. No individual who receives, or is supervised by anyone who receives, or supervises any individual who receives funding to act as a PBO Archive, Analysis Center, or the Analysis Center Coordinator may serve in a voting capacity on the PBO advisory working group overseeing that aspect of PBO.

6. Any member of any PBO advisory working group who wishes to submit a proposal in response to any PBO RFP or RFQ must declare his or her intention to do so in writing to the PBO Director, and will be required to resign his or her voting position if he or she is the successful bidder.

7. Prior to reviewing PBO RFPs, RFQs, and Statements of Work, members of the UNAVCO Board of Directors and Standing Committee will be required to sign a non-disclosure agreement (NDA) governing the particular document. Those Board or Standing Committee members who may be interested in bidding on a given RFP or RFQ will not have to sign the NDA and will be excluded from any discussion of that RFP or RFQ. Board or Standing Committee members who do sign an NDA and participate in the review of the given RFP or RFQ will be prohibited from discussing any aspect of that RFP/RFQ with any member who has not signed the NDA, and will be prohibited from bidding on the final RFP/RFQ.

8. Exceptions to this policy may be granted by the PBO Director, with concurrence of the UNAVCO President and Board of Directors, when necessary to meet PBO science, budget, schedule, and/or reporting requirements.

Chapter 3

Oversight

3.1 Standing Committee

The PBO Standing Committee (PBOSC) is charged with the role of advising the UNAVCO, Inc. Board of Directors in overseeing all functions and all aspects of PBO management. The members of the Standing Committee are appointed by, and report to, the UNAVCO Board. The Standing Committee represents the PBO scientific data user community to ensure that the scientific goals of the PBO are met, to the extent possible within fiscal and other practical limits. The Standing Committee meets regularly to review in detail the progress on all aspects of PBO operations, facilities, and contracts. In particular, the Standing Committee reviewed the PBO Data Management Plan and the RFPs for the Analysis Centers and Analysis Center Coordinator. The Standing Committee will identify issues of concern and recommend actions to the UNAVCO Board, which will in turn advise the UNAVCO, Inc. President and PBO Director. The current Standing Committee members, as of April 3, 2006, are:

- Paul Segall (Chair), Stanford University
- John Beavan, Institute of Geological and Nuclear Sciences
- Greg Beroza, Stanford University
- Brad Hager, Massachusetts Institute of Technology
- William Holt, State University of New York, Stony Brook
- Susan Owen, University of Southern California
- Evelyn Roeloffs, United States Geological Survey, Vancouver
- Mark Simons, California Institute of Technology

A current list of the Standing Committee membership will be maintained on the PBO web site.

3.2 Data Products Working Group

The PBO Data Products Working Group is charged with helping the PBO Data Products Manager (DPM) and PBO Director in defining high-level requirements for collection, archiving, analysis of data from the PBO GPS and strainmeter installations, and derived data products. Working Group members represent the EarthScope science user community and are appointed at the sole discretion of the PBO Director. The Working Group will meet at least annually to evaluate PBO data and data products and to review, and if necessary suggest revisions to, the PBO data management plan. After each meeting, the Working Group Chair will submit a written evaluation report to the PBO Data Products Manager and Director, which they will use in evaluating progress and initiating necessary changes to PBO data management systems. The current Working Group members, as of April 3, 2006, are:

- Jeff Freymueller (Chair), University of Alaska, Fairbanks
- Duncan Agnew, University of California, San Diego
- Rick Bennett, University of Arizona
- Fran Boler (ex-officio non-voting), UNAVCO
- James Davis (ex-officio non-voting), Harvard-Smithsonian Center for Astrophysics
- Elizabeth Hearn, University of British Columbia
- Thomas Herring (ex-officio non-voting), Massachusetts Institute of Technology
- Nancy King, United States Geological Survey, Pasadena
- John Langbein, United States Geological Survey, Menlo Park

A current list of the Working Group membership will be maintained on the PBO web site.

3.3 GPS Data Analysis Working Group

The PBO GPS Data Analysis Working Group (DAWG) will assist the PBO Data Products Manager and Analysis Center Coordinator in developing detailed standards and methods for the analysis of PBO GPS data and production of PBO GPS data products. The DAWG represents the scientific community in ensuring that PBO GPS data products are generated using community-accepted, state-of-the-art methods and standards. The DAWG will review, at least annually, the detailed GPS analysis plan developed by the Analysis Center Coordinator and the Analysis Center Directors (see Appendix D on page 56), making recommendations for changes as needed. The DAWG Members and Chair will be appointed by and will report to the PBO Director.

As of April 3, 2006, the members of the GPS Data Analysis Working Group are

- James Davis (Chair), Harvard-Smithsonian Center for Astrophysics
- Eric Calais, Purdue University

- Danan Dong, Jet Propulsion Laboratory
- Tim Dixon, University of Miami
- Jeff Freymueller (ex-officio non-voting), University of Alaska, Fairbanks
- Tom Herring (ex-officio non-voting), Massachusetts Institute of Technology
- Tim Melbourne (ex-officio non-voting), Central Washington University
- Mark Murray (ex-officio non-voting), New Mexico Institute of Mining and Technology
- Giovanni Sella, National Geodetic Survey

A current list of the Working Group membership will be maintained on the PBO web site.

3.4 Strainmeter Working Group

The PBO Strainmeter Working Group, among other roles, advises PBO on strainmeter analysis strategies and standards. Members of the Strainmeter Working Group include the Strainmeter Data Analyst, the Director of the Laser Strainmeter Analysis Center, and experts on strainmeter data analysis drawn from the EarthScope community. The Strainmeter Working Group will review, at least annually, the detailed borehole strainmeter analysis plan developed by the Strainmeter Data Analyst, making recommendations for changes as needed.

As of April 3, 2006, the Strainmeter Working Group members are:

- Duncan Agnew, University of California, San Diego
- Roger Bilham, University of Colorado, Boulder
- John Beavan, Institute of Geological and Nuclear Sciences, New Zealand
- Mick Gladwin, GTSM Technologies, Australia
- Malcolm Johnston, United States Geological Survey, Menlo Park
- Alan Linde, Carnegie Institution of Washington
- Mike Lisowski, United States Geological Survey, Vancouver
- I. Selwyn Sacks, Carnegie Institution of Washington
- Frank Wyatt, University of California, San Diego

A current list of the Strainmeter Working Group membership will be maintained on the PBO web site.

Part II

PBO Stations and Data Products

Chapter 4

Naming Conventions and Metadata

4.1 Site/Data Naming Conventions

This section outlines PBO's conventions for naming stations and data. Appendix C on page 52 gives much more detail on these new conventions.

PBO has a specific terminology describing a hierarchy from sites to data channels:

- **Site**
  The property on which one or more stations are installed, delineated by the legal definition in each site's permit.
- **Station**
  A collection of sensors and a reference point. For example, a PBO GPS station consists of the GPS monument and antenna, and optional met sensors, and the reference point for a PBO continuous GPS station is the tip of the center-support screw on the D3 adaptor atop the GPS monument. A given site may have more than one station installed.
- **Sensor**
  A particular instrument, such as a GPS antenna, each component of a borehole seismometer, or a fiber anchor for a laser strainmeter. A given station may contain multiple sensors.
- **Channel**
  A channel is a specific data stream from a specific sensor, recorded at a specific sample rate; a given sensor may have more than one channel. For example, a GPS station may have channels at 15-sec, 1-sps, and 5-sps, or a borehole strainmeter's north-south sensor could have channels sampled at 20-sps, 10-sps, 1-sps, and so on.

PBO stations have a 4-character identifier as has been common practice, but also have two longer identifiers: a 16-character "short name" based on a common geographic reference, and a 12-character geocode based on the station's precise position:

- **Station Short Name**
  The primary station identifier for PBO will be a 16-character name, based on some nearby landmark

as is current practice. The short name will have the format `ccccccccccrryyyy`, where `cccccccccc` is a 10 alphanumeric character station name string, `yyyy` is the 4-digit year in which the station was installed, and `rr` is a 2-character region code. An example is `Marshall__CO2004` for the station installed at Marshall Field, in Boulder, Colorado in 2004.

- **Station Location Geocode**
  The secondary station identifier for PBO will be a geocode that converts the position of the station's reference point, when installed, into a compact representation that is suitable for filenames and the like. We will use the GHAM geocodes proposed by Duncan Agnew [*Agnew*, 2005], which can specify locations to arbitrary accuracy and are structured such that stations located near each other will appear near each other in a lexicographically-sorted list of geocodes (see Appendix B on page 50 for more details on how the GHAM system works). PBO Station Location Geocodes are 12 characters long; an example geocode for a site at 32.867772°N, -117.252331°E would be `E4I8U3W2V7I3`.

Files containing PBO data and derived data products will be named using the station short name and the station's four-character identifier, to allow users to find data using either standard.

4.2 Metadata

Metadata about PBO sites, stations, sensors, data, and derived data products are collected and stored in the PBO Operational Database (POD). Metadata contained in the POD includes such classes as:

- Standard station information: location, point of contact, etc.
- Site equipment information: equipment type, UNAVCO ID, serial number, telecommunications paths, etc.
- Reconnaissance, siting, and permitting: pictures, pointers to paper field reconnaissance logs and site permits, scanned images of site permits, etc.
- Installation and construction data: Dates installed, who by, geology type, materials used, etc.
- Network state of health: history of data received, battery voltages, temperature, etc.
- Derived parameters from data products: long-term GPS station velocity; noise parameters; seasonal signals; offsets and outliers; borehole and laser strainmeter tidal admittance and scale factors; etc.

Table 4.1 gives some more details on the types of metadata stored in the POD.

Metadata related to reconnaissance, siting, permitting, and installation are submitted manually using the PBO Site Reconnaissance Report (SRR) and Site Installation Report (SIR), a set of tools based on the SCIGN Site Evaluation and Construction Reports and the current UNAVCO Facility Permanent Station Database interface. Metadata related to network state of health are gathered automatically on a subdaily basis. PBO Analysis Centers will also upload various metadata derived from data analysis and product generation using a similar system.

In order to facilitate metadata queries and help prevent metadata loss, the PBO Archives act as backup repositories for public metadata from PBO. The POD contains the master copy of PBO metadata, and

Table 4.1: Sample Metadata Classes[1]

Metadata Type	Examples	Source	Update frequency
Station	Station Name	SIR[2]	Never
	Station Location Geocode	SIR	Never
	Point of Contact	SIR	As needed
	Descriptive Location	SIR	As needed
	ITRF Location	Data Processing	Monthly
	Installation date	SIR	Never
	Geology type	SRR[3]/SIR	As needed
Equipment	Receiver/antenna type	SIR	As needed
	Rcvr/antenna serial #	SIR	As needed
	Date installed	SIR	As needed
	Antenna height	SIR	As needed
	Station IP address	SIR	As needed
	Telecommunication path	SIR	As needed
Recon , Siting , Permitting	Recon log index pointer	PBO staff	Never
	Site permit pointer	PBO staff	Never
	Scanned recon logs	PBO staff	Never
	Scanned site permits	PBO staff	Never
	Site pictures	PBO staff	As needed
Network SOH[4]	History of data rec'd	PBO staff	Regularly
	Battery voltages	PBO staff	Regularly
Parameters from Data Processing	GPS station velocity	Data Processing	Monthly
	Tidal admittance	Data Processing	Monthly
	Time series noise	Data Processing	Monthly
	Seasonal signals	Data Processing	Monthly
	Offsets and outliers	Data Processing	Monthly
	BSM/LSM scale factors	Data Processing	As needed

[1] This is not an exhaustive list, but merely intended to give examples of the kinds of metadata to be collected in the POD.
[2] PBO Site Installation Report, derived from the SCIGN Site Construction Report.
[3] PBO Site Recon Report, derived from the SCIGN Site Evaluation Report.
[4] SOH: State of Health

each of the PBO Archives has a secondary copy derived from, and kept in sync with, the master copy. PBO works with the PBO Archives to develop and implement appropriate methods for distribution, synchronization, and storage of PBO metadata.

It is important to note that not all metadata related to PBO sites are public. Those metadata, including receiver IP addresses and lockbox key numbers, that are critical to PBO operational security are strictly limited to PBO staff only. As there is no need for anyone outside PBO operational staff to have access to these metadata, this restriction should have no impact on PBO end users.

EarthScope community members will access public PBO metadata using a read-only interface to the POD, which will ultimately be part of the PBO web site and eventually the EarthScope Integrated Data Access System. This interface will also allow users to sign up for automated e-mail notification of changes to metadata stored in the POD. Also, PBO Archives may make public metadata available on a read-only basis through their own access methods, provided the Archives prominently identify the metadata as generated by PBO.

Chapter 5

PBO GPS Stations and Data

5.1 Continuous GPS Stations: Primary Data Flow and Products

5.1.1 Receiver Configuration

PBO continuous GPS installations consist of a geodetic GPS receiver, choke-ring antenna, auxiliary power and communications equipment, and a SCIGN-type deep-braced monument or shallow-braced monument; some stations, particularly in Alaska, may require other monument types. PBO continuous GPS receivers have 1 GB of internal RAM, with about 800 MB free for data. This is typically split into two independent ring buffers: a 600 MB partition for data collected at 5 samples/sec (5-sps) and a 200 MB partition for data collected at 15 seconds/sample (15-sec), giving on-board storage for approximately 200 days of 15-sec data and 10 days of 5-sps data. The remaining memory is used for maintenance space (e.g., firmware upgrades). Sites with limited access or telecommunications need the longest buffer possible for the nominal 15-sec data, and thus are typically configured to record only 15-sec data.

5.1.2 Data Flow

Where possible, 15-sec continuous GPS data are downloaded from PBO continuous GPS stations to the PBO Boulder Data Center (BDC) at least once per day. Some stations, particularly in Alaska, require periodic manual downloads instead. We plan no routine downloads of 5-sps GPS data; rather, we will retrieve 5-sps GPS data in response to specific triggering events (e.g., large earthquakes, volcanic eruptions, traveling ionospheric disturbances), defined through discussions with the PBO Data Products Working Group and the community.

Level 0 data from all PBO continuous GPS stations with communications infrastructure are downloaded to the PBO Boulder Data Center at least once per day using a PBO-generated version of wget. In order to meet the requirement for timely, reliable delivery of data products for the EarthScope community, PBO will develop and maintain a second data center (DC) at the USArray Array Operations Facility in Socorro, NM. This will help ensure reliable data flow in the event of system failure at either DC.

The data transport system uses a variety of data paths, in the following order of priority:

1. Direct Internet connectivity is the preferred choice of data communications at each station, but also one of the rarest solutions, given the remoteness of most PBO stations.

2. If a suitable Internet connection is not available at a given station, we typically use a Proxicast 1xRTT cellular router connection. The bulk of PBO stations currently use this communications path.

3. If 1 and 2 are not available, we typically use radio modem/radio repeaters to transfer data to an Internet or cellular router node.

4. If none of the above solutions is practical, we rely on a Very Small Aperture Terminal (VSAT) system.

5. If none of the above options for quasi-real time data communications are viable, PBO staff perform periodic manual downloads as often as compatible with PBO Operations and Maintenance funding and other practical limits.

Once at the DC, continuous GPS data will undergo two-part quality checks. First, we will compare metadata in the raw data files with the most up-to-date station metadata in the PBO Operational Database. If the POD and raw data headers are in conflict, the file in question will be flagged and removed from the forward stream, and the data will be examined to determine the proper course of action, including updating the POD to reflect new metadata correctly defined in the data file. The second part is to run the Translate, Edit, and Quality-Check (teqc) tool developed by UNAVCO to examine such measures of quality as the number of actual vs. anticipated observations per file, the number of cycle slips, and the signal-to-noise ratio as a function of azimuth and elevation; further quality-checking thresholds will be fixed through discussion with the Data Analysis Working Group. A full history of all quality measures will be stored in the POD.

The process of quality-checking converts PBO Level 0 data to Level 1 data. Those Level 1 data files passing the quality checks will be delivered to the GPS Analysis Centers and Archives using a data distribution system based on the UNIDATA Local Data Manager (LDM) (`http://www.unidata.ucar.edu/packages/ldm/index.html`) or similar software. This ensures that the Archives and Analysis Centers receive data simultaneously and that the Analysis Centers operate on the same Level 1 data files.

Figure 5.1 outlines the primary data flow from PBO continuous GPS stations.

5.1.3 Flagging Bad Data and Regenerating New Files

The quality checking system is responsible for certifying data files sent to the Archive and Analysis Centers. This raises the issue of what happens when metadata will be changing due to maintenance visits, or are found to be in error after the fact. Rather than holding data back when maintenance is scheduled for a given station, or recreating full BINEX files with correct metadata when errors are found, we will take advantage of the way BINEX handles metadata to supply only corrected metadata.

The BINEX format allows for metadata changes to be sent in the form of a BINEX metadata record that is prepended to a BINEX file when it is distributed. This allows us to simply generate a record of correct metadata to send to the Archives; this record will be prepended to the BINEX files for which the metadata are applicable as those files are distributed. Users wishing to recreate work using the incorrect

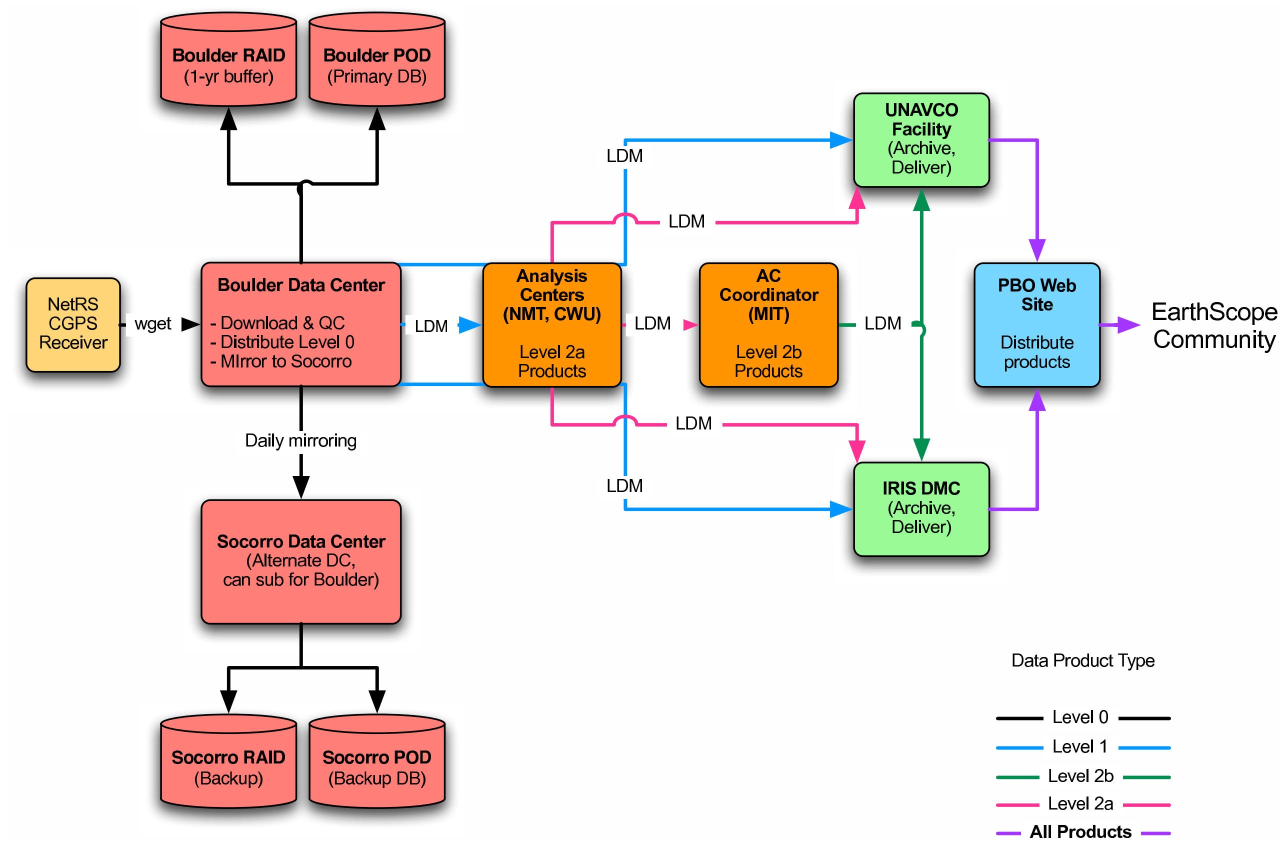

Figure 5.1: Data flow from GPS stations to the Boulder Data Center (BDC) via wget. At the BDC, the data will undergo initial quality checks (including metadata checks and updates of the Boulder POD as needed) and are copied into the ring buffer stored on the Boulder RAID. The BDC system will distribute data via LDM to the Analysis Centers and Archives. The Socorro Data Center (SDC) will be kept synchronized with the BDC on a regular basis, and will be capable of taking over the above activities in the event of system failure in Boulder. The Analysis Centers analyze the data and create Level 2a products, which are distributed to the Analysis Center Coordinator and both archives via LDM. The Analysis Center Coordinator creates Level 2b products and distributes them to both archives via LDM. The two archives mirror each other on a regular, subdaily basis. Either archive can serve data to end users directly, via the PBO web site, or through the EarthScope Integrated Data Access System (IDAS; still to be developed).

metadata will need only to ignore or strip off the additional record to obtain the original file. Corrected metadata will be included by the PBO Analysis Centers when they perform their rerun solutions (see below).

5.1.4 Data Analysis and Products

5.1.4.1 Overview

Two Analysis Centers process PBO GPS Level 0/1 data into Level 2a data products. Central Washington University operates one Analysis Center using the GIPSY processing package, and the New Mexico Institute of Mining and Technology (New Mexico Tech) operates the other using GAMIT (this was previously at the University of California, Berkeley). The Analysis Centers are overseen by the GPS Analysis Center Coordinator, an independent scientist who combines the individual Analysis Center Level 2a products to generate PBO Level 2b GPS data products.

Table 5.1 gives a summary of PBO GPS data products.

5.1.4.2 Level 0/1 data products: Raw and quality-checked data and associated metadata

Continuous GPS Level 0 products will include:

- daily raw (T00 or BINEX) files for each station containing 24 hours of raw data with 15-sec sampling;
- 5-sps raw files, downloaded as needed;
- metadata related to PBO sites, stations, sensors, raw data, and derived data products (see Section 4.2 on page 12).

UNAVCO provides converters to translate the raw data files to RINEX format. All these raw files are stored in the PBO Archives, and metadata will be stored in the PBO Operational Database, with public metadata also stored in backup databases at each PBO Archive. PBO Level 1 GPS data products are identical to Level 0 products, except for whatever metadata changes may be required following quality checks.

5.1.4.3 Level 2 data products: Low-level Derived Products and Associated Metadata

Level 2a: Loosely constrained geodetic positions and related products

Two GPS Analysis Centers, one at New Mexico Tech and the other at Central Washington University, process PBO Level 0/1 continuous GPS data products to generate Level 2a GPS data products. The New Mexico Tech Analysis Center, directed by Mark Murray, uses the GAMIT processing package, while the Central Washington University Analysis Center, directed by Tim Melbourne, uses the GIPSY processing package.

The Analysis Centers receives Level 0/1 data products from UNAVCO, and for each data file received, the Analysis Centers create three broad classes of geodetic position solutions:

Table 5.1: PBO GPS Data Products

Level	Data Product	Format	Station or network	Generation Frequency
0	15-sec raw data	T00/BINEX	Station	Daily
	5-sps raw data	T00/BINEX	Station	As needed
	SGPS Raw data	T00/BINEX	Station	Variable
	Station metadata	POD[1]	Station	Daily
1	15-sec quality checked data	T00/BINEX	Station	Daily
	5-sps quality checked data	T00/BINEX	Station	As needed
	SGPS quality checked data	T00/BINEX	Station	Variable
	Station metadata	POD[1]	Station	Daily
2a	Loosely-constrained geodetic position solution	SINEX	Network	Daily[2]
	GPS processing input and auxiliary output files	Various	Varies	As needed
	Station quality metadata	POD[1]	Station	Daily
2b	Combined loosely-constrained geodetic position solutions	SINEX	Network	Daily[2]
	Combined constrained geodetic position solutions	SINEX	Network	Daily[2]
	Combined time series of daily position	ASCII[3]	Station	Daily[2]
	Geodetic velocity solution	SINEX	Network	Periodically
	Simplified geodetic velocity solution	ASCII[4]	Network	Periodically
	Time series noise properties	ASCII[5]	Station	Periodically
	Properties of periodic time series components	ASCII[5]	Station	Periodically
	Coseismic offsets for significant events	ASCII[6]	Network	As needed
	GPS baseline time series	ASCII[5]	Station	On request
	Station quality metadata	POD[1]	Station	Daily

[1] POD: PBO Operational Database, the central repository for all PBO metadata. Metadata will also be made available in an XML format for metadata, still to be defined.

[2] "Rapid" position estimates are produced daily using rapid orbit products within 18–24 hours of receipt of data at the GPS Analysis Centers. "Final" estimates are produced approximately 15 days later with final orbit products. There will be periodic reruns with all data collected since the final rerun for which there are final orbit products. See Appendix D on page 56 for details.

[3] GPS station position time series format is defined in Appendix G.

[4] Network velocity solution format is defined in Appendix H.

[5] ASCII format to be defined.

[6] GPS coseismic offset file format is defined in Appendix I.

1. Daily rapid solutions
   These solutions are produced using the IGS rapid orbit product within six hours of the availability of that product and data for a given UTC day. The IGS rapid orbit product is usually available by 1800 UTC. These solutions include all PBO, PBO Nucleus, and selected IGS reference stations and CORS stations whose data for a given day are available at the time the IGS orbit is made available (see Appendix D on page 56 for details). This solution is primarily intended to ensure that station and network instrumentation are working properly.

2. "Final" solutions
   These are produced for a given UTC day using the IGS final orbit product, within 24 hours of the availability of that product. These solutions include any data for a given UTC day that have become available since the rapid solutions were computed. The IGS final orbits are generated weekly and are normally available on Friday for the week that ended Saturday 13 days before. The seven days of data for the week are processed with the IGS final orbits with data from GPS stations that have become available since the rapid solution added to the analysis.

3. "Supplemental" solutions
   These solutions are generated twelve weeks behind real time and are intended to allow timely analysis of data from stations that were not available at the time of the original "final" solution for a given day. The ACs only include data from stations that were not originally processed in the finals runs and three to five stations from the finals processing that are used to tie the new processing to the final solution. These supplemental solutions are run weekly after the finals for the week are complete. The ACC will combine the supplemental files with the final solution from 12 weeks before to generate SINEX and time series files with the designation "suppl".

4. Rerun solutions
   These will be produced when new models or significant amounts of new data require it, as determined by the GPS Analysis Center Coordinator and PBO Data Products Manager. For example, full reruns would be needed for model changes such as the proposed move to absolute phase center models or changes in short-period tidal loading models. These solutions will include all data received for a given station by the time of the beginning of the rerun solution. Typically, full reruns would be run once per year.

For each of these solutions, the combined PBO, Nucleus, and frame stabilization networks are subdivided into subnetworks to allow efficient processing. For the GAMIT solution, typically this includes about 30–40 stations per subnetwork. The GIPSY solutions are created using point positioning, with results formed into subnetworks for bias fixing. The subnetworks include one subset of stations spanning the entire PBO region and others focused on specific areas; these focused subnetworks incorporate a few stations from the PBO-wide subnetwork to allow the subnetworks to be joined.

For each of the solutions listed above, each Analysis Center creates a single daily (0–24 hours UTC) merged SINEX file. This file contains all the above subnetworks combined into a single loosely constrained network, with orbits constrained to the IGS orbit product. This file is the primary Level 2a data product for the given day, and is sent to the PBO GPS Archives and the Analysis Center Coordinator as rapidly as possible after each solution is completed. The Archives store these files and make them available.

In addition to the merged SINEX file, each Analysis Center will produce several additional Level 2a data products:

1. Product generation and station quality metadata
   A report giving metadata derived from each analysis run, including data completeness, latency of product generation, and station quality information such as the number of phase measurements per station, the RMS scatter of the dual-frequency phase residuals per station, and other similar information. Other relevant metadata will be determined by the Data Products Manager through consultation with the Analysis Centers, Analysis Center Coordinator, and Data Analysis Working Group. These metadata will be sent to PBO via a system still to be developed, and from PBO, changes in metadata will be made available to the Archives and the community.

2. Input and output files from processing
   All input and output files necessary for an independent user to reproduce the work of the given Analysis Center. These files will be made available in the native format of the processing software (i.e., GAMIT or GIPSY) used by the given Analysis Center. These files will be sent to the GPS Archives when processing is first started, and when a given file is updated, the updated file will be sent to the Archives at the same time as the products made using the updated file. The updated file will be kept in the Archives along with all previous versions.

Level 2b: Combined geodetic positions, velocities, and related products

The Analysis Centers send their Level 2a SINEX files to the Analysis Center Coordinator, Dr. Thomas Herring of MIT. He combines them using GLOBK to create a range of Level 2b products, including:

1. Daily rapid solutions
   Merged SINEX files based on each day's rapid SINEX files. These merged files contain network solutions that are loosely-constrained and constrained to a standard reference frame. Additionally, time series derived from the merged and constrained SINEX file; these are available in an ASCII format defined in Appendix G on page 77, and will be available in an as-yet undefined XML format. These time series will be examined automatically for significant deviations from the expected station position for a given day.

2. "Final" solutions
   Once each week, for each day of the week, the Analysis Center Coordinator generates loosely-constrained merged SINEX files by combining, using GLOBK, the "final" Level 2a SINEX files produced by each Analysis Center. The Analysis Center Coordinator also generates a merged SINEX file rotated and translated onto a North America-fixed frame; since the PBO solutions contain several IGS reference stations, these North America-fixed solutions can be easily transformed to ITRF2000. Also, the primary PBO station position time series will be derived from this combined solution. These time series are also in the format defined in Appendix G.

3. Periodic time series and velocity solution updates
   As needed, the Analysis Center Coordinator will estimate a secular velocity solution derived from the combined SINEX files, and will update the station position time series as needed. The network velocity field will be released in SINEX and in a simplified ASCII format defined in Appendix H on page 81.

   Some PBO stations may be affected by transients which make estimating a single secular velocity inappropriate. For those stations, the Analysis Center Coordinator will estimate a pre-transient and during- or post-transient velocity, and both velocities will appear in the PBO velocity field

Table 5.2: Distance/Magnitude relationship for PBO coseismic offset estimation

If magnitude is less than...	Estimate offset if distance to event is less than ...
4	7 km
5	13 km
6	44 km
7	200 km
8	1,000 km

releases, flagged with the range of epochs from which they have been estimated and the epoch on which they were estimated. These suspect stations will also be highlighted in a prominent fashion.

4. Estimates of time series noise properties and properties of periodic time series components
   The Analysis Center Coordinator will estimate long-term time series noise parameters and amplitudes and phases of periodic time series components for each station with at least six months of data available. The time series noise parameters and parameters describing periodic time series components for each station will be contained in two separate XML files, and will be versioned similarly to the velocity field XML files.

   Magtitude Distance (km)

5. Coseismic offsets for significant events inside or near the PBO network
   The Analysis Center Coordinator will also estimate coseismic offsets caused by significant earthquakes inside or near the PBO network; these offsets will be estimated for all stations with a theoretical horizontal offset of at least one mm. PBO uses an empricial magnitude/distance relationship, derived from studies of the 1999 Hector Mine, 2003 San Simeon, and 2004 Parkfield earthquakes, to determine the theoretical coseismic offsets:

   $$d_{1mm} = 0.0025 * 5^M, \tag{5.1}$$

   where M is the body wave magnitude reported by the National Earthquake Information Center (NEIC), and d_{1mm} is the distance from the hypocenter at which the estimated coseismic offset is one millimeter; hypocenters are assumed to be 5 km deep. This equation gives Table 5.2.

   These offsets will be released as rapidly as possible following the given event in an ASCII format defined in Appendix I.

6. Rerun solutions
   These will be produced whenever significant model changes, such as a new official ITRF release, warrant reprocessing of SINEX files, or when the Analysis Centers run a full rerun solution. These solutions will include all data received for a given station by the time of the beginning of the rerun solution. Typically, rerun solutions will be generated once per year.

All these products will be delivered to the GPS Archives, using LDM or similar software, as rapidly as possible after the products are generated.

The Analysis Center Coordinator will also work with PBO staff and staff of the Analysis Centers and GPS Archives to develop software to produce GPS baseline time series from the SINEX files. This software will be made available to end users so that they may produce baseline time series for selected stations and time periods; the baseline series will neither be produced routinely nor archived.

The Analysis Center Coordinator will also provide required metadata describing his/her analysis work, including parameters describing station quality, data completeness, and latency of product generation. Other relevant metadata will be determined by the Data Products Manager through consultations with the Analysis Centers, Analysis Center Coordinator, and other interested parties. All these metadata will be sent to PBO via a system still to be developed, and from PBO, changes in metadata will be made available to the Archives and the community.

Further details of the generation of PBO GPS data products can be found in the detailed GPS Analysis Plan developed by the Analysis Center Coordinator (see Appendix D on page 56).

5.1.5 Continuous GPS Data Product Archiving and Delivery

5.1.5.1 Archiving

All PBO GPS 15-sec and 5-sps Level 0 and Level 1 data products, all Level 2 data products, and all public metadata will be archived at two GPS Archives, operated by the UNAVCO Facility with the support of the IRIS Data Management Center (DMC). One archive will be located at the UNAVCO Facility in Boulder and the other at the DMC in Seattle. PBO GPS data products will be delivered to the Archives via LDM as rapidly as possible after generation, and will be freely available as rapidly as the Archives can make them available.

Level 0 and 1 GPS data products will be archived in BINEX format, with nominally one BINEX file per station per day. Some stations may be downloaded more than once per day (e.g., hourly), in which case, all subdaily files will be made available for at least 14 days. After 14 days, only daily 24-hour files need be kept. Level 2 GPS data products will be archived in SINEX, XML, and plain text formats (see Table 5.1 on page 19). Public metadata will be held in the database systems of the UNAVCO Facility (installed at both Archives).

Each Archive will have two copies of all these products (for a total of four copies). One copy at each Archive will be rapidly and freely accessible to end users at all times, barring equipment failure or downtime for necessary system maintenance. The other copy for each archive will be maintained on-site, on separate media, with data available to replace the on-disk copy within three (3) business days of failure of any part of the primary copy.

5.1.5.2 Delivery

All PBO GPS data products will be available from the GPS Archives directly, via the PBO web site, and eventually through the EarthScope Integrated Data Access System (still to be developed). PBO GPS data products will be delivered to the community through the current distribution systems of the UNAVCO Facility, including ftp and http access and the GPS Seamless Archive Center (GSAC) software tools, developed by the University of California San Diego and UNAVCO. There will also be enhancements to the current systems, particularly including a Web services-enabled version of GSAC.

Ultimately, all EarthScope raw data, derived data products, and public metadata (including those from PBO) will be available through the EarthScope Integrated Data Access System (IDAS). While

IDAS is still being planned by the EarthScope Data Working Group, it will most likely be based on Web services, and its primary goal will be to allow external groups to develop data retrieval interfaces that can "talk" to all EarthScope Archives through a common language. IDAS will provide a uniform interface to all EarthScope data and data products, regardless of which Archive holds those data/data products, without requiring special knowledge of how to access or translate those data.

Among the PBO-related items which may be available through the EarthScope IDAS are:

- General information on PBO, with links to related projects such as USArray and SAFOD
- Network installation timeline and status reports
- Site installation documentation
- Graphical and text-based network reports
- PBO data, metadata, and derived data products
- Links to modeling software and public domain PBO software
- Acknowledgments and contacts

5.2 Continuous GPS Stations: Real-time High-rate Data Flow and Products

PBO science goals, as defined in the PBO White Paper, require neither real-time data (defined as data retrieved from a GPS receiver as it is recorded by the receiver, with minimal latency) nor high-rate data (defined as data sampled at 1-sps or above). Also, the PBO Standing Committee is on record as saying that they do not feel real-time data access is a priority for PBO. Here is the relevant section of the report from their 25 February 2004 meeting:

> The PBO scientific goals do not call for real time data.... On the other hand real time access to data is useful to the surveying community. The PBO plan is to cooperate with such groups, if they provide assistance in siting and permitting, and the real time data streaming does not impact or threaten PBO data. The PBOSC supports this plan, but are concerned that the needs of the real time surveying community not become a burden to PBO staff. We urge PBO management to proceed cautiously on this.

However, PBO management recognizes that some members of the surveying, space physics, meteorology, and hazard response communities, among others, would like more rapid access (latencies from seconds to about 15 minutes) to PBO GPS data. It is also clear that continuous GPS networks worldwide are evolving toward real-time data access; PBO needs to make similar efforts to avoid obsolescence before it is completed.

5.2.1 IP-Based Real-time High-rate Data Streams

Throughout PBO's operations, PBO continuous GPS receivers will be configured as described in Section 5.1.1 on page 15. Beginning in PBO Year 3, some PBO continuous GPS stations will be configured to also stream 1-sps BINEX data in real time to a real-time system (RTS) at the Boulder Data Center. Over PBO Years 3, 4, and 5, as large a segment of the PBO continuous GPS network as can be converted without significant negative impacts on PBO budget or schedule targets or staff time will be converted to stream 1-sps data.

The RTS will split the data into two paths. First, there will be a 60-day 1-sps ring buffer, stored on disk in the PBO RAID to allow for rapid data retrieval if needed. Second, the RTS will forward the 1-sps stream to two machines, `binex.earthscope.org` and `rtcm.earthscope.org`, for distribution to end users. `binex.earthscope.org` will stream each station's BINEX data from a unique network port; users will simply connect to that port, and the 1-sps BINEX stream will flow freely to their computer with minimal latency. `rtcm.earthscope.org` will convert the BINEX streams to RTCM streams, and again stream each station's RTCM data from a unique network port for users to download freely and with minimal latency. Figure 5.2 outlines the IP-based real-time high-rate data flow model for PBO.

As these are secondary data for PBO, neither the BINEX nor the RTCM streams will be quality checked. Also, neither the BINEX nor the RTCM streams will be archived in the PBO GPS Archives, and if other groups wish to archive these streams, PBO will not fund those archiving tasks. Due to budget, schedule, and staff constraints, PBO will not guarantee the availability of any particular stream; streams that break down will be restored as rapidly as possible given other PBO constraints. Finally, these data streams may not be used for profit or gain as is, or from directly derived products, by any entity (individual, group, organization, or institution) that receives funding through PBO.

We plan to develop this real-time data flow through a collaborative process with the community. The Data Products Manager will form a PBO Real-Time Working Group (RTWG), including representatives of groups who have an interest in real-time data streams, of groups who have developed similar software, and UNAVCO and PBO staff. This working group will develop the standards and protocols to be used in developing software necessary to make real-time data streams available from PBO stations. Participation in the group will be subject to two principal restrictions:

- All standards, protocols, and software developed to generate PBO real-time data streams and make them available in usable formats will be made available as open source products under the PBO Software License.
- UNAVCO will pay for development of the software needed by PBO and will specify where the development is done and under what timeline.

Anyone not willing to agree in writing to these restrictions will be excluded from the RTWG.

5.2.2 Serial-based Real-time High-rate Data Streams

PBO GPS receivers can also support serial-based communications which PBO may use to make real-time high-rate data available in a number of possible formats (see Table 5.3 on page 27). Access to such streams will be subject to the following restrictions:

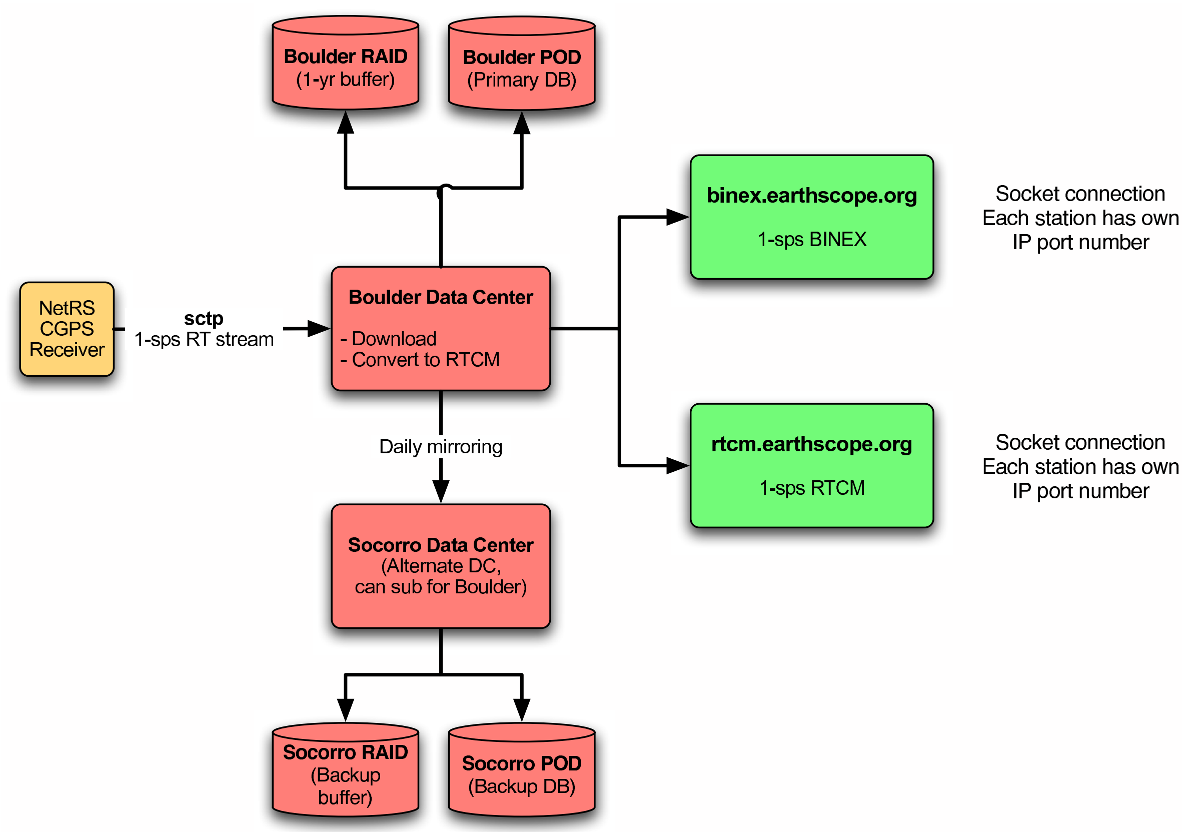

Figure 5.2: IP-based real-time high-rate GPS data flow. Trimble NetRS receivers will stream 1-sps data in real-time, in BINEX format, to `pbo.unavco.org`. This machine will fork each station's stream into a 60-day short-term ring buffer and two real-time streams, one to `binex.earthscope.org` and the other to `rtcm.earthscope.org`. `binex.earthscope.org` will make real-time BINEX streams available, with each station's stream being output from a different pre-determined network socket. `rtcm.earthscope.org` will convert the BINEX streams to RTCM and make each station's RTCM stream available from a different pre-determined network socket. Both the BINEX and RTCM streams will be freely available, but neither will be quality-checked, and stream availability is not guaranteed. 5-sps and 15-sec data will dealt with as shown in Figure 5.1.

Table 5.3: Real-time High-Rate Streaming Formats

Format	Available Sample Rates
Trimble Compact Measurement Record (CMR Standard)	1-sps; 15-, 30-sec
Trimble Compact Measurement Record v2 (CMR Plus)	1-sps; 15-, 30-sec
Trimble RS-232 real-time stream format (RT17)	1-sps; 2-, 5-, 10-, 15-, 30-sec
RTCM DGPS+RTK v2.1	N/A
RTCM DGPS+RTK v2.2	N/A
RTCM DGPS+RTK v2.3	N/A

1. Requests will be sent through the EarthScope Change Control Process by the responsible Regional Engineer to determine the possible impacts on PBO science, budget, and schedule targets. Only those requests which pass the Change Control Process may be granted.

2. Direct access to PBO serial high-rate data streams, via an external port on the equipment enclosure, may be granted to landowners in exchange for a PBO installation, operations, and maintenance permit. At their discretion, landowners may allow outside groups physical access to the enclosure output port of, or redistribution of serial data streams from, PBO stations on their land, subject to the terms of this policy.

3. Parties requesting access to PBO serial data streams will provide all necessary equipment, including power and communications equipment. All installed equipment must be properly isolated in order to avoid damaging PBO equipment.

4. The serial data stream will be output only. Absolutely no command and/or control access will be made available from the serial port.

5. There will be no physical access to the receiver. Instead, there will be an access port on the enclosure which all broadcast or recording equipment must use to gain access to the serial data stream.

6. PBO will not quality check or verify the serial data stream in any way.

7. PBO will not guarantee the availability of the serial data stream. Restoring streams that have failed due to PBO equipment will be handled as PBO staff time allows.

8. PBO data may not be used for profit or gain as is, or from directly derived products, by any person who receives funding through PBO.

5.3 Survey-mode GPS Stations: Data Flow and Products

5.3.1 Survey-mode data collected by PBO staff

Survey-mode data collected by PBO personnel, such as in response to an earthquake, will be treated as other PBO GPS data. PBO staff will download receivers to the data centers and enter metadata into the POD. From that point forward, survey-mode GPS data will flow to the Archives and Analysis

Centers through the same systems as are used for continuous GPS data. PBO survey-mode GPS data will be processed and analyzed by the PBO Analysis Centers using the same systems as will handle PBO continuous GPS data. These data, metadata, and derived data products will be made available through the same mechanisms as PBO continuous GPS data.

5.3.2 Survey-mode data collected using PBO survey-mode receiver pool

The PBO survey-mode equipment sets will be available for use by the community through scientific proposals to the EarthScope science program at NSF. PBO wishes to make a survey-mode GPS data processing service available to community members who use this equipment.

Raw data collected using PBO survey-mode equipment, and all supporting metadata, must be submitted to PBO within six months of initial data collection. PBO will treat these in an analogous fashion to data collected by PBO staff, uploading metadata, performing quality checks, and distributing those data to the PBO Archives as soon as soon as possible after data and supporting information are received by PBO. PIs may request these raw data and metadata be withheld from public access for a period of no more than two years after initial data collection; otherwise, these data will be publicly accessible immediately.

If a project PI desires the PBO Analysis Centers to analyze raw data he/she collected using the PBO survey-mode receiver pool, he/she must supply to PBO sufficient funding to cover this analysis; analogous conditions hold for analysis by the PBO Analysis Center Coordinator. The Analysis Centers and/or Analysis Center Coordinator will not process PI-driven survey-mode data collected using the PBO survey-mode pool without such funding being supplied to PBO. If Analysis Center (and/or Analysis Center Coordinator) processing is requested, and sufficient funding is made available by the PI, the Analysis Centers (and/or Analysis Center Coordinator) will process those data similarly to the methods used for continuous 15-sec data.

If these data arrive at the Analysis Centers before the “final” position solutions are estimated, the data will be included in those solutions; otherwise, they will be included in the next annual rerun. The processed results will be sent to the project PI and to the GPS Archives; the results will be made freely available immediately after their generation.

Chapter 6

PBO Strainmeter Stations and Data

6.1 Borehole Strainmeters

PBO borehole strainmeter stations consist of one Gladwin tensor strainmeter, one three-component seismometer, one pore-fluid pressure sensor, one surface barometer, one surface thermometer, one rainfall gauge, and two downhole thermistors.

From these instruments, PBO generates three levels of strainmeter data products. Level 0 products are the raw data and associated metadata. Level 1 products consist of data that have been converted to geophysical units. Level 2 is divided into two sublevels: 2a and 2b. Level 2a products include cleaned gauge data, derived time series of areal and shear strain, and time series corrections for borehole effects and the earth tides. Level 2b products include, in addition to the strain measurements, a reassessment of the information used to produce areal and shear strain such as instrument calibration and borehole relaxation and curing parameters.

Figure 6.1 outlines the borehole strainmeter data flow. Table 6.1 summarizes the borehole strainmeter data products, while Table 6.2 outlines the steps in generating the products.

6.1.1 Level 0 data products: Raw data and associated metadata

PBO Level 0 borehole strainmeter data products include raw, uncorrected gauge data from the strainmeters, raw data from the seismometer and other sensors, and metadata describing all of these data and the station.

The Gladwin strainmeters record four components of strain, with one being a redundant backup. Data from each component are recorded on-site by the GTSM Technologies datalogger at sampling rates of 20-sps, 1-sps, and 600-sec; this gives a total of 12 channels of strain data. The GTSM logger also records 36 channels of diagnostic and environmental data collected in daily files at much lower sample rates. Overall, there are 48 channels of data to be downloaded from the GTSM datalogger. In addition PBO collects atmospheric pressure at the surface and pore pressure and temperature at the pore pressure sensor at each strainmeter site.

The environmental channels collected by the GTSM datalogger at each strainmeter are (all sampled once every 30 minutes):

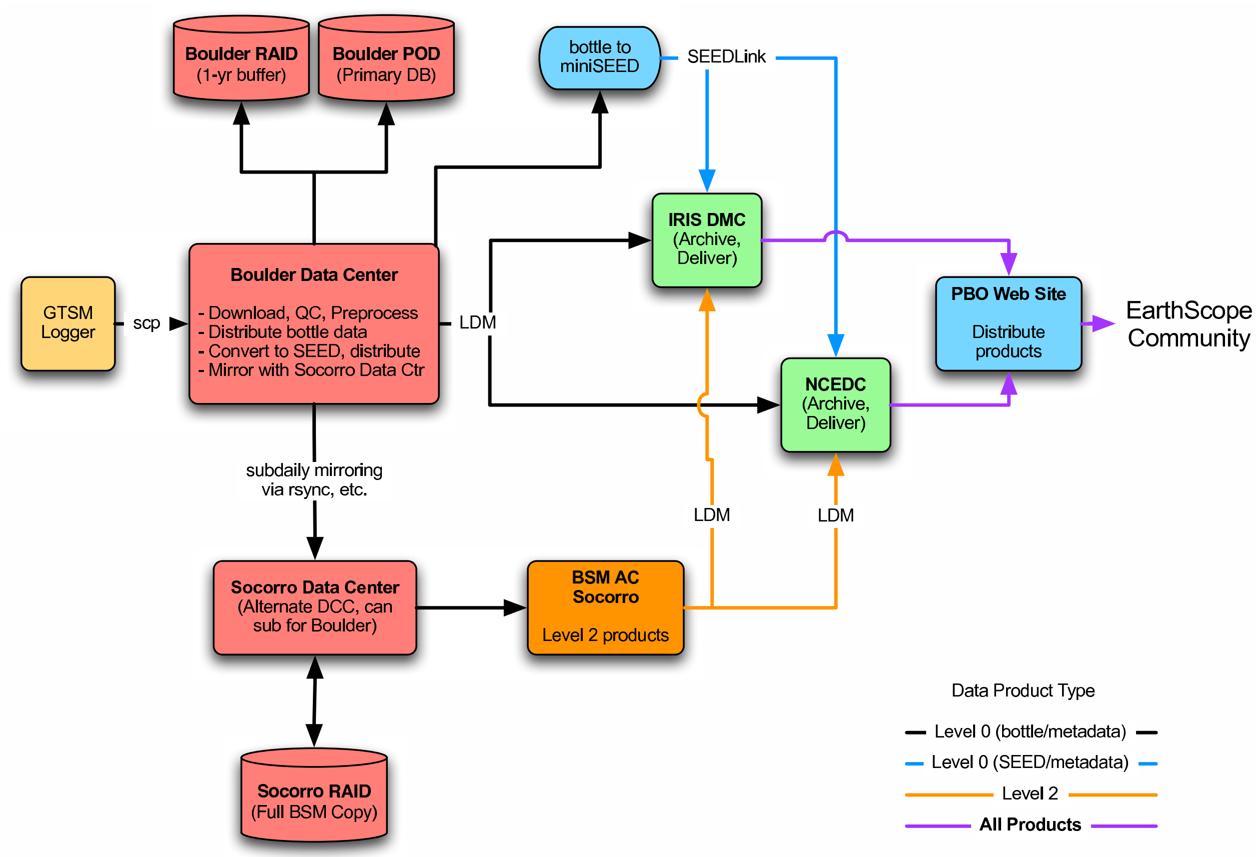

Figure 6.1: Routine borehole strainmeter data flow. The Boulder Data Center (BDC) downloads data from borehole strainmeter stations using scp; performs initial QC; and copies the data to the Boulder RAID system. All these data and metadata will also be mirrored to the Socorro Data Center (SDC) on a regular basis. The BDC distributes Level 0 and 1 strain data products to the Borehole Strainmeter Analysis Center at the UNAVCO Socorro Facility and to the Archives at the Northern California Earthquake Data Center (operated by UC Berkeley) and the IRIS Data Management Center in Seattle. In the event of system failure at the BDC, the SDC will be able to take over these functions, helping to ensure continued stable data flow. The Borehole Strainmeter Analysis Center analyzes the Level 0/1 data products to create Level 2a and 2b strain data products, which are distributed to the archives. Either archive can serve data to end users directly, via the PBO web site, or through the EarthScope Integrated Data Access System (still to be developed).

Table 6.1: PBO Borehole Strainmeter Data Products

Level	Data Product	Format	Generation Frequency
0	4 channels each of 20-sps, 1-sps, and 600-sec strain gauge data	bottle/SEED	Hourly
	31 channels of diagnostic and SOH data	bottle/SEED	Daily
	1 channel of 1-sps surface atmospheric pressure data	SEED	Hourly
	1 channel each of 10-sec downhole pore pressure and temperature	ASCII/SEED	Hourly
	1 channel each of 30-min surface rainfall and atm pressure	bottle/SEED	Daily
	1 channel each of 30-min temperature at logger, powerbox, and strainmeter	bottle/SEED	Daily
	Station metadata	POD/XML[1]/dSEED[2]	Daily
1	20-sps, 1-sps, and 600-sec auto-scaled gauge data	bottle/miniSEED	Hourly
	20-sps, 1-sps, and 600-sec quick-look areal and shear strain series	bottle/miniSEED	Hourly
	Auto-scaled 1-sps pore pressure	ASCII/SEED	Hourly
	Auto-scaled 1-sps barometric pressure and downhole temperature	bottle/miniSEED	Hourly
	Auto-scaled 300-sec surface rainfall and temperature	bottle/miniSEED	Hourly
	Auto-scaled environmental data	bottle/miniSEED	Daily
2a	300-sec cleaned linearized gauge data	XML	14 days[3]
	300-sec cleaned and scaled environmental data	XML	14 days
	300-sec areal and shear strain time series computed from cleaned gauge data	XML	14 days
	300-sec borehole and grout curing correction time series	XML	14 days
	300-sec areal and shear strain tidal correction time series	XML	14 days
2b	300-sec cleaned linearized gauge data	XML	3 months[4]
	300-sec cleaned and scaled environmental data	XML	3 months
	300-sec areal and shear strain time series computed from cleaned gauge data	XML	3 months
	300-sec borehole and grout curing correction time series	XML	3 months
	300-sec areal and shear strain tidal correction time series	XML	3 months
	Phase and amplitude for main tides for each gauge and tensor strain	XML, POD	3 months
	Borehole relaxation and curing coefficients for each gauge	XML, POD	3 months
	Strain calibration matrices[5]	XML, POD	3 months
	Outlier criterion[6]	XML, POD	3 months

[1] XML: Strainmeter XML data format defined in Appendix N.
[2] dSEED: Dataless SEED format.
[3] These series will be updated no less often than once every two weeks, using curing models and other parameters estimated at Level 2b.
[4] These series and correction model parameters will be updated every 3 months. The model parameters computed here will also be used to update the Level 2a data products.
[5] See Appendix J.1 on page 89.
[6] The amount by which adjacent data points are permitted to differ. Initially set to five times the inter-quartile range based on 1 month of cleaned data.

Table 6.2: PBO Borehole Strainmeter Data Processing Steps

Step	Procedure	Output
	Receiving and archiving the raw data	
1	Import tar files produced by the data logger	
2	Unpack and merge data logger files	
3	Convert raw data to miniSEED format	
4	Send miniSEED and bottle data to archives	Level 0 Data Set
5	Raw data are scaled to natural units (performed by user)	Level 1 Data Set
	Rapid data quality check (every 24 hours)	
6	Linearize the gauge data	
7	Remove tidal signal and borehole effects from linearized gauge data	
8	Autocheck the detrended and detided data for outliers	
	Manual cleaning and verification of data (no more than 14 days latency)	
9	Reduce the raw 1-sps data to 300-second data	
10	Linearize the 300 s gauge data	
11	Remove tidal signal and borehole effects and inspect the residual gauge data using field metadata, field system diagnostics, and outlier criteria	
12	Apply edits to raw 1-sps data, linearize, and reduce to 300 s intervals	
13	Generate and remove borehole trends	
14	Generate the tidal correction	
15	Compute areal and shear strain	
16	Convert to XML	Level 2a Data Set
	Post-processing and derivation of processing parameters (every 3 months)	
17	Reprocess all 300-sec data including additional field information and editing if required (Steps 11 to 15 repeated)	
18	Recalculate the borehole relaxation and grout curing parameters using entire series	
19	Compute the tidal signal in the clean gauge, areal and shear strain data	
20	Recalibrate	
21	Reassess the outlier criterion	
22	Recalculate areal and shear strain and corrections using output from steps 18, 19, 20, and 21	
23	Convert to XML, store processing parameters	Level 2b Data Set

- 1 channel of surface rainfall
- 1 channel of surface atmospheric pressure
- 2 channels of surface temperature
- 1 channel of temperature at the strainmeter

In addition we record:

- 1 channel of surface atmospheric pressure, sampled once per second (collected on the Q330)
- 1 channel of downhole pore fluid pressure, sampled once per 10 seconds (collected on a local PC)
- 1 channel of temperature at the pore fluid pressure, sampled once per 10 seconds (collected on a local PC)

The GTSM data logger collects data in bottle format. Each 20-sps channel is stored in bottle files that span 1 minute of time, giving four 1-minute-long bottle files per minute of 20-sps data. Each hour, there are four 1-hour-long bottle files containing data sampled at 1-sps. Each hour, the 240 1-minute-long 20-sps bottle files are assembled on the logger into a single gzip-ed tar file for ease of downloading. Similarly, the four 1-hour-long 1-sps bottle files are assembled into a single gzip-ed tar file for ease of downloading. The four 600-sec channels and 36 diagnostic channels are stored in 40 1-day-long bottle files. The pore pressure related data are collected in 1-hour long ASCII formatted files while the 1-sps atmospheric data are collected in miniSEED format.

The GTSM logger has onboard storage for 14 days of 20-sps data, 30 days of 1-sps data, and 1 year of 600-sec and diagnostic data. To allow for additional storage in case of communications failure, PBO will install at each strainmeter station a local buffer computer with 10–40 GB of onboard disk storage. This computer will use scp, wget, rsync, or other similar data mirroring tools to download data from the GTSM logger every hour, thus acting as a local mirror. Every hour, the Boulder Data Center will download data from the local buffer (or the GTSM logger, if the buffer computer fails), using wget, scp, rsync, or similar software.

Once in Boulder, the 20-sps and 1-sps gzip'ed tar files are unpacked and uncompressed. The 60 one-minute-long bottle files for each 20-sps channel are then merged into a single 60-minute long bottle file for each channel; thus 240 1-minute files become four 60-minute bottle files. The 1-sps, daily 600-sec and diagnostic files do not require merging.

Once the files have been downloaded (and merged, in the case of the 20-sps data), they will undergo initial quality checks similar to the GPS data. Metadata in the borehole strainmeter data files are compared with the most up-to-date station metadata in the PBO Operational Database. If the POD and raw data headers are in conflict, the file in question will be flagged and removed from the forward stream, and the data will be examined to determine the proper course of action, including updating the POD to reflect new metadata correctly defined in the data file. The borehole strainmeter data will be checked for the time of last data returned, the completeness of data over some time period, the ability of the strainmeters to track tidal signals, and data outliers; further criteria will be determined by the PBO Borehole Strainmeter Data Analyst in consultation with GTSM Technologies and the PBO Strainmeter Working Group.

All the 60-minute-long and one-day-long bottles are stored in a one-year ring buffer on a RAID storage system at the Boulder Data Center to allow easy correction and replacement of data files found to be in error, and any metadata updates will be stored in the POD via an automated process. The one-hour 1-sps, one-hour merged 20-sps, 1-sps, and 1-day bottle files (600-sec and diagnostic) will be mirrored to the RAID systems at the Socorro Data Center every hour via rsync or similar means, to allow generation of the Level 1 and 2 borehole strain data products, as described below.

Currently, the one-hour-long 1-sps bottles and the 1-day bottles (600-sec and diagnostic) are converted to miniSEED via software originally developed by staff of the Northern California Earthquake Data Center. The miniSEED data are delivered to the Archives using SEEDLink. Three day-long tars are made of the raw bottle data; a 20-sps tar, a 1-sps tar and a low frequency data tar file. Other than bundling them into the tar files the data are in the form they were pulled from the datalogger. The bottle tars are delivered, using LDM, to the PBO Strainmeter Archives at the Northern California Earthquake Data Center and the IRIS Data Management Center. Data delivery typically starts within two weeks of instrument installation, with the delay allowing for initial instrument and data flow testing.

6.1.2 Level 1 data products: Scaled data and associated metadata

Level 1 products are the raw data converted to natural units, and except for the conversion to natural units, are identical to the Level 0 data. Because the Level 1 data can be obtained simply by multiplying the Level 0 data by the appropriate scale factor, they are not archived independently. Rather, users can obtain Level 0 data and scale factors from the archives and apply the scale factor to create the Level 1 data products.

Level 1 products will also include a "quick look" areal and shear strain data set. PBO will provide the tools to transform the Level 1 scaled gauge data into areal and shear strain. The user will therefore be able to obtain unedited tensor strain data rapidly. The matrices required to combine the individual gauge data into tensor strain will be derived from a tidal analysis of the gauge data and will not be available for at least three months after installation.

GTSM Technologies have developed a system that produces several streams of automatically processed data as soon as the raw data arrives at the Borehole Strainmeter Analysis Center [Gladwin, November 2004 documentation]. These data streams will be retained at PBO while the system is being tested and will be available to the community upon request.

Another aspect of Level 1 processing is an initial data quality assessment, performed at the Borehole Strainmeter Analysis Center. The output of this quality check will be used in generating the Level 2 products. The goal of this initial processing is to rapidly identify potential problems, such as spurious or missing data points, in the data from each sensor. If any such problems are found, an e-mail will be sent to the Strainmeter Data Analyst, alerting her to examine the particular data segment with extra care. The goal is not to correct any of these potential problems; those corrections will occur as part of the Level 2 processing.

The rapid data quality check processing steps for strain data are shown in Figure 6.2 and outlined in Table 6.2, steps six through eight; analogous checks will be performed for the environmental and diagnostic data sets. Here is a summary of the quality checks:

- Linearize the data (Step 6)

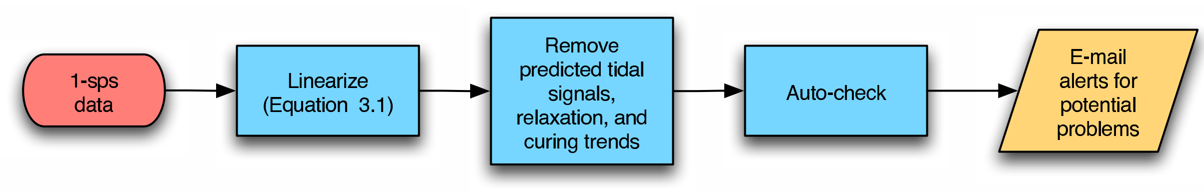

Figure 6.2: Rapid borehole strainmeter data check flowchart.

The raw measurements from each gauge are output from capacitance bridges. To convert the gauge output to change in length along the axis of the gauge, the following equation is used

$$U = \frac{R_{raw}}{1 - R_{raw}} - \frac{R_0}{1 - R_0}, \quad (6.1)$$

where U is the linearized data point, R_{raw} is the raw data point, and R_0 is a data point early in the record.

- Remove the tidal and borehole relaxation and grout curing trends from linearized data (Step 7)
  The tidal signal, borehole relaxation and grout curing trends are removed from the gauge data to facilitate checking for spurious data. The tidal signal and borehole effect trends are generated using parameters derived as Level 2b products.

- Check the detrended and de-tided data for outliers and offsets (Step 8)
  The data are checked for outliers and offsets by examining the difference between adjacent data points. If two adjacent differences exceed an outlier criterion, the common data point is flagged as an outlier. The outlier criterion will be recalculated every three month for each instrument (Level 2b product). Email will be generated reporting any detected problems such as, offsets, outliers and missing data. These findings will be examined when producing the level 2 data set.

6.1.3 Level 2 data products: Low-level Derived Products and Associated Metadata

Level 2 borehole strainmeter data products include strain gauge and environmental data, scaled and corrected with human input, and derived areal and shear (γ_1 and γ_2) strain time series. Also included in Level 2 products are a variety of correction time series, including earth tides, borehole and grout curing, and edits. All Level 2 series are decimated down to 300-sec sampling, and are archived in an eXtensible Markup Language (XML) format defined in Appendix N. The XML files are versioned so that previously released versions remain available to the community, and each file contains one year of data.

There are two sublevels of Level 2 derived products. Level 2a products will be no more than 14 days out of date, and involve initial corrections to data available at that time. Level 2b products involve more extensive corrections, and will thus be delayed by four months.

Level 2 products require the phase and amplitudes of the M2 and O1 tides to be known. These parameters are required to compute the matrices used to combine the gauge measurements into areal and shear strain. The tidal amplitudes are also required to compute the tidal correction. It takes two to three months of strainmeter data to estimate these tidal parameters. The borehole relaxation and curing rates are largest in the months immediately following installation and can complicate tidal analysis. Level 2 products will therefore not be available for three to six months after instrument installation.

Table 6.1 summarizes the Level 2 products, and Table 6.2 lists the steps needed to create them.

6.1.3.1 Level 2a: Initial Human-Corrected Products

Level 2a processing consists of editing the data and forming the tensor strain and time series corrections for the earth tides and borehole and grout curing effects. The steps are shown in Figure 6.3 and outlined in Table 6.2, steps nine through sixteen. They are summarized as follows:

- Downsample raw gauge data to 300-second sampling (Step 9)
  The 1-sps data from the borehole strainmeter are lowpassed and decimated to produce 300-second samples using a series of filter-and-decimate stages. At each stage, we use lowpass filters designed to provide a flat passband, minimal aliasing at zero frequency, and a minimum-phase finite impulse response (FIR filters). The filtering and decimation are implemented causally, because this minimally distorts the onset arrivals from earthquakes [*Scherbaum and Bouin*, 1997]. The phase response can be computed from the filter weights given at each stage, which will be provided in the SEED and XML metadata.

- Linearize the data (Step 10)
  Linearize the 300-sec gauge data as in Equation 6.1.

- Data editing (Step 11)
  Level 2a data cleaning and editing are not automated processes. In addition to using an outlier criterion to detect bad data, metadata such as information from the field engineers and the systems diagnostic data steams are used as input to the cleaning process. Problems detected during the rapid data quality check indicating the presence of spurious data will be investigated by inspecting the data. Auxiliary data such as instrument voltages and environmental data are also checked. As for the rapid data quality check, the tidal signals and curing trends are removed from the gauge data to facilitate cleaning. Any data that are considered outliers are recorded and flagged as such in the Level 2 XML file. The GTSMs provide a relative measurement of the instrument diameter and, unlike hydraulic strainmeters, do not have reset offsets. For this reason, only offsets that are known to be non-tectonic in origin, such as those caused during site maintenance, will be removed. The magnitude and time of any offset that is removed will be recorded in the level 2 XML file. Coseismic offsets will not be removed but their magnitude will be recorded in the XML file for those who wish to do so.

- Apply edits to raw data (Step 12)
  The edits recorded in step 11 are applied to the 1-sps data. The data are then linearized and reduced to 300-sec sampling. The tidal and borehole effect signals are not removed from the data.

- Generate and remove borehole trends (Step 13)
  Time series of the borehole relaxation and grout curing effects are generated for the gauge data. The borehole effects are subtracted from the edited gauge data to form a set of gauge residuals.

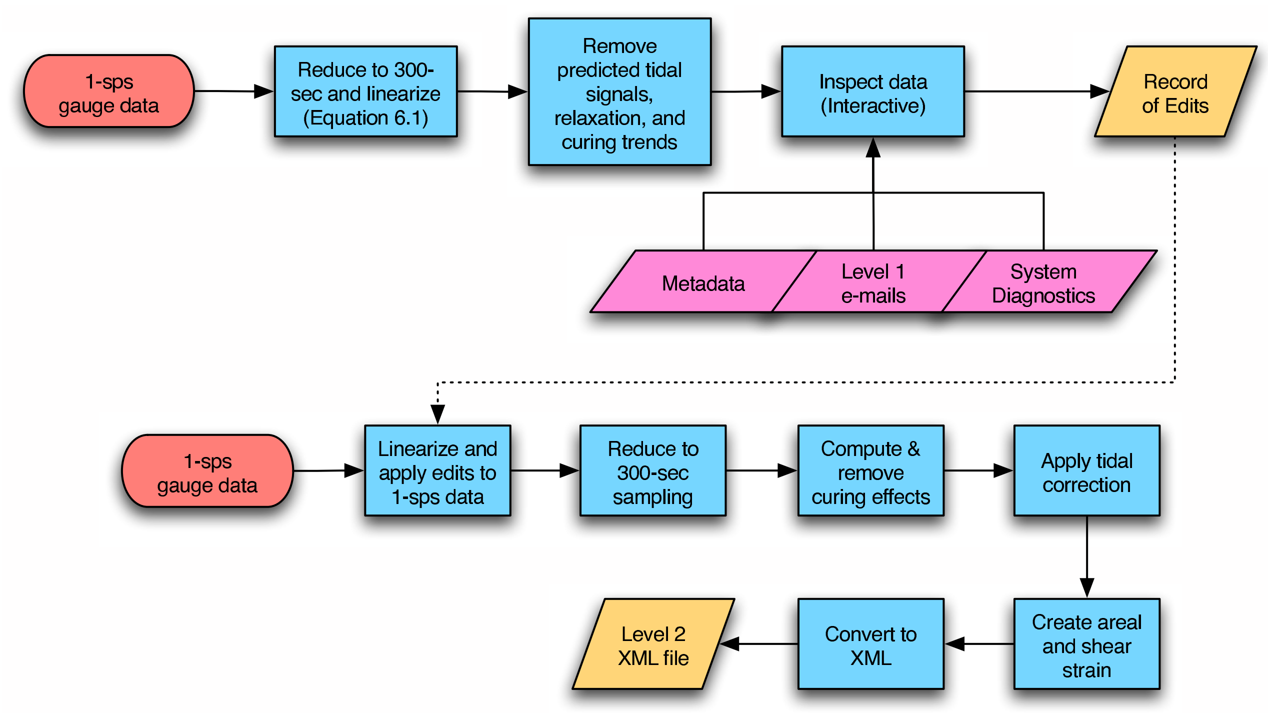

Figure 6.3: Level 2a processing steps.

- Generate the tidal correction (Step 14)
  A time series of tidal corrections are generated for the areal and shear strain data. The parameters used to generate the corrections will be determined at Level 2b and reevaluated every four months.

- Compute areal and shear strain (Step 15)
  The areal and shear strain measurements are computed from both the residual gauge data and gauge data that has not had the borehole effects removed. The tidal signal is not removed from this data set. See Appendix J.1 for details.

- Conversion to XML (Step 16)
  The processed data are converted to XML. The XML files contain all the Level 2 cleaned strain data and edit information. They also list the processing parameters used to detect outliers, estimate the tidal and borehole curing effects and matrices used to combine the gauge data into areal and shear strain.

These products are sent to the strain archives as rapidly as possible after generation using the LDM software suite.

6.1.3.2 Level 2b: Final Human-Corrected Strainmeter Products

Level 2b processing provides an opportunity for much more detailed postprocessing of each station's entire dataset. At four-month intervals the entire 300-second interval data set will be reprocessed to form the Level 2b data set. The clean gauge and tensor strain time-series are regenerated by applying the edits found at Level 2a and any additional editing found necessary. This allows detection of long-term component problems, allow better communication of results from field maintenance, and so on.

In addition to reprocessing the strain time series, Level 2b processing includes the following analyses of the cleaned data, summarized in steps 17 through 23 of Table 6.2:

- Reprocess entire data 300-second data set (Step 17)
  Additional edits are made to the data that were not made at Level 2a.
- Recomputation of the borehole relaxation and grout curing detrending parameters (Step 18)
  The detrending parameters will be recalculated using the entire data record and applied to the Level 2b data set. The parameters determined are then used to calculate the gauge corrections for the next 3 months of Level 2a data.
- Computation of the observed tides in the gauge, areal and shear strain data (Step 19)
  The observed amplitude and phases of the earth tide signal will be calculated for the last three months of measurements. These parameters are used to calculate the tidal correction for the strain data.
- Recalculation of calibration matrices (Step 20)
  The instrument will be recalibrated using the last three months of observed tidal parameters.
- Reassess the outlier criterion (Step 21)
  The criteria for detecting outliers is recalculated and updated if needed.
- Recalculate areal and shear strain (Step 22)
  Areal and shear strain are recalculated using the output from step 20. The borehole trends and tidal corrections are recalculated using the output from steps 18 and 19.
- Convert to XML (Step 23)
  Data and processing parameters, borehole and grout curing parameters, tidal amplitude and phases, calibration matrices will be converted to XML and archived. We expect that the processing parameters will change with time and the entire series of parameters will be contained in the XML header.

Level 2b borehole strainmeter data products are sent, using LDM, to the strainmeter archives as rapidly as possible after product generation.

6.1.4 Borehole Strainmeter Data Archiving and Delivery

All PBO borehole strainmeter data products will be archived at the PBO Strain Archives at the Northern California Earthquake Data Center (NCEDC) and IRIS Data Management Center. As of April 3, 2006, the strain archives receive the following Level 0/1 data files from PBO:

- Bottle format
  - 4 channels of 20-sps data, in four sets of 60 1-minute-long files per hour;

  - 4 channels of 1-sps data, in four 60-minute-long files per hour;
  - 4 channels of 600-sec data, in four 1-day-long files per day; and
  - 36 channels of lower sample rate diagnostic data and environmental data, in 36 1-day-long files per day

- ASCII format
  - 1 channel of 10-sec pore pressure data, one 60-minute long file per hour
  - 1 channel of 10-sec temperature at the pore pressure sensor, one 60-minute long file per hour

- miniSEED format
  - 4 channels of 1-sps data, in four 60-minute chunks per hour;
  - 4 channels of 600-sec data, in four one-day chunks per day;
  - 36 channels of lower sample rate diagnostic data and environmental data, in 36 1-day-long chunks per day;
  - 1 channel of 10-sec pore pressure data, one 60-minute-long file per hour;
  - 1 channel of 10-sec sps temperature at the pore pressure sensor, one 60-minute long file per hour
  - 1 sps atmospheric pressure data (from the Q330)

The bottle format data are delivered to the archives using LDM on a daily basis. The 1-sps miniSEED data are delivered on an hourly basis and the lower sample rate miniSEED data delivered on a daily basis using the free SEEDLink software. The miniSEED format data are delivered to the archives on a similar schedule, using the free SEEDLink software. The Archives will store both bottle and miniSEED files in order to serve the widest range of users possible — the seismic community commonly uses the SEED standard for data exchange, while the borehole strainmeter community commonly uses the bottle files directly — as well as to ensure that any corruption that may occur during the translation from bottle to miniSEED formats may be corrected at a later date. Level 1 products are not archived separately, but instead the scaling factors necessary to convert from Level 0 to Level 1 will be stored in the POD and PBO strain SEED volumes. Level 2 products are delivered to the archives using LDM, no less often than once every 14 days, and are stored in the XML format defined in Appendix N. Finally, the archives also hold a copy of the PBO public metadata related to strainmeter stations and data, in dataless SEED format.

Each archive will store at least two copies of the full set of PBO strain data products. One copy will be rapidly and freely accessible to end users at all times, barring equipment failure or downtime for necessary system maintenance. The other copy will be maintained on-site, on separate media, with data available to replace the primary copy nominally within 96 hours of failure of the primary copy. The archives simply store the Level 0 bottle files and Level 2 XML files, and make them available through anonymous ftp and simple http access. They will also fully manage, archive, and distribute the Level 0/1 miniSEED data to EarthScope users through the Data Handling Interface, NetDC, and other similar tools and installed at both facilities. In addition, these groups will make strain data available via the PBO web site and will make them available through the EarthScope Integrated Data Access System (still to be developed).

6.1.5 Borehole Seismic Data

In addition to the strainmeter data, each PBO borehole strainmeter station will collect three channels of 100-sps seismic data. The USArray element of EarthScope will manage these data in a very similar manner to that from the Transportable Array element of USArray. Seismic data from PBO borehole seismometers will flow into a series of Antelope Object Ring Buffers (Orbs). The USArray Array Network Facility will receive data from those Orbs, perform quality checks similar to that for the Transportable Array, and will maintain metadata for PBO seismic data. The IRIS DMC will receive waveforms in real time directly from the Orbs, run the Quality Assurance Framework on all PBO seismic data, and archive all PBO seismic data and products derived from them. All PBO seismic data and products derived from them will be accessible from the IRIS DMC either directly or through the EarthScope Integrated Data Access System.

PBO will create no Level 1 or 2 data products from PBO seismometers. PBO will generate raw waveform data only, not such higher-level products as phase picks, amplitudes, earthquake catalogs, etc. PBO waveforms may be integrated into other data products, such as those from USArray, and PBO seismic data may be used in conjunction with PBO strainmeters for calibration or other purposes, but higher-level products will not be generated by PBO directly.

6.1.6 Physical Samples and Borehole Logs

During the installation of the PBO borehole strainmeter stations, crews collect physical samples and conduct geophysical logging, both in two phases. As the initial hole is drilled, samples are collected at depth intervals of approximately three to four meters, with approximately 500 cc of rock cuttings in each sample. These are washed, set out to dry, and packaged in standard 4-inch by 6-inch sample bags, each of which is labeled with the station name, date, and depth range from which the same was collected. PBO crews will collect approximately 50 samples per borehole.

Once the primary drilling is complete, the PBO collects a series of geophysical well logs, including:

- Full sonic suite (Vp, Vs, tube wave velocity, and Poisson's ratio)
- Acoustic televiewer
- Caliper
- Natural gamma
- Spontaneous potential (SP)
- Resistivity
- Temperature

These logs are collected digitally and are stored both in digital and paper form. Digital form is stored on CD-ROM, with one CD-ROM per hole.

Once the hole drilling, logging, and casing are completed, crews drill the borehole to the target depth, collecting continuous core samples as drilling progresses. Crews collect 30 to 100 feet of PQ (3.3-inch outer diameter) whole core in 10-feet-long segments. Each segment is laid out on a temporary storage rack and marked with the core number for that particular hole, the depth at the top of the core and in onefoot intervals thereafter, and with the orientation marked in the usual standard of red and black lines running vertically down the core (with red always on the right when the core is oriented with the shallowest part of the core at the top and the deepest part at the bottom); crews also mark an arrow pointing toward the shallow end of the core. These labels allow the PBO to orient the core segments properly in a vertical sequence. The core is then cut into 2.5-foot lengths, and adjacent lengths are stored in a 32-inch long, 8-inch wide, 4-inch tall box with two internal divisions to hold the lengths; thus each box contains 5 feet of core, and crews collect approximately 6 to 20 boxes per borehole.

Once the coring is completed, crews collect another series of geophysical well logs with the same parameters as the first; these are also stored digitally and in paper form, with the digital copy on individual CD-ROMs for each hole.

All these physical samples and digital logs will be stored by the Houston Research Center, an NSF-funded core and log repository that is part of the Texas Bureau of Economic Geology. The Houston Research Center is funded by NSF to hold physical samples and logs from all NSF-funded terrestrial coring projects in perpetuity, without cost to the investigator who collects and submits the samples to the Houston Research Center. The Houston Research Center then makes samples and logs available to users under terms of a physical sample policy jointly developed by EarthScope, NSF, and the Houston Research Center Advisory Council.

6.2 Laser Strainmeters

Each PBO laser strainmeter collects 12 channels of 16-bit digital data, sampled at nominal rates of 1 sample/second and 300 seconds/sample. These channels include

- 2 channels for the laser fringe count;
- 2 channels for correction signals from the optical fiber anchors at each end monument;
- 5 environmental channels: rainfall, outside air and ground temperatures, barometric pressure, pyranometer;
- and 3 instrument health channels: vacuum pressure, and temperature readings from either end of the strainmeter.

PBO produces three levels of laser strainmeter data products. Level 0 products are the raw data and associated metadata. Level 1 products consist of data that have been converted to geophysical units. Level 2 is divided into two sublevels: 2a and 2b. Level 2a products include cleaned fringe data, derived time series of linear strain, and time series corrections for environmental, tidal, vacuum pressure, laser frequency drift, and other effects. Level 2b products include, in addition to the strain measurements, a reassessment of the information used to produce linear strain, such as instrument calibration, environmental, vacuum pressure, laser frequency drift, and tidal parameters.

Figure 6.4 outlines the laser strainmeter data flow, and Table 6.3 summarizes the laser strainmeter data products.

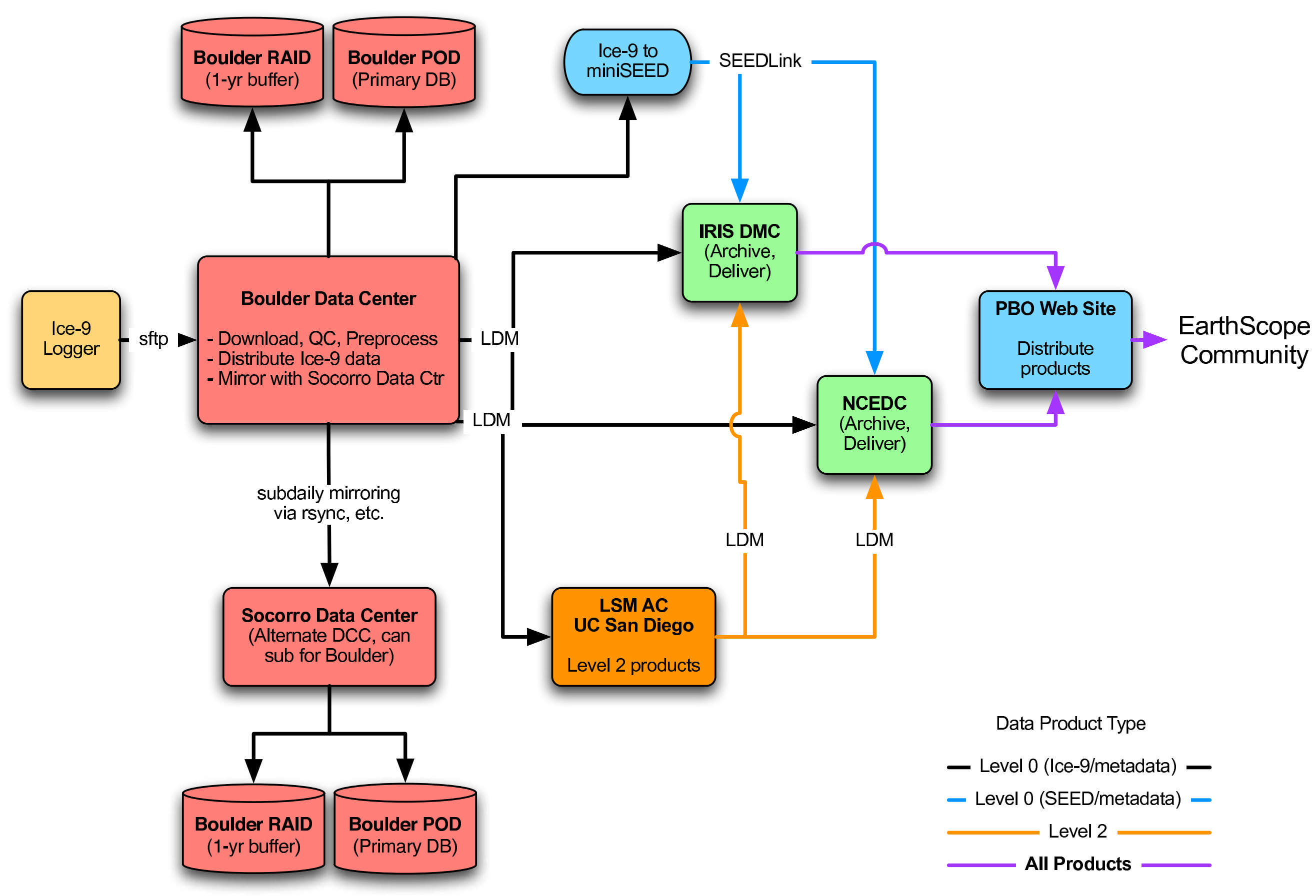

Figure 6.4: Routine laser strainmeter data flow. The Boulder Data Center (BDC) receives data from laser strainmeter stations, performs initial QC, and copies the data to the Boulder RAID system; all these data and metadata will also be mirrored to the Socorro Data Center (SDC) on a regular, subdaily basis. The BDC distributes Level 0 and 1 strain data products to the Laser Strainmeter Analysis Center at UC San Diego and to the archives at the Northern California Earthquake Data Center and the IRIS Data Management Center in Seattle. In the event of system failure at the Boulder DC, the Socorro DC will be able to take over these functions, helping to ensure continued stable data flow. The Laser Strainmeter Analysis Center analyzes the Level 0 and 1 data products to create Level 2a and 2b strain data products, which are distributed to the archives using LDM. Either archive can serve data to end users directly, via the PBO web site, or through the EarthScope Integrated Data Access System (still to be developed).

Table 6.3: PBO Laser Strainmeter Data Products

Level	Data Product	Format	Generation Frequency
0	12 channels of 1-sps LSM fringe count, environmental, and instrument state of health data	Ice9/miniSEED	Daily
	12 channels of 300-sec LSM fringe count, environmental, and instrument state of health data	Ice9/miniSEED	Daily
	Station metadata	POD/XML[1]/dSEED[2]	Daily
1	1-sps auto-scaled laser fringe data	Ice9/miniSEED	Daily
	300-sec auto-scaled laser fringe data	Ice9/miniSEED	Daily
	1-sps auto-corrected and -scaled environmental data	Ice9/miniSEED	Daily
	300-sec auto-corrected and -scaled environmental data	Ice9/miniSEED	Daily
	1-sps auto-corrected and -scaled state of health data	Ice9/miniSEED	Daily
	300-sec auto-corrected and -scaled state of health data	Ice9/miniSEED	Daily
2a	300-sec human-corrected and -scaled linear strain data	XML	2-week latency
	300-sec human-corrected and -scaled environmental data	XML	2-week latency
	300-sec human-corrected and -scaled state of health data	XML	2-week latency
2b	300-sec human-corrected and -scaled linear strain data	XML	3-month latency
	300-sec human-corrected and -scaled environmental data	XML	3-month latency
	300-sec human-corrected and -scaled state of health data	XML	3-month latency
	Earth tide model	XML	3-month latency
	Laser frequency drift correction	XML	3-month latency
	Vacuum pressure drift correction	XML	3-month latency
	Optical anchor correction	XML	3-month latency

[1] XML: Strainmeter XML data format defined in Appendix N.
[2] dSEED: Dataless SEED format.

6.2.1 Level 0 data products: Raw data and associated metadata

Level 0 laser strainmeter data products include raw laser fringe counts, fiber anchor series, and environmental series. Each PBO long-baseline laser strainmeter collects 12 channels of 16-bit digital data, sampled at nominal rates of 1 sample/second and 300 seconds/sample.

Laser strainmeter data are recorded by each station's Ice9 data logger, which can store more than one year of laser strainmeter data as an on-site buffer. Each day, the logger pushes two one-day-long Ice9 data files, one for the 1-sps data and the other for the 300-sec data, to the Boulder Data Center (or the Socorro Data Center if the Boulder facility is down) using sftp. Once at the Boulder Data Center, the data will undergo initial quality checks similar to the borehole strainmeter data. All the Ice9 files are stored in a one year ring buffer on a RAID storage system at the Boulder Data Center to allow easy correction and replacement of data files found to be in error, and any metadata updates will be stored in the POD via an automated process. The Ice9 files will also be mirrored to the RAID systems at the Socorro Data Center every day via rsync or similar means.

The Ice9 files are sent, using LDM, to the Laser Strainmeter Analysis Center, run by the Agnew/Wyatt research group at Scripps Institution of Oceanography. The Laser Strainmeter Analysis Center staff uses these data to create Level 2 laser strainmeter data products.

The Ice9 files are sent to the PBO strain archives at the Northern California Earthquake Data Center and the IRIS Data Management Centers using LDM. Also, the Ice9 files are converted to miniSEED data, which are sent to the strain archives using SEEDLink. Metadata related to laser strainmeter data are sent to the archives in dataless SEED format, and will be updated in the POD through a system still to be determined. Data delivery will start within two months of instrument installation, with the delay allowing for initial instrument and data flow testing.

6.2.2 Level 1 data products: Scaled data and associated metadata

Level 1 laser strainmeter data products are the raw data converted to natural units, and except for the conversion to natural units, Level 1 data products are identical to the Level 0 data. Because the Level 1 data can be obtained simply by multiplying the Level 0 data by the appropriate scale factor, they are not archived independently. Rather, users may obtain Level 0 data and scale factors from the Archives and apply the scale factor to create the Level 1 data products.

Level 1 products will also include a "quick look" linear strain data set, with corrections applied for laser frequency drift, vacuum pressure drift, the motion of the end-monuments of the strainmeter (via optical fiber anchors), and optionally earth tides. The end-monument correction is available from the beginning of data collection, while the laser frequency and vacuum pressure corrections require at least four months of data and are thus only available after the first Level 2b data products are released (see below). These corrections will be re-evaluated with each release of Level 2b data products.

Another aspect of Level 1 processing is preparing to produce the Level 2 products. The first step of this is for all strain and environmental data arriving at the Laser Strainmeter Analysis Center to undergo an initial data quality assessment. The goal of this initial processing is to rapidly identify potential problems in the data from each sensor, for later correction as part of the Level 2 processing.

6.2.3 Level 2 data products: Low-level Derived Products and Associated Metadata

Level 2 laser strainmeter data products include linear strain time series, and environmental data, all scaled, edited, and corrected with human input. Also included in Level 2 products are a variety of correction time series, including earth tides, laser frequency and vacuum pressure drift, and offsets. All Level 2 products are created by the Laser Strainmeter Analysis Center at UC San Diego and are archived in the same XML format as the borehole strainmeter data (see Appendix N for details). The XML files are versioned similarly to the XML files for the borehole strainmeter data products, and each laser strainmeter XML file contains one year of data.

There are two sublevels of Level 2 derived products. Level 2a products will be no more than 14 days out of date, and involve initial corrections to data available at that time. Level 2b products involve more extensive corrections, and are thus delayed by four months. The first Level 2a data products will not be available for two months following instrument installation, while the first Level 2b products will not be available until four months after the instrument is installed.

6.2.3.1 Level 2a: Initial Human-Corrected Products

Level 2a processing consists of editing, scaling, and correcting the 300-sec data streams from the Ice9 logger. The steps involved are as follows (see Figure 6.5 on page 46):

1. Unpack the 300-sec Ice9 data file and convert into niolib format.
2. Remove a tidal model determined using parameters from Level 2b processing
   The laser strainmeters do not have any drift in their record of earth tides, and thus a predictive model (based on parameters determined at Level 2b) can be subtracted from the observed data without loss of fidelity.
3. Clean the de-tided 300 second data manually
   Level 2a data cleaning and editing are done manually using information derived from the rapid automated data checking procedure, metadata such as information from the field engineers, instrument state of health information, and so on. At this stage, offsets and outliers are flagged and removed. The magnitude and time of any offsets or outliers removed are stored in the Level 2 XML file.
4. Subtract correction series from cleaned data
   The Analysis Center staff correct the cleaned data for the effects of laser frequency and vacuum pressure drift. The time series recorded from the optical anchors at each end monument will also be subtracted, to account for motion of the end-monuments. The correction and environmental series will be stored in their own XML file.
5. Add the earth tide model back into the cleaned and corrected strain series
6. Create XML file for cleaned, corrected, and re-tided series, including the scaling factor from data-logger units to linear strain in the XML file.

These products are sent to the strain archives as rapidly as possible after generation using LDM.

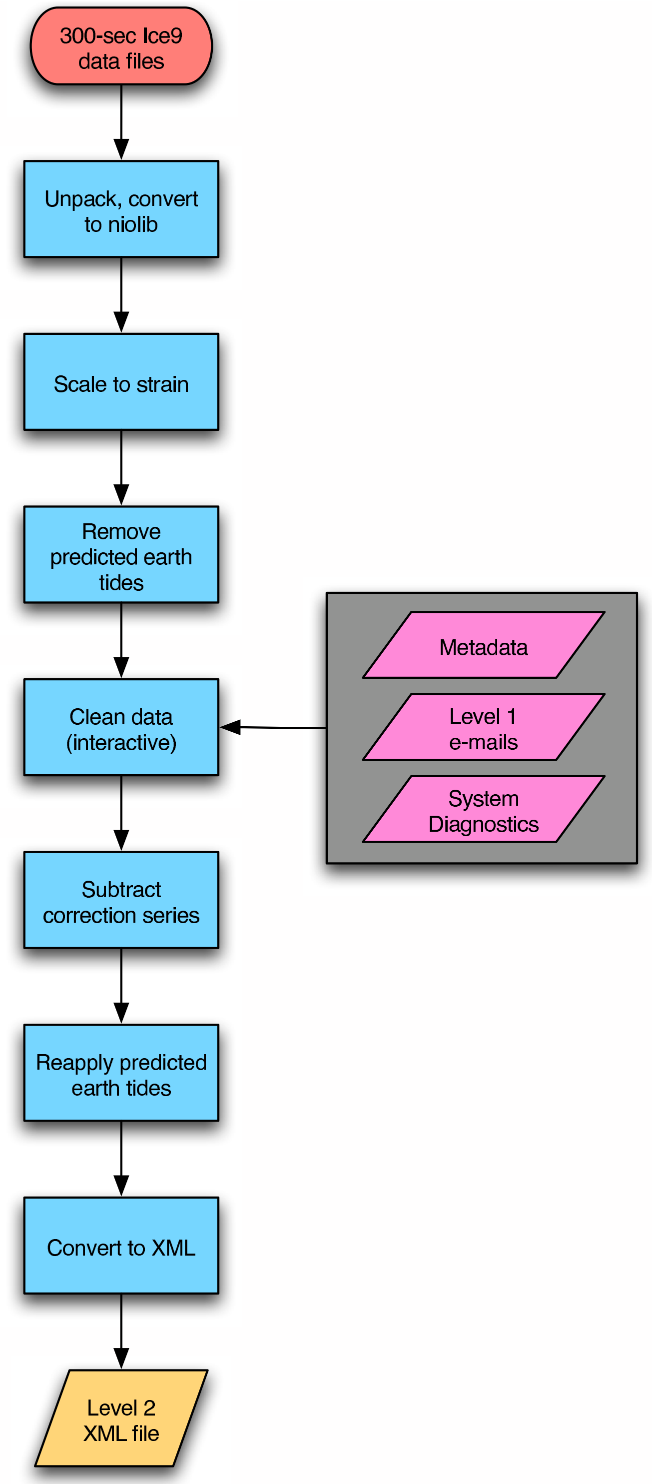

Figure 6.5: Level 2a processing steps for laser strainmeter data.

6.2.3.2 Level 2b: Final Human-Corrected Strainmeter Products

At four-month intervals, the entire 300-second interval data set will be reprocessed to form the Level 2b data set. The clean strain series will be regenerated by applying the edits found at Level 2a and any additional editing found necessary. This allows correction of long-term instrument problems not detected earlier.

In addition to re-doing the steps for Level 2a processing, the Level 2b processing includes

1. Monitoring the parameters used to derive the earth tide model
   Experience with previous laser strainmeters suggests that the earth tide model estimated for the first Level 2b processing run should be accurate for over a decade. However, the Laser Strainmeter Analysis Center will recompute these parameters every four months to find any possible significant changes, and will continue to monitor agreement between the data and the tidal model.

2. Reassess the outlier criterion
   The criteria for detecting outliers will be recalculated and updated as needed.

Level 2b products are stored in XML format and sent to the strain archives as rapidly as possible after generation using LDM.

6.2.4 Laser Strainmeter Data Archiving and Delivery

All PBO laser strainmeter data products will be archived at the PBO strain archives at the Northern California Earthquake Data Center and IRIS Data Management Center. The strain archives receive two 24-hour Ice9 files per day via LDM, one containing 1-sps data and the other 300-sec data. The archives also receive 12 channels of 1-sps miniSEED format data in a single 24-hour chunk per day, as well as 12 channels of 300-sec miniSEED data in another single 24-hour chunk per day; these data are sent using SEEDLink.

The archives store both Ice9 and miniSEED data, in order to guard against errors in data translation and to serve the widest range of users possible. Level 1 products are not archived separately, but instead the scaling factors necessary to convert from Level 0 to Level 1 will be stored in the POD and PBO strain SEED volumes. Level 2 products are delivered to the archives using LDM, no less often than once every 14 days, and are stored in the XML format defined in Appendix N. The archives will also hold a copy of the PBO public metadata related to strainmeter stations and data. These data are made available in the same fashion as the borehole strainmeter data products.

Appendix A

Glossary

N-sps	Data sampled N times per second
N-sec	Data sampled once every N seconds
BINEX	Binary Exchange Format, `http://binex.unavco.org`
BSM	Borehole Strainmeter
CGPS	Continuously-operating GPS
CMR	Trimble Compact Measurement Record
DGPS	Differential GPS
DPM	Data Products Manager
GAMIT	GPS at MIT, one of the three major GPS data processing packages (the others being Bernese and GIPSY)
GLOBK	GLOBal Kalman filter software (companion package to GAMIT)
GHAM	Global, Hierarchical, Alphanumeric, Morton-encoded; a geocode system developed by Duncan Agnew
GIPSY	GPS Inferred Positioning System, another major GPS processing package, developed by JPL
GPS	Global Positioning System
GSAC	GPS Seamless Archive System
GTSM	Gladwin Tensor Strainmeter
https	Secure HyperText Transfer Protocol, a secure web interface
IERS	International Earth Rotation and Reference Systems Service, `www.iers.org`
IGS	International GNSS Service
IRIS DMC	Incorporated Research Institutions for Seismology Data Management Center, one of the main archives of global seismological data
ITRF2000	International Terrestrial Reference Frame, 2000 realization
LDM	UNIDATA Local Data Manager software, `http://www.unidata.ucar.edu/packages/ldm/index.html`
LSM	Laser Strainmeter
PBO	Plate Boundary Observatory
PBOSC	PBO Standing Committee
POD	PBO Operational Database
RINEX	Receiver-Independent EXchange Format, `http://gauss.gge.unb.ca/RINEX/rinex2.txt`

RT17	Trimble RS-232 real-time stream format
RTCM	Radio Technical Commission for Maritime Services
RTK	Real-time Kinematic
SCIGN	Southern California Integrated GPS Network
scp	Secure Copy, a program for secure transfer of data using the Secure Shell (ssh) system
SEED	Standard for the Exchange of Earthquake Data, a standard seismic data format
sftp	Secure FTP, a secure version of the widely-used File Transfer Protocol (ftp)
SGPS	Survey-mode GPS
SINEX	Solution-Independent EXchange Format, `http://tau.fesg.tu-muenchen.de/~iers/web/sinex/format.php`
SIR	Site Installation Report, derived from the SCIGN Site Construction Report
SOPAC	Scripps Orbit and Permanent Array Center
SPOTL	Some Programs for Ocean-Tide Loading, a tide-modeling package from Duncan Agnew
SRR	Site Reconnaissance Report, derived from the SCIGN Site Evaluation Report
VSAT	Very Small Aperture Terminal, a satellite data-transfer system
WGS 84	World Geodetic System, 1984 version, a very widely-used standard geodetic coordinate system
XML	eXtensible Markup Language

Appendix B

GHAM Geocode Structure

The PBO station naming convention proposed here is based on the GHAM (Global, Hierarchical, Alphanumeric, Morton-encoded) geocode system developed by Duncan Agnew [*Agnew*, 2005]. GHAM geocodes can be used to uniquely identify PBO sites (provided the coding is done to a fine enough resolution), and have the distinct advantage that nearby sites tend to appear near each other in sorted lists considerably more often than in sorted lists of latitude and longitude. They are also considerably more compact representations of latitude and longitude for the same spatial resolution than simply writing out the coordinates in decimal degrees.

The following algorithm converts a latitude ϕ and longitude λ to a GHAM geocode:

1. Convert ϕ (in degrees from -90 to 90, with north positive) and λ (in degrees from -180 to 180, with east positive) into x and y coordinates in an equal-area cylindrical projection as:

$$x = (\lambda + 180)/360 \quad \text{(B.1)}$$

$$y = [\sin(\phi) + 1]/4 \quad \text{(B.2)}$$

   This gives $0 \leq x \leq 1$ and $0 \leq y \leq 0.5$.

2. Compute two new numbers, $r_x = 16x$ and $r_y = 16y$ and store the integer parts of r_x and r_y as 4-bit integers i_x and i_y, which run from 0 to 15.

3. Interleave the bits of i_x and i_y to form an 8-bit integer i_c, which will be in the range $0 \leq i_c \leq 255$.

4. Express i_c as $10N + M$, with M being between 0 and 9.

5. Encode this value as a letter-number pair using the $N + 1$-th letter of the Latin alphabet and the numbers 0–9, so that A0 encodes 0 and Z5 encodes 255.

6. Define $x' = r_x - i_x$ and $y' = r_y - i_y$, both of which are in the range 0–1, and substitute these values in step 2 above.

7. Repeat this process until the desired level of resolution is reached, storing each new letter-number pair at the end of a growing string.

By repeating this process as long as needed, any desired level of resolution can be achieved. For example, repeating this process four times gives an 8-character Level 4 GHAM geocode that represents an area of 230,400 square meters, while repeating it six times gives a geocode that represents an area of about 3.6 square meters. PBO station location geocodes will be Level 6 GHAM codes.

Appendix C

Naming Convention Details

This appendix gives more details on the PBO naming conventions, including details of the site, station, sensor, and channel identifiers, how files will be named, and how these conventions will be used.

C.1 Naming Hierarchy

PBO's naming hierarchy explicitly distinguishes between "sites" and "stations". In the PBO lexicon, a site is a parcel of land on which PBO holds a permit to install one or more stations, and is defined by the permit for that site, while a station is a collection of one or more sensors and a common reference point. There may be multiple stations at a given site, and therefore we want to distinguish between stations and sites. Sites have a single identifier: a descriptive name, usually based on some nearby geographic feature; Table C.1 describes this identifier in more detail.

PBO continuous GPS stations consist of the GPS monument and antenna, and optional met sensors; the receiver is not a sensor *per se*, but instead a datalogger. The reference point for a PBO continuous GPS station is the tip of the center-support screw on the D3 adaptor atop the GPS monument. The reference point for a PBO survey-mode GPS station would be the center punch on the disk, or a similar point on other types of marks. For borehole strainmeter stations, the reference point would be the center of the top of the wellhead. Each station has multiple identifiers, as described in Table C.1.

A given station may have multiple sensors, such as in the case of a borehole strainmeter station, where there are four individual borehole strainmeter sensors (one for each component), three borehole seismometer sensors, a pore pressure sensor, and so on. Each of these sensors may record data at multiple sample rates; for example, a GPS station may have channels at 15-sec, 1-sps, and 5-sps, or a borehole strainmeter's north-south sensor could have channels sampled at 20-sps, 10-sps, 1-sps, and so on. Sensors and channels have multiple identifiers, described in Table C.1.

Table C.1: Site/Station/Sensor/Channel Identifiers

	Site Identifiers	
Site Descriptive Name	Format	45 alphanumeric characters and underscores
	Source	Responsible PBO Regional Engineer
	Usage	Appears in POD and as a comment in raw data files
	Example	`Carson_Sink_Nevada`
	Station Identifiers	
Station Long Name	Format	60 alphanumeric characters and underscores
	Source	Based on Site Descriptive Name, with additional space to separate stations at a single site, and including installation dates to handle resets
	Usage	Appears in POD and as a comment in raw data files and derived data products
	Example	`Carson_Sink_Nevada_2005_CGPS`
Station Short Name	Format	16 alphanumeric characters: `ccccccccccrryyyy` `cccccccccc`: 10 alphanumeric character name string `rr`: 2-character region code `yyyy`: 4-digit station installation year
	Source	Responsible PBO Regional Engineer
	Usage	Primary station name in filenames and PBO data products
	Example	`CARSONSNK_NV2004`
Station Location Geocode	Format	12 alphanumeric characters: `CNCNCNCNCNCN` `C`: uppercase English letter `N`: single digit from 0–9
	Source	12-character Level 6 GHAM geocode (see Appendix B on page 50), computed from station precise latitude and longitude
	Usage	Secondary station identifier, appears in all PBO data products
	Example	`E4I8U3W2V7I3`, corresponding to 32.86772°N, -117.252331°E
Station Type code	Format	4 alphanumeric characters
	Source	Responsible PBO Regional Engineer
	Usage	Distinguishes co-located stations (such as a borehole strainmeter station with a GPS monument installed on the wellhead). Will appear in the POD and as comments in PBO raw data files and derived data products
	Example	`CGPS`: continuous GPS, `BSM_`: borehole strainmeter

Site/Station/Sensor/Channel Identifiers (cont'd)

IERS DOMES Number	Format	9 alphanumeric characters, AAANNMPPP AAA: 3-character "area code" NN: 2-digit station number, unique inside a given area M: indicates the PBO station reference point is a "ground mark" PPP: 3-digit sequential point number that allows for resets and the like
	Source	Assigned by IERS based on station location
	Usage	Only applicable for PBO GPS stations. Will appear in the POD, as a comment in PBO BINEX data files, and in the MARKER NUMBER field in PBO RINEX data files
	Example	40479M001, DOMES number for the SCIGN station BLYT
PBO Dot Number	Format	4 alphanumeric characters
	Source	PBO proposal
	Usage	Backward compatibility with current standards. Will appear in POD, PBO BINEX data files, and in the MARKER NAME field in PBO RINEX data files. Alternative filename identifier.
	Example	P041 for the station at Marshall Field, Colorado (which has the Short Name Marshall__CO2004 and the Geocode E507A7N2D4W4)

Sensor Identifiers

Sensor Type code	Format	24 alphanumeric characters and underscores
	Source	PBO
	Usage	Will appear in POD and as comment in PBO raw data files and derived data products.
	Example	Gladwin_TSM_1_component_, for the 1-component of a Gladwin Tensor Strainmeter

Site/Station/Sensor/Channel Identifiers (cont'd)

Channel Identifiers

Channel code	Format	`ssssssssssssssssssssssss+rrrrrrrr` `ssssssssssssssssssssssss`: Sensor Type Code described above, same for all channels from a particular sensor `rrrrrrrr`: 8-digit channel sample rate, specified in samples per 86400 seconds.
	Source	PBO
	Usage	Will appear in the POD and as a comment in PBO raw data files and derived data products.
	Example	`Gladwin_TSM_1_component_+01728000`, for the 1-component of a Gladwin Tensor Strainmeter, sampled at 20-sps.
SEED Code	Format	3-character codes in SEED-standard format.
	Source	SEED conventions
	Usage	Will appear in the POD and will be used in creating SEED-format files from data from PBO strainmeter stations.
	Example	`BSN` for a 20-sps channel from a borehole strainmeter's north-south sensor

Appendix D

GPS Data Analysis Plan

D.1 Introduction

As described in Section 5.1.4.3, PBO Level 2 GPS products are generated in two stages. Level 2a products are created by Analysis Centers (ACs) at Central Washington University and the New Mexico Institute of Mining and Technology (NMT); formerly the NMT AC was located at the University of California, Berkeley Seismological Laboratory (BSL). These ACs process GPS phase and pseudorange data to generate estimates of the coordinates of the stations averaged over the 24-hour duration of the data in the daily processing. The models used in this processing are given for each of the ACs in Tables D.3 and D.4. In these analyses, the GPS satellites orbits are fixed either to the orbits generated by the Jet Propulsion Laboratory (JPL) (CWU analysis) or the International GNSS Service (NMT). The CWU analysis also uses the JPL clock products.

Level 2b products are generated by the Analysis Center Coordinator using the Level 2a products from each AC. The Level 2a products are aligned to the PBO realization of the Stable North America Reference Frame (SNARF). In this frame, motions of stations on stable North America are due solely to glacial isostatic adjustment (GIA). The AC results are also combined and the combined analysis (denoted `pbo`) aligned to this same frame. The standard PBO position and velocity results are the combined, frame-realized product The Analysis Center ID (`AC_ID`) and Product ID (`PROD_ID`) are `pbo` and either `final_frame` or `rapid_frame` depending on the age of the data (see Appendix E for details). (Final products are not available until 6 to 13 days after data collection; rapid results are available with 1-day latency). Products are made available in SINEX format and a plain text files; see Appendices F–H for details on the format of these products.

D.2 Analysis Characteristics

The station coordinates estimated by the ACs are not referred to any particular reference frame; they can be rotated and translated (to some extent) to align them with a reference frame chosen for PBO products. The station coordinates in the NMT products are loosely constrained and the loose SINEX files have large standard deviations for the coordinates (approximately 0.5 meters). The covariance matrix for the NMT products has rotational uncertainty and because of the relative small aperture of the network

Table D.1: 4-character codes for stations used in the PBO daily frame definition

AGMT	BAMO	BBRY	BEMT	BEPK	BILI	BILL	BKAP	BKMS	BLW2	BMHL	BRMU
BSRY	BVPP	CAST	CAT1	CAT2	CCCC	CDMT	CEDA	CHMS	CIRX	CNPP	COPR
CORV	CRHS	CRRS	CRU1	CSCI	CSST	CTDM	CTMS	DRAO	DUBO	ECCO	ECFS
ECHO	EDPP	ELKO	EWPP	FERN	FGST	FLIN	FMVT	FOOT	FRED	GARL	GLRS
GMRC	GNPS	GOSH	GTRG	HCMN	HEBE	HNPS	HVYS	IID2	KELY	KOD1	KTBW
LAPC	LDES	LDSW	LEWI	LFRS	LKCP	LNMT	LORS	MAT2	MAUI	MDO1	MIG1
MKEA	MLFP	MPWD	MUSB	NBPS	NEWP	NHRG	NLIB	NRC1	OPBL	OPCL	OPCP
OPCX	OPRD	ORES	ORMT	P474	P562	P584	PABH	PBPP	PHLB	PPBF	PVRS
RAMT	RDMT	REDM	RSTP	RUBY	SBCC	SCOO	SCIA	SDHL	SEAT	SFDM	SHIN
SHLD	SHOS	SLMS	SMEL	SNI1	SOMT	SPIC	SPMS	SRS1	THCP	THU3	TOIY
TONO	TUNG	UPSA	USGC	VDCY	VIMT	VNCX	WGPP	WILL	WOMT	YELL	

Note: LDSW was not used between 2005-10-17 and 2005-12-03 due an unexplained anomalous offset (10 mm North

relatively large translational uncertainty of about 50 mm. The CWU loose SINEX files are determined by point positioning in which clocks and satellite orbits are assumed perfectly known, however the reference frame for the orbits in not well determined. The standard deviation of the coordinates in the CWU "loose" SINEX files are very small and reflect the uncertainty in a well defined reference frame. In the analysis of the CWU SINEX files, a 1-meter translation and 0.3-meter rotational uncertainty covariance matrix is added to the CWU SINEX file to covariance to allow the coordinates to rotate and translate into different reference frames.

The basic models used by both PBO ACs are consistent with IGS analysis standards with one major exception. NMT is using the adopted IGS values for the positions of the phase center of the transmission antenna on the satellite relative to the center of mass of the satellite. CWU is using the values adopted by JPL that differ from the IGS ones by 63 cm in the Z-component (radial) for Block IIR GPS satellites. This difference in the models introduces scale differences between the NMT and CWU solutions of about 2.5 parts-per-billion that needs to be accounted for when the NMT and CWU results combined. In the PBO product generation, scale is explicitly included in the estimated parameters, with different scale estimates for the NMT and CWU analyses.

In PBO analyses, rotations, translations and scales are estimated and the final values of these parameters set to best align the PBO daily and long-term results with the PBO realization of the SNARF reference frame. For PBO daily analyses, SNARF is realized by aligning with the positions of 130 stations in North America (127), Eastern Russia (1) and central Pacific (2). The distant stations serve to maintain the orientation of the daily reference frame realizations. The list of stations used in the daily realization is given in Table D.1. Their locations are shown in Figure D.1. The position and velocities of these stations are taken from the current PBO network velocity field estimate.

The PBO reference frame itself is determined from a Kalman filter combination of all finals SINEX files from NMT and CWU for the period between 1 Jan 2004 (GPS week 1251, day 4) and 27 August 2005 (GPS week 1337 day 6) in which the position and velocities of all stations were estimated. The velocity and positions of the 13 IGb00 reference stations listed in Table D.2 were used to align the frame with the SNARF reference frame. These locations are also shown in Figure D.1.

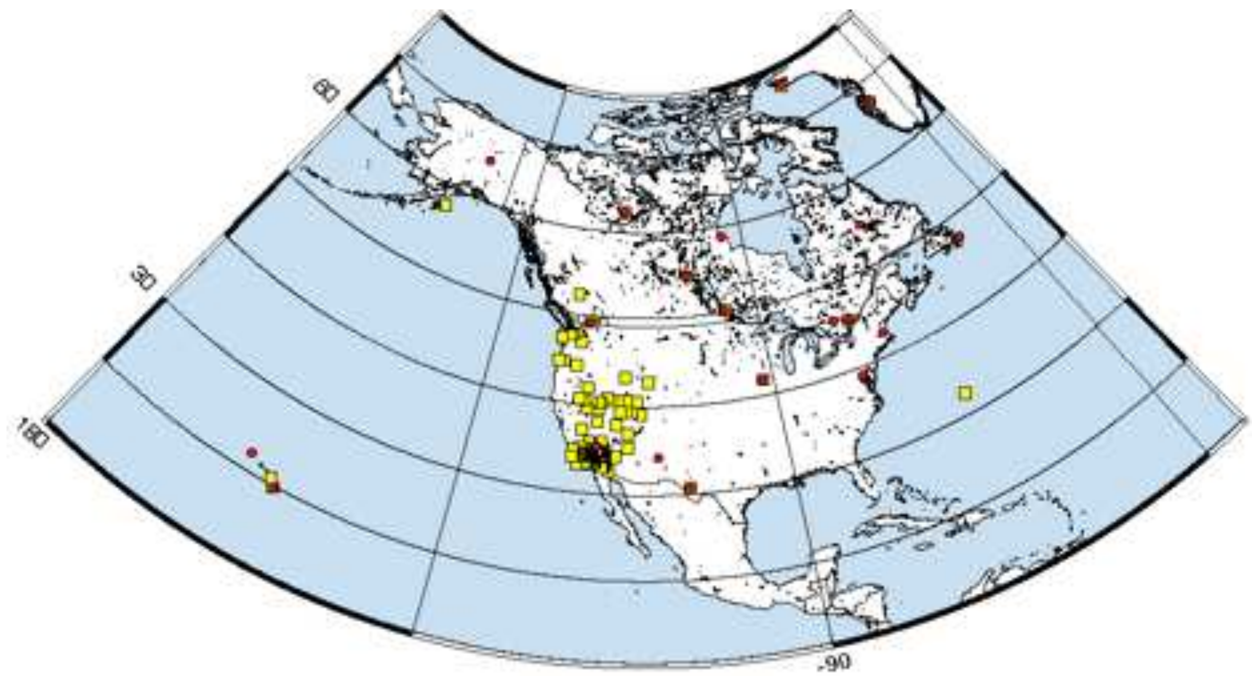

Figure D.1: Locations of PBO reference frame stations (yellow squares) and the IGb00 North America reference stations (red circles). Not shown on the plot is BILI (166.4 deg longitude, 68.1 deg latitude).

Table D.2: 4-character codes of the IGb00 stations used to align the PBO reference frame to the SNARF position and velocities

ALGO	BILI	CHUR	DRAO	DUBO	FLIN	GODE	GOLD	KELY	MDO1	MKEA	NLIB
NRC1	PIE1	SCH2	STJO	THU3	WES2	YELL					

In all the frame alignments, only rotations, translations and scale are estimated and therefore the internal geometry and evolution of the station positions is retained in the transformations. The current reference frame definitions will be retained until the next large re-analysis that is expected during 2006 when the IGS adopted new absolute antenna phase center models for both ground and space antennas.

D.3 AC SINEX file weighting

The NMT and CWU ACs use different models for the phase noise in the GPS processing and different decimation factors. These differences are accounted for multiplying the covariance matrices from the two center's SINEX by factors that make the chi-squared per degree of freedom fits to the PBO reference frame near unity (on a daily basis) for the two centers. The estimated factors derived by averaging results from the first 18 months of data analysis are 0.7 for NMT and 4.8 for CWU. When the results from the two centers are combined, these factors are doubled so that the standard deviations of the position estimates from the combined analysis are similar to those from the individual centers.

D.4 Product Generation Schedules

The PBO rapid analyses are performed with 1-day latency. The generation of these products takes place when the IGS rapid orbits become available. Generally, the analysis is completed near 0 hrs UT for the previous day of data (i.e., the products are ready 18 hrs EST). The final products are generated weekly when the IGS final orbit for the week becomes available. Usually, these orbits are ready by Friday each week for the GPS two weeks before. The latency of the products varies between 14 to 20 days depending on which day of the week the GPS data were collected.

In the event on an earthquake in the PBO area of sufficient size to displace stations by more than 1 mm (see Table 5.2 in Chapter 5), the Analysis Center Coordinator will product a set of estimated coseismic offsets for the affected stations using the last full day of data before the earthquake and first full day of data after the earthquake. There will be no position estimates for the affected stations on the day of the earthquake. The position-offset products will be generated from the rapid solution with a maximum 48-hour latency. Offsets will only be available for stations operating the day before and the day after the earthquake. See Appendix I for details on the format of these products.

D.5 Details of Analysis Center Processing Models and Methods

This section gives very detailed information on the processing methods used by each Analysis Center.

Note that the IGS has announced that FES2004 will be the official IGS ocean tidal loading model. The PBO Analysis Centers will adopt this model as soon as feasible.

```
Table D.3: Central Washington University Processing Strategy

================================================================================
                                 PLATE BOUNDARY OBSERVATORY
```

```
                      CWU Processing Specifications
=============================================================================
Analysis Center   | Central Washington University Geodesy Lab
                  |
                  | Dept. of Geological Sciences
                  | Central Washington University
                  | 400 University Way
| Ellensburg, WA 98926
                  | Phone:  ++ 1 509 963 2799
                  | Fax:    ++ 1 509 963 1109
-----------------------------------------------------------------------------
Contact Person(s) | Tim Melbourne       e-mail: tim@geology.cwu.edu
                  |                      phone : ++ 1 509 963 2799
                  | Marcelo Santillan    e-mail: marcelo@geology.cwu.edu
                  |                      phone : ++ 1 509 963 1108
                  | Craig Scrivner       e-mail: scrivner@geology.cwu.edu
                  |                      phone : ++ 1 509 963 1446
-----------------------------------------------------------------------------
Software Used     | Gipsy/Oasis R4 V2, developed at Jet Propulsion Laboratory
-----------------------------------------------------------------------------
Preparation Date  | September 30, 2005
-----------------------------------------------------------------------------
Modification Dates| None
-----------------------------------------------------------------------------
Effective Date for| September 30, 2005
Data Analysis     |
=============================================================================


-----------------------------------------------------------------------------
|                            MEASUREMENT MODELS                             |
|---------------------------------------------------------------------------|
| Observable         | RAPID: non-differenced estimates of ionosphere-free   |
|                    | combinations of L1 and L2 carrier phase (point-       |
|                    | positioning, see Zumberge, (1997).                    |
|                    |                                                       |
|                    | FINAL: Double-differenced, ionosphere-free combination |
|                    | L1, L2 carrier phases.  Pseudoranges are used only    |
|                    | to obtain receiver clock offsets and in ambiguity     |
|                    | resolution.                                           |
|---------------------------------------------------------------------------|
| Data weighting     | Sigma on doubly difference LC phase: Site and elevation|
|                    | dependent based on iterated analysis.                 |
|                    | Cleaning at 30-second rate.                           |
|                    | Sampling rate: 5 minutes                              |
```

```
|                   | Elevation angle cutoff : 15 degrees                        |
|----------------------------------------------------------------------------|
| Data Editing      | Cycles slips detected and fixed.                           |
|                   | Unresolved cycle slips estimated in solution.              |
|                   | Iterative editing using sequential slip deletion while     |
|                   | slips exceed RMS                                           |
|----------------------------------------------------------------------------|
| RHC phase         | Phase polarization effects applied (Wu et al, 1993)        |
| rotation corr.    |                                                            |
|----------------------------------------------------------------------------|
| Ground antenna    | Elevation-dependent phase center corrections are           |
| phase center      | applied according to the model IGS_01 and info.003.        |
| calibration       | These corrections
|                   | are given relative to the Dorne Margolin T antenna.        |
|----------------------------------------------------------------------------|
| Troposphere       | A priori zenith delay: nominal constant;  2-hour piece-|
|                   | wise linear function estimated, 1 NS and EW gradient       |
|                   | per day.                                                   |
|                   |--------------------------------------------------------|
|                   | Met data input: none                                       |
|                   |--------------------------------------------------------|
|                   | Mapping functions: (Niell, 1996)                           |
|----------------------------------------------------------------------------|
| Ionosphere        | Not modeled (ionosphere eliminated by forming the          |
|                   | ionosphere-free linear combination of L1 and L2).          |
|----------------------------------------------------------------------------|
| Plate motions     | ITRF2000 velocities                                        |
|----------------------------------------------------------------------------|
| Tidal             | Solid earth tidal displacement:                            |
| displacements     | constant Love number tides                                 |
|                   | frequency dependent radial tide (K1)                       |
|                   |--------------------------------------------------------|
|                   | Pole tide: Applied to Mean IERS pole position              |
|                   |--------------------------------------------------------|
|                   | Ocean loading: Applied (Scherneck Model)                   |
|----------------------------------------------------------------------------|
| Atmospheric       | Not applied                                                |
|  loading          |                                                            |
|----------------------------------------------------------------------------|
| Earth orientation| IERS Bulletin B plus diurnal and semidiurnal variations|
|                   | in x,y, and UT1  models (EOP)  R. Ray [1995], IERS         |
|                   | Tech. Note 21 [1996]                                       |
|----------------------------------------------------------------------------|
| Satellite center  | Sat. Phase centers hardwired into GIPSY R4. Values can     |
|                   | be found at #  \    |
```

```
|                    | satellites. Values used are:
|                    | Block I      x,y,z:  0.2100, 0.0000, 0.8540 m          |
| of mass            |---------------------------------------------------------|
| correction         | Block II/IIA x,y,z:  0.2790, 0.0000, 0.09679 m         |
|                    |---------------------------------------------------------|
|                    | Block IIR    x,y,z:  0.0580, -0.0680, -0.06300 m       |
|------------------------------------------------------------------------------|
| Satellite phase    | Not applied                                             |
| center calibrat    |                                                         |
--------------------------------------------------------------------------------
| Relativity         | Relativistic corrections applied                        |
| corrections        |                                                         |
--------------------------------------------------------------------------------
| GPS attitude       | Yaw computed using model of Bar-Sever (1996), using     |
| model              | nominal rates or estimates supplied by JPL              |
--------------------------------------------------------------------------------

--------------------------------------------------------------------------------
|                              ORBIT MODELS
| CWU USES STANDARD JPL RAPID AND FINAL SATELLITE ORBITS
|
|------------------------------------------------------------------------------|
| Geopotential       | EGM96 degree and order 9                                |
|                    |---------------------------------------------------------|
|                    | GM = 398600.4415 km**3/sec**2                           |
|                    |---------------------------------------------------------|
|                    | AE =   6378.1363 km                                     |
|------------------------------------------------------------------------------|
| Third-body         | Sun and Moon as point masses                            |
|                    |---------------------------------------------------------|
|                    | Ephemeris: CfA PEP NBODY 740                            |
|                    |---------------------------------------------------------|
|                    | GMsun  = 132712440000 km**3/sec**2                      |
|                    |---------------------------------------------------------|
|                    | GMmoon =    4902.7989 km**3/sec**2                      |
|------------------------------------------------------------------------------|
| Solar radiation    | A priori: nominal block-dependent constant direct       |
| pressure           |           acceleration; corrections to direct, y-axis,  |
|                    |           and B-axis constant and once-per-rev terms    |
|                    |           estimated                                     |
|                    |---------------------------------------------------------|
|                    | Earth shadow model: umbra and penumbra                  |
|                    |---------------------------------------------------------|
|                    | Earth's albedo: applied                             |
|                    |---------------------------------------------------------|
```

	Satellite attitude model not applied
Tidal forces	Solid earth tides: frequency independent Love number K2= 0.300
	Ocean tides: None
Relativity	Applied (IERS 1996, Chapter 11, Eqn.1)
Numerical Integration	Adams-Moulton fixed-step, 11-pt predictor-corrector with Nordsieck variable-step starting procedure (see Ash, 1972 and references therein)
	Integration step-size: 75 s; tabular interval: 900 s
	Arc length: 24 hours

ESTIMATED PARAMETERS (A PRIORI VALUES & SIGMAS)	
Adjustment	Square-root information filter with smoothing of terminal estimate backwards through time over all data
Station coordinates	16 networks of ~20 stations per network, dynamically updated daily as PBO expands. 2 common sites between networks. Weak constraints applied to site coordinates.
Satellite clocks bias	Standard JPL Satellite clocks error estimates are used
Receiver clock bias	Time estimated from pseudoranges.
Orbital parameters	Initial Position and Velocity (IC) plus 9 radiation-pressure terms: constant and sin/cos once-per-rev terms for a direct,y-axis, and b-axis acceleration. ICs estimated each day. Radiation parameters treated as random walk with process noise based on independent daily estimates. ICs fixed to IGS Final orbit values.
Troposphere	Piece-wise linear function in zenith delay estimated once per 2-hr for each station constrained by random walk process to 5e-8 km/sqrt(sec); one N-S and one

	E-W gradient parameter per day for each station
Ionospheric correction	Not estimated (first-order effect eliminated by linear combination of L1 and L2 phase)
Ambiguity	Resolution attempted for all baselines but resolving Melbourne-Webena Widelanes for L2-L1 using pseudo-ranges with differential code biases applied, and then L1 from geodetic solution using ionosperic free observable.
Earth Orient. Parameters (EOP)	Pole X/Y and their rates, and UT1 rate estimated once per day.
GPS attitude model	Not estimated

REFERENCE FRAMES	
Inertial	Geocentric; mean equator and equinox of 2000 Jan 1 at 12:00 (J2000.0)
Terrestrial	ITRF2000, No constrained sites coordinates.
Interconnection	Precession: IAU 1976
	Nutation: IAU 1980

REFERENCES:

Ash, M. E., Determination of Earth satellite orbits, Tech. Note 1972-5, Lincoln Laboratory, MIT, 19 April 1972.
Bar-Sever, Y. E., A new module for GPS yaw attitude, in Proc. IGS Workshop: Special Topics and New Directions, edit. G. Gendt and G. Dick, pp. 128-140, GeoForschungsZentrum, Potsdam, 1996.
Dong, D., and Y. Bock, Global Positioning System network analysis with phase ambiguity resolution applied to crustal deformation studies in California, Journal of Geophysical Research, 94, 3949-3966, 1989.
Dong, D., T. A. Herring, and R. W. King, Estimating Regional Deformation from a Combination of Space and Terrestrial Geodetic Data, J. Geodesy, 72, 200-214, 1998.

Niell, A. E., Global mapping functions for the atmospheric delay, J. Geophys. Res., 101, 3227-3246, 1996.
Ray, R.D., ftp://maia.usno.navy.mil/conventions/chapter8/ray.f (IERS Standards), 1995
Schaffrin, B., and Y. Bock, A unified scheme for processing GPS phase observations, Bulletin Geodesique, 62, 142-160, 1988.
Springer, T. A., G. Beutler, and M. Rothacher, A new solar radiation pressure model for the GPS satellites, IGS Analysis Center Workshop, Darmstadt, 9-11 February 1998.
Wu, J. T., S. C. Wu, G. A. Hajj, W. I. Bertiger, S. M. Lichten, Effects of antenna orientation on GPS carrier phase. Manuscripta Geodaetica 18, 1993, 91-98, 1993.
Zumberge, J.F., M.B. Heflin, D.C. Jefferson, M.M. Watkins, and F.H. Webb, Precise point positioning for the efficient and robust analysis of GPS data from large networks, Journal of Geophysical Research, B, Solid Earth and Planets, 102 (3), 5005-5017, 1997.

Table D.4: Berkeley Seismological Laboratory processing strategy

```
==============================================================================
PLATE BOUNDARY OBSERVATORY
BSL Processing Strategy Summary
==============================================================================
Analysis Center   | Berkeley Seismological Laboratory (BSL)
                  |
                  | 215 McCone Hall
                  | University of California
                  | Berkeley, CA 94720-4760
                  | Phone:  ++ 1 510 642 3977
                  | Fax:    ++ 1 510 643 5811
------------------------------------------------------------------------------
Contact Person(s) | Mark Murray          e-mail: murray@nmt.edu
                  |                      phone : ++ 1 505 835 6930
                  | Douglas Neuhauser    e-mail: doug@seismo.berkeley.edu
                  |                      phone : ++ 1 510 642 0931
                  | Barbara Romanowicz   e-mail: barbara@seismo.berkeley.edu
                  |                       phone : ++ 1 510 643 5690
------------------------------------------------------------------------------
Software Used     | GAMIT v. 10.2, GLOBK v. 10.2,   developed at MIT/SIO
------------------------------------------------------------------------------
Preparation Date  | September 23, 2005
------------------------------------------------------------------------------
Modification Dates| March 3, 2006
------------------------------------------------------------------------------
Effective Date for| September 23, 2005
```

Data Analysis |

===

MEASUREMENT MODELS	
Observable	Doubly differenced, ionosphere-free combination of L1 and L2 carrier phases. Pseudoranges are used only to obtain receiver clock offsets and in ambiguity resolution.
Data weighting	Sigma on doubly difference LC phase: Site and elevation dependent based on iterated analysis. Cleaning at 30-second rate. Sampling rate: 2 minutes Elevation angle cutoff : 10 degrees
Data Editing	Cycles slips detected and fixed. Unresolved cycle slips estimated in solution. Postfit editing using 4 times RMS deletion.
RHC phase rotation corr.	Phase polarization effects applied (Wu et al, 1993)
Ground antenna phase center calibrations	Elevation-dependent phase center corrections are applied according to the model IGS_01. The corrections are given relative to the Dorne Margolin T antenna.
Troposphere	A priori zenith delay: nominal constant; 2-hour piece-wise linear function estimated, 1 NS and EW gradient per day.
	Met data input: none
	Mapping functions: (Niell, 1996)
Ionosphere	Not modeled (ionosphere eliminated by forming the ionosphere-free linear combination of L1 and L2).
Plate motions	ITRF2000 velocities
Tidal displacements	Solid earth tidal displacement: constant Love number tides frequency dependent radial tide (K1)

	Pole tide: Applied to Mean IERS pole position
	Ocean loading: Applied (Scherneck Model)
Atmospheric loading	Not applied
Earth orientation	IERS Bulletin A plus diurnal and semidiurnal variations in x,y, and UT1 models (EOP) R. Ray [1995], IERS Tech. Note 21 [1996]
Satellite center of mass correction	Block I x,y,z: 0.2100, 0.0000, 0.8540 m
	Block II/IIA x,y,z: 0.2790, 0.0000, 1.0230 m
	Block IIR x,y,z: 0.0000, 0.0000, 0.0000 m
Satellite phase center calibrat	Not applied
Relativity corrections	Relativistic corrections applied
GPS attitude model	Yaw computed using model of Bar-Sever (1996), using nominal rates or estimates supplied by JPL

ORBIT MODELS	
Geopotential	EGM96 degree and order 9
	GM = 398600.4415 km**3/sec**2
	AE = 6378.1363 km
Third-body	Sun and Moon as point masses
	Ephemeris: CfA PEP NBODY 740
	GMsun = 132712440000 km**3/sec**2
	GMmoon = 4902.7989 km**3/sec**2
Solar radiation	A priori: nominal block-dependent constant direct

pressure	acceleration; corrections to direct, y-axis, and B-axis constant and once-per-rev terms estimated (Beutler et al., 1994; Springer et al. 1998)
	Earth shadow model: umbra and penumbra
	Earth's albedo: not applied
	Satellite attitude model not applied
Tidal forces	Solid earth tides: frequency independent Love number K2= 0.300
	Ocean tides: None
Relativity	Applied (IERS 1996, Chapter 11, Eqn.1)
Numerical Integration	Adams-Moulton fixed-step, 11-pt predictor-corrector with Nordsieck variable-step starting procedure (see Ash, 1972 and references therein)
	Integration step-size: 75 s; tabular interval: 900 s
	Arc length: 24 hours

ESTIMATED PARAMETERS (A PRIORI VALUES & SIGMAS)	
Adjustment	Weighted least squares plus Kalman filter
Station coordinates	11 networks of ~38 stations per network. 2 common sites between networks. Weak constraints applied to site coordinates.
Satellite clocks bias	Initial values from linear fit to Broadcast emphemeris. Values estimated during data cleaning.
Receiver clock bias	Time estimated from pseudoranges.
Orbital parameters	Initial Position and Velocity (IC) plus 9 radiation-pressure terms: constant and sin/cos once-per-rev terms for a direct,y-axis, and b-axis acceleration. ICs

	estimated each day. Radiation parameters treated as random walk with process noise based on independent daily estimates. ICs fixed to IGS Final orbit values.
Troposphere	Piece-wise linear function in zenith delay estimated once per 2-hr for each station constrained by a random-walk process to 20mm/sqrt(hr); one N-S and one E-W gradient parameter per day for each station, constrained to 10 mm at 10 deg elevation angle
Ionospheric correction	Not estimated (first-order effect eliminated by linear combination of L1 and L2 phase)
Ambiguity	Resolution attempted for all baselines but resolving Melbourne-Webena Widelines for L2-L1 using pseudo-ranges with differential code biases applied, and then L1 from geodetic solution using ionosperic free observable.
Earth Orient. Parameters (EOP)	Pole X/Y and their rates, and UT1 rate estimated once per day.
GPS attitude model	Not estimated

REFERENCE FRAMES	
Inertial	Geocentric; mean equator and equinox of 2000 Jan 1 at 12:00 (J2000.0)
Terrestrial	ITRF2000, No constrained sites coordinates.
Interconnection	Precession: IAU 1976
	Nutation: IAU 1980

REFERENCES:

Ash, M. E., Determination of Earth satellite orbits, Tech. Note 1972-5, Lincoln Laboratory, MIT, 19 April 1972.

Bar-Sever, Y. E., A new module for GPS yaw attitude, in Proc. IGS Workshop: Special Topics and New Directions, edit. G. Gendt and G. Dick, pp. 128-140,

GeoForschungsZentrum, Potsdam, 1996.
Beutler, G., E. Brockmann, W. Gurtner, U. Hugentobler, L. Mervart, and M. Rothacher, Extended Orbit Modeling Techniques at the CODE Processing Center of the International GPS Service for Geodynamics (IGS): Theory and Initial Results, Manuscripta Geodaetica, 19, 367-386, 1994.
Dong, D., and Y. Bock, Global Positioning System network analysis with phase ambiguity resolution applied to crustal deformation studies in California, Journal of Geophysical Research, 94, 3949-3966, 1989.
Dong, D., T. A. Herring, and R. W. King, Estimating Regional Deformation from a Combination of Space and Terrestrial Geodetic Data, J. Geodesy, 72, 200-214, 1998.
Niell, A. E., Global mapping functions for the atmospheric delay, J. Geophys. Res., 101, 3227-3246, 1996.
Ray, R.D., ftp://maia.usno.navy.mil/conventions/chapter8/ray.f (IERS Standards), 1995
Schaffrin, B., and Y. Bock, A unified scheme for processing GPS phase observations, Bulletin Geodesique, 62, 142-160, 1988.
Springer, T. A., G. Beutler, and M. Rothacher, A new solar radiation pressure model for the GPS satellites, IGS Analysis Center Workshop, Darmstadt, 9-11 February 1998.
Wu, J. T., S. C. Wu, G. A. Hajj, W. I. Bertiger, S. M. Lichten, Effects of antenna orientation on GPS carrier phase. Manuscripta Geodaetica 18, 1993, 91-98, 1993.

Appendix E

GPS Data Product File Naming Conventions

E.1 Common File Name Elements

E.2 SINEX File Names

PBO SINEX file names have the following structure:

```
<AC_ID><GPSWK><GPSDAY>.<PROD_ID>.snx
```

where the elements are as defined in Table E.1.

For example, the SINEX file derived from final combined solutions generated by the Analysis Center Coordinator for 22 July 2005 (GPS week and day are 1332 and 5, respectively) would be named `pbo13325.final_frame.rms`. The SINEX file derived from Berkeley's rapid processing for the same day, in Berkeley's own loosely-constrained reference frame, would be named `bsl13325.rapid_loose.rms`.

E.3 GPS Station Phase RMS File

Each GPS station phase RMS file generated by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

```
<AC_ID><GPSWK><GPSDAY>.<PROD_ID>.rms
```

where the elements are as defined in Table E.1.

For example, the GPS station phase RMS file derived from final combined solutions generated by the Analysis Center Coordinator for 22 July 2005 (GPS week and day are 1332 and 5, respectively) would be named `pbo13325.final_frame.rms`.

Table E.1: PBO GPS Product File Naming Common Elements

Element	Definition	Possible Values
`<AC_ID>`	Analysis Center identifier, 3-letter string that identifies which Analysis Center produced the given product	`BSL` (Berkeley Seismological Laboratory), `CWU` (Central Washington University), `PBO` (PBO combined solution, produced by the Analysis Center Coordinator at MIT).
`<GPSWK>`	GPS week number, 4 digits	0001 to XXXX
`<GPSDAY>`	GPS day number, 1 digit	0 (Sunday) through 6 (Saturday)
`<PROD_ID>`	Product type identifier. This is an 11-letter string that identifies what processing the product was derived from. This identifier allows users to distinguish between products derived from processing runs using rapid orbit products vs. final orbit products, as well as loose vs. tightly-constrained reference frames.	`rapid_loose` (product derived from processing runs using rapid orbit products, in loose reference frame); `rapid_frame` (rapid product, rotated into common PBO frame); `final_loose` (product derived from processing runs using final orbit products, in loose reference frame); `final_frame` (final product, rotated into common PBO frame); `rerun_loose` (product derived from reprocessing runs, in loose reference frame); `rerun_frame` (rerun product, rotated into common PBO frame)
`<STATION>`	4-character station name	Any valid 4-character station name.

E.4 GPS Position Time Series File

Each of the time series products created by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

```
<STATION>.<AC_ID>.<PROD_ID>.pos
```

where the elements are as defined in Table E.1.

For example, the time series for station P041 derived from final solutions generated by Central Washington University, and in the standard PBO reference frame, would be named `P041.CWU.final_loose.pos`. The time series for station P511 derived from the final combined solution produced by MIT would be named `P511.PBO.final_frame.pos`.

E.5 PBO GPS Velocity Field File

Each velocity field estimate generated by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

```
<AC_ID>.<PROD_ID>.vel
```

where the elements are as defined in Table E.1.

For example, the network velocity field derived from final combined solutions generated by the Analysis Center Coordinator would be named `PBO.final_frame.vel`.

E.6 PBO Coseismic Offset File

Each coseismic offset file created by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

```
<EQ_ID>.<PROD_ID>.off
```

where <PROD_ID> is defined in Table E.1, and <EQ_ID> is an event identifier based on the origin time of the earthquake for which the coseismic offsets have been estimated. The format for <EQ_ID> is `YYYYMMDDhhmmss`, where `YYYY`, `MM`, `DD`, `hh`, `mm`, `ss` are the year, month, day of month, hour, minute, and second (in UTC) of the origin time of the given event. The origin time considered definitive is that given by the National Earthquake Information Center.

For example, the file giving coseismic offsets estimated from the 28 September 2004 Parkfield earthquake, based on the “final” solutions rotated into the common PBO reference frame would be named `20040928171524.final_frame.off`.

Appendix F

GPS Station Phase RMS File Description

PBO's Analysis Center Coordinator generates an estimate of the phase RMS for each station for each solution based on the Level 2b products he creates. This chapter describes the file format for PBO GPS station phase rms files.

F.1 File Name Convention

Each GPS station phase RMS file generated by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

`<AC_ID><GPSWK><GPSDAY>.<PROD_ID>.rms`

where

- `<AC_ID>`
  A string that identifies the PBO Analysis Center from whose work the velocity field is derived. Values are one of `CWU` (Central Washington University), `BSL` (Berkeley Seismological Laboratory), or `PBO` (Combined solution from MIT).

- `<GPSWK><GPSDAY>`
  A string that gives the GPS week and day of week for which the RMS values are valid. `<GPSWK>` is a 4-digit number and `<GPSDAY>` is a 1-digit number from 0 to 6, where 0 represents Sunday and 6 represents Saturday.

- `<PROD_ID>`
  A string that identifies the product type. Values are
  `rapid_loose`: Rapid product from a given AC, in the loose reference frame used by that AC.
  `final_loose`: Final product from a given AC, in the loose reference frame used by that AC.
  `rerun_loose`: Rerun product from a given AC, in the loose reference frame used by that AC.
  `rapid_frame`: Rapid product from a given AC, rotated into a common reference frame.

`final_frame`: Final product from a given AC, rotated into a common reference frame.
`rerun_frame`: Rerun product from a given AC, rotated into a common reference frame.

For example, the GPS station phase RMS file derived from final combined solutions generated by the Analysis Center Coordinator for 22 July 2005 (GPS week and day are 1332 and 5, respectively) would be named `pbo13325.final_frame.rms`.

F.2 PBO GPS Station Phase RMS File Format

PBO GPS Station Phase RMS files are made available in ASCII format. The file begins with a set of header lines:

```
* RMS File from  ../BSL_snx/bsl12996.20041204.b.rms
* Created 20050929072528
```

where `RMS File from` gives the original source file and `Created` gives the creation date for the RMS file in `YYYYMMDDhhmmss` format, where `YYYY` is the 4-digit year, `MM` is the 2-digit month, `DD` is the 2-digit day of month, `hh` is the 2-digit hour, `mm` is the 2-digit minute, and `ss` is the 2-digit second of file generation.

Following the header are a series of lines giving phase RMS estimates for each station for the given day. These lines have the structure

```
*Site  COM #    RMS    CWU #    RMS     BSL #    RMS      A       B
*               mm              mm               mm      mm      mm
```

The entries are described in Table F.1.

F.3 Example phase RMS file

An example file would have the following internal structure

```
* RMS File from  ../BSL_snx/bsl12996.20041204.b.rms
* Created 20050929072528
*Site  COM #    RMS    CWU #    RMS     BSL #    RMS      A       B
*               mm              mm               mm      mm      mm
 AB07  24780    8.1        -     -      24780    8.1     2.7     3.2
 AB37  26860    6.0        -     -      26860    6.0     0.4     2.3
 AC59  24982    7.3        -     -      24982    7.3     2.3     2.9
 AC62  26713    7.1        -     -      26713    7.1     1.7     2.7
 AC63  26620    5.4        -     -      26620    5.4     1.3     2.1
 AC64  19950    9.5        -     -      19950    9.5     5.7     3.3
 AC65  23876    9.1        -     -      23876    9.1     5.0     3.0
```

Table F.1: PBO GPS Station Phase RMS File Header Format

Entry	Definition
`Site`	4-character PBO Dot Number for a given station
`Com #`	Average number of phase observations for the solutions from the two Analysis Centers, weighted by the processing frequency of the two groups. CWU uses five-minute sampling in their processing, while BSL uses 30-second sampling. Thus the CWU phase observations are multiplied by 10 in the averaging process. If only one Analysis Center has provided values, `Com #` is the value for that center.
`RMS`	Average phase RMS for the two solutions, again weighted by the processing frequency as in `Com #`.
`CWU #`	Number of phase observations reported by CWU, scaled from five-minute to 30-second sampling.
`RMS`	Phase RMS for CWU solutions
`BSL #`	Number of phase observations reported by BSL
`RMS`	Phase RMS for CWU solutions
`A, B`	Parameters for the elevation angle-dependent RMS model used by PBO. PBO uses $R^2(\theta) = A^2 + B^2/\sin(\theta^2)$, where R is the phase RMS and θ is the elevation angle from zero to 90 degrees (horizon to zenith).

and so on.

This file was named `pbo12996.final_frame.rms`, indicating it was produced from the combined solution produced by the Analysis Center Coordinator for GPS week 1299, day 6, and in the final PBO reference frame. Information from the CWU solution was lacking for this day, giving the `-` for each CWU entry.

Station `AC59` had 24982 phase RMS observations in the BSL solution, and since there was no CWU information, it also had 24982 phase RMS observations in the combined solution. The overall phase RMS for `AC59` on this day was 7.3 mm, and the elevation angle-dependent model for this day would be

$$R^2(\theta) = 2.3^2 + 2.9^2/\sin(\theta) \tag{F.1}$$

or

$$R^2(\theta) = 5.29 + 8.41/\sin(\theta). \tag{F.2}$$

Appendix G

GPS Position Time Series Format

This chapter describes version 1.0.1 of the PBO GPS station position time series file format.

G.1 File Name Convention

Each of the time series products created by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

`<STATION>.<AC_ID>.<PROD_ID>.pos`

where

- `<STATION>`
  4-character PBO Dot Number or 16-character PBO station name
- `<AC_ID>`
  A string that identifies the PBO Analysis Center from whose work the time series is derived. Values are one of `CWU` (Central Washington University), `BSL` (Berkeley Seismological Laboratory), or `PBO` (Combined solution from MIT).
- `<PROD_ID>`
  A string that identifies the product type. Values are
  `rapid_frame`: Rapid product from a given AC, rotated into a common reference frame.
  `final_frame`: Final product from a given AC, rotated into a common reference frame.
  `rerun_frame`: Rerun product from a given AC, rotated into a common reference frame.

For example, the time series for station P041 derived from final solutions generated by Central Washington University, and in the standard PBO reference frame, would be named `P041.CWU.final_loose.pos`. The time series for station P511 derived from the final combined solution produced by MIT would be named `P511.PBO.final_frame.pos`.

Table G.1: PBO GPS Station Position Time Series Header Format

Line	Entry	Definition
1	PBO Station Position Time Series	
2	Format Version: 1.0.1	
3	`4-character ID: SSSS`	4-character PBO Dot Number
4	`Station name: CCCCCCCCCCCCCCCC`	16-character station name
5	`First epoch: YYYYMMDD HHMMSS`	Date and time of first data point in the file, in UTC
6	`Last epoch:  YYYYMMDD HHMMSS`	Date and time of last data point in the file, in UTC
7	`Release date: YYYYMMDD HHMMSS`	Date and time at which the time series file was released, in UTC
8	`XYZ reference position: X Y Z`	Reference position of the station, in ITRF standard Cartesian coordinates with units of meters.
9	`NEU reference position: N E U`	Reference position of the station with respect to the WGS-84 standard ellipsoid. N is latitude in decimal degrees, with North positive. E is longitude in decimal degrees, with East positive. U is elevation with respect to the WGS-84 ellipsoid, in meters.

G.2 PBO GPS Station Position Time Series Format

PBO GPS station position time series are made available in ASCII format. The file begins with a set of header lines as given in Table G.1.

Following the header are a series of lines giving position information as a function of time. Each line gives the position estimate at a particular epoch. These lines have the structure

```
YYYYMMDD HHMMSS JJJJJJ X Y Z xx yy zz xy xz yz  N E U Ndel Edel Udel nn ee uu ne nu eu <quality>
```

The entries are described in Table G.2. Note that all times and dates given are in UTC.

G.3 Example GPS station position time series file

An example file derived from the Berkeley final solution rotated into the common PBO reference frame would be named `P041.BSL.final_frame.pos` and have the following internal structure

```
PBO Station Position Time Series
Format Version: 1.0.1
4-character ID: P041
Station name: Marshall_CO2004
First epoch: 20040120 000000
Last epoch:  20050930 000000
Release date: 20051001 000000
XYZ reference position: -2382952.20260893 -3688233.91924966 4610508.43527586
```

Table G.2: PBO GPS Station Position Time Series Format

Entry	Definition
YYYY	4-digit year for the given position epoch
MM	2-digit month of year for the given position epoch
DD	2-digit day of month for the given position epoch
HH	2-digit hour for the given position epoch
MM	2-digit minute for the given position epoch
SS	2-digit second for the given position epoch
JJJJJ	Modified Julian day for the given position epoch
X Y Z	ITRF Cartesian coordinates, meters
xx	Standard deviation of the X position, meters
yy	Standard deviation of the Y position, meters
zz	Standard deviation of the Z position, meters
xy	Correlation of the X and Y position
xz	Correlation of the X and Z position
yz	Correlation of the Y and Z position
N	North latitude, decimal degrees, relative to WGS-84 ellipsoid
E	East longitude, decimal degrees, relative to WGS-84 ellipsoid
U	Elevation, meters, relative to WGS-84 ellipsoid
Ndel	Change in North component relative to NEU reference position, meters. If the station moves northward, Ndel is positive.
Edel	Change in East component relative to NEU reference position, meters. If the station moves eastward, Ndel is positive.
Udel	Change in vertical component relative to NEU reference position, meters. If the station moves upward, Ndel is positive.
nn	Standard deviation of Ndel, meters
ee	Standard deviation of Edel, meters
uu	Standard deviation of Udel, meters
ne	Correlation of Ndel and Edel
nu	Correlation of Ndel and Udel
eu	Correlation of Edel and Udel
<quality>	'final' or 'rapid', corresponding to products generated from final or rapid orbit products

```
NEU reference position: 39.949481 -105.193900 1809.6159

20040120 000000 53024.0 -2382952.20260893 -3688233.91924966 4610508.43527586...
    0.01234 0.05678 0.09012 0.03456 0.07890 0.01234 ...
    39.949481 -105.193900 1809.6159 0.00000 0.00000 0.00000 0.01234 0.05678...
    0.09012 0.03456 0.07890 0.01234 final
20040121 000000 53025.0 -2382952.20260942 -3688233.91924927 4610508.43530291...
    0.01234 0.05678 0.09012 0.03456 0.07890 0.01234 ...
    39.949481 -105.193900 1809.6159 0.00000 0.00000 0.00000 0.01234 0.05678...
    0.09012 0.03456 0.07890 0.01234 final
```

and so on, where ... indicates that a single line has been broken for clarity.

Appendix H

GPS Network Velocity Field Format

This chapter describes the format of version 1.0.1 of PBO GPS network velocity field products.

H.1 File Name Convention

Each velocity field estimate generated by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

<AC_ID>.<PROD_ID>.vel

where

- <AC_ID>
  A string that identifies the PBO Analysis Center from whose work the velocity field is derived. Values are one of `CWU` (Central Washington University), `BSL` (Berkeley Seismological Laboratory), or `PBO` (Combined solution from MIT).

- <PROD_ID>
  A string that identifies the product type. Values are
  `rapid_loose`: Rapid product from a given AC, in the loose reference frame used by that AC.
  `final_loose`: Final product from a given AC, in the loose reference frame used by that AC.
  `rerun_loose`: Rerun product from a given AC, in the loose reference frame used by that AC.
  `rapid_frame`: Rapid product from a given AC, rotated into a common reference frame.
  `final_frame`: Final product from a given AC, rotated into a common reference frame.
  `rerun_frame`: Rerun product from a given AC, rotated into a common reference frame.

For example, the network velocity field derived from final combined solutions generated by the Analysis Center Coordinator would be named `PBO.final_frame.vel`.

H.2 PBO GPS Network Velocity Field Format

PBO GPS network velocity fields are made available in ASCII format. The file begins with a set of header lines:

```
PBO Network Velocity Field
Release date: YYYYMMDD hhmmss
```

where the release date is the UTC date and time at which the velocity field file was written. In this date, `YYYY` is the 4-digit year, `MM` is the 2-digit month, `DD` is the 2-digit day of month, `hh` is the 2-digit hour, `mm` is the 2-digit minute, and `ss` is the 2-digit second.

Following the header are a series of lines giving one or more velocity estimates for each station. Each velocity estimate is valid for a specified range of epochs and was estimated on a given date. These lines have the structure

```
Dot# station_name ref_epoch ref_jday ref_XYZ ref_NEU...
   X Y Z xx yy zz xy xz yz  N E U nn ee uu ne nu eu first_epoch last_epoch
```

where `...` indicates the line has been wrapped for ease of printing. The entries are described in Table H.1. Note that all times and dates given are in UTC.

A given station may have more than one velocity in a given velocity file, for example if the station is near a large earthquake and is affected by postseismic deformation. In this case, there will be two different velocity lines in the given file, with different `first_epoch`, `last_epoch`, and `valid_epoch`. A given velocity is valid from until a new `valid_epoch` is given for the same station.

H.3 Example GPS network velocity field file

An example file would have the following internal structure

```
PBO Network Velocity Field
Release date: 20051001 000000

P041 Marshall__CO2004 20040330,000000 53094 39.949481 -105.193900 1809.6159...
   0.0012 0.0034 -0.0056 0.0001 0.0023 0.0045 0.0067 0.0089 0.0001...
  -0.0023 0.0045 -0.0067 0.0002 0.0003 0.0004 0.0005 0.0006 0.0007...
  20040101,000000 20040517,235945
P511 CoxcombMtnCS2005 20050901,000000 53614 33.886935 -115.296100  273.3447...
   0.0023 0.0045 -0.0067 0.0012 0.0034 0.0056 0.0078 0.0090 0.0012...
  -0.0010 0.0029 -0.0038 0.0047 0.0056 0.0065 0.0074 0.0083 0.0092...
  20050801,000000 20050927,235945
P067 CleggRanchCS2004 20040715,000000 53201 35.551751 -121.002958  106.9845...
   0.0012 0.0034 -0.0056 0.0001 0.0023 0.0045 0.0067 0.0089 0.0001...
  -0.0023 0.0045 -0.0067 0.0002 0.0003 0.0004 0.0005 0.0006 0.0007...
  20040101,000000 20040927,235945
```

Table H.1: PBO GPS Network Velocity Field Format

Entry	Definition
Dot#	4-character PBO Dot Number for a given station
station_name	16-character station name
ref_epoch	Date and time at which the station position is as given in ref_XYZ and ref_NEU. Also the date and time for which the given velocity is first valid. Format is YYYYMMDD,hhmmss. YYYY is the 4-digit year, MM is the 2-digit month, DD is the 2-digit day of month, hh is the 2-digit hour, mm is the 2-digit minute, and ss is the 2-digit second.
ref_jday	Reference epoch, represented as Modified Julian day.
ref_XYZ	Reference position, in ITRF Cartesian coordinates, in meters
ref_NEU	Reference position, given as latitude, longitude and elevation relative to the WGS-84 ellipsoid. Latitude is in decimal degrees, North positive. Longitude is in decimal degrees, East positive. Elevation is in meters.
X	X component of station velocity, in meters/yr. Positive if the station moves in the positive X direction as defined for the ITRF standard Cartesian coordinate system.
Y	Y component of station velocity, in meters/yr. Positive if the station moves in the positive Y direction as defined for the ITRF standard Cartesian coordinate system.
Z	Z component of station velocity, in meters/yr. Positive if the station moves in the positive Z direction as defined for the ITRF standard Cartesian coordinate system.
xx	Standard deviation of X velocity, meters/yr
yy	Standard deviation of Y velocity, meters/yr
zz	Standard deviation of Z velocity, meters/yr
xy	Correlation of X and Y velocity, meters/yr
xz	Correlation of X and Z velocity, meters/yr
yz	Correlation of Y and Z velocity, meters/yr
N	North component of station velocity, meters/yr. Positive if the station moves northward.
E	East component of station velocity, meters/yr Positive if the station moves eastward.
U	Vertical component of station velocity, meters/yr Positive if the station moves downward.
nn	Standard deviation of North velocity, meters/yr
ee	Standard deviation of East velocity, meters/yr
uu	Standard deviation of vertical velocity, meters/yr
ne	Correlation of North and East velocity, meters/yr
nu	Correlation of North and vertical velocity, meters/yr
eu	Correlation of East and vertical velocity, meters/yr
first_epoch	Epoch of first data used to derive the station velocity, in the same format as ref_epoch.
last_epoch	Epoch of last data used to derive the station velocity, in the same format as ref_epoch.

```
P067 CleggRanchCS2004 20040929,000000 53201 35.551752 -121.002960  106.9852...
  -0.0008 0.0022  0.0022 0.0010 0.0018 0.0105 0.0102 0.0310 0.0508...
  +0.0058 -0.0032 0.0011 0.0002 0.0008 0.0114 0.0039 0.0206 0.0647...
  20050929,000000 20050915,235945
```

and so on, where ... indicates that a single line has been broken for clarity.

In this case, stations Marshall__CO2004 (P041) and CoxcombMtnCS2005 (P511) have only one velocity. Station CleggRanchCS2004 (P067) was near the 28 Sept 2004 Parkfield earthquake and has two different velocities, one determined from data prior to the event, and one determined from data following the event. These are given on separate lines.

Appendix I

PBO Coseismic Offset File Format

PBO estimates coseismic offsets for any event meeting the minimum criteria given in Table 5.2. These offsets will be released as rapidly as possible following a given event, in the format described in this chapter. This chapter describes the format for version 1.0 of PBO coseismic offset files.

I.1 File Name Convention

Each coseismic offset file created by the GPS Analysis Center Coordinator is identified with a unique file name, with the structure

<EQ_ID>.<PROD_ID>.off

where

- <EQ_ID>
  Identifier based on the origin time of the earthquake. Format is `YYYYMMDDhhmmss`, where `YYYY`, `MM`, `DD`, `hh`, `mm`, `ss` are the year, month, day of month, hour, minute, and second (in UTC) of the origin time of the given event. The origin time considered definitive is that given by the National Earthquake Information Center.

- <PROD_ID>
  A string that identifies the product type. Values are
  `rapid_frame`: Rapid product from a given AC, rotated into a common reference frame.
  `final_frame`: Final product from a given AC, rotated into a common reference frame.
  `rerun_frame`: Rerun product from a given AC, rotated into a common reference frame.

.

For example, the file giving coseismic offsets estimated from the 28 September 2004 Parkfield earthquake, based on the "final" solutions rotated into the common PBO reference frame would be named `20040928171524.final_frame.off`.

Table I.1: PBO Coseismic Offset File Header Format

Line	Entry	Definition
1	PBO Coseismic Offsets	
2	Format Version: 1.0	
3	Earthquake Origin Time (from NEIC):	NEIC-determined origin time for event, in UTC. Format is `YYYYMMDD hhmmss`
4	Earthquake Latitude (from NEIC):	NEIC-determined latitude for event, in decimal degrees with North positive.
5	Earthquake Longitude (from NEIC):	NEIC-determined longitude for event, in decimal degrees with East positive.
6	Earthquake Magnitude (Mw, from NEIC):	NEIC-determined moment magnitude for event, format M.M
7	Offsets release date:	UTC date on which offsets were released, in format `YYYYMMDD hhmmss`.

I.2 PBO Coseismic Offset Format

PBO coseismic offset files are made available in ASCII format. The file begins with a set of header lines as given in Table I.1.

Following the header are a series of lines giving coseismic offset information for each station. These lines have the structure

```
SSSS CCCCCCCCCCCCCCCC X Y Z x y z xx xy xz yy yz zz N E U n e u nn ne nu ee eu uu
```

The entries are described in Table I.2. Note that all times and dates given are in UTC.

I.3 Example coseismic offsets file

An example file for an event occurring at 18:15:35 UTC on 29 September 2005, with a moment magnitude of 7.5, latitude 5.660 degrees South, longitude 76.385 degrees West, and derived from final solution rotated into the common PBO reference frame would be named `20050929181535.final_frame.off` and have the following internal structure

```
PBO Coseismic Offsets
Format Version: 1.0
Earthquake Origin Time (from NEIC): 20050926 015537
Earthquake Latitude (from NEIC):     -5.660 N
Earthquake Longitude (from NEIC):   -76.385 E
Earthquake Magnitude (Mw, from NEIC): 7.5
Offsets release date: 20050929 181535
```

Table I.2: PBO Coseismic Offset File Header Format

Entry	Definition
SSSS	4-character PBO station identifier
CCCCCCCCCCCCCCCC	16-character PBO station identifier
X Y Z	Reference position in Cartesian coordinates, meters
x	Coseismic offset in X direction, meters
y	Coseismic offset in Y direction, meters
z	Coseismic offset in z direction, meters
xx	Standard deviation of the X offset, meters
yy	Standard deviation of the Y offset, meters
zz	Standard deviation of the Z offset, meters
xy	Correlation of the X and Y offset, meters
xz	Correlation of the X and Z offset, meters
yz	Correlation of the Y and Z offset, meters
N	Reference north latitude, decimal degrees, relative to WGS-84 ellipsoid
E	Reference east longitude, decimal degrees, relative to WGS-84 ellipsoid
U	Reference elevation, meters, relative to WGS-84 ellipsoid
n	Coseismic offset in N direction, meters
e	Coseismic offset in E direction, meters
u	Coseismic offset in U direction, meters
nn	Standard deviation of N offset, meters
ee	Standard deviation of E offset, meters
uu	Standard deviation of U offset, meters
ne	Correlation of N offset and E offset, meters
nu	Correlation of N offset and U offset, meters
eu	Correlation of E offset and U offset, meters

```
P041 Marshall__CO2004 -2382952.20260893 -3688233.91924966 4610508.43527586...
0.0013 0.0045 0.0009 0.0012 0.0034 0.0045 0.0067 0.0089 0.0001 ...
39.949481 -105.193900 1809.6159 0.0013 0.0045 0.0009 0.0012 0.0034 0.0045 ...
0.0067 0.0089 0.0001
```

and so on, where ... indicates that a single line has been broken for clarity.

Appendix J

Borehole Strainmeter Calibration and Correction

J.1 Borehole Strainmeter Calibration

PBO borehole strainmeters will be calibrated using the following equation:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{g_1} & 0 & 0 & 0 \\ 0 & \frac{1}{g_2} & 0 & 0 \\ 0 & 0 & \frac{1}{g_3} & 0 \\ 0 & 0 & 0 & \frac{1}{g_4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2\cos 2\theta_1} & \frac{1}{2\sin 2\theta_1} \\ \frac{1}{2} & \frac{1}{2\cos 2\theta_2} & \frac{1}{2\sin 2\theta_2} \\ \frac{1}{2} & \frac{1}{2\cos 2\theta_3} & \frac{1}{2\sin 2\theta_3} \\ \frac{1}{2} & \frac{1}{2\cos 2\theta_4} & \frac{1}{2\sin 2\theta_4} \end{bmatrix} \begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix} \begin{bmatrix} t_1 & t_1 & t_1 \\ t_2 & t_2 & t_2 \\ t_3 & t_3 & t_3 \end{bmatrix} \begin{bmatrix} e_a \\ \gamma_1 \\ \gamma_2 \end{bmatrix}^{regional} \tag{J.1}$$

where u_i are the gauge readings, g_i the gauge weightings, θ_i the gauge orientations, c and d the areal and shear strain scale factors, respectively, and t_i the corrections for the topography and geology. The regional strain is described by e_a, γ_1, and γ_2. Initially the topography and geology matrix will be set to the identity matrix (Equation J.1 is then equivalent to Equation B4 of *Hart et al.* [1996]).

The c and d values and relative gauge weightings g_2/g_1, g_3/g_1, and g_4/g_1 will be solved for by comparing the real and imaginary components of the M2 and O1 tides as observed on each gauge with the regional M2 and O1 strain tides predicted using theoretical models. The parameters will be found via a least squares inversion of Equation J.1. If the least squares solution gives unrealistic areal and shear strain factors (following parameter ranges described in *Gladwin and Hart* [1985] then Equation J.1 will be solved in an iterative fashion using starting values of 1 for the weightings and 1.5 and 3 for c and d

respectively. Once the values are found the strain measured by the instrument is calculated using

$$\begin{bmatrix} e_a \\ \gamma_1 \\ \gamma_2 \end{bmatrix}^{regional} = \begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2} & \frac{1}{2\cos 2\theta_1} & \frac{1}{2\sin 2\theta_1} \\ \frac{1}{2} & \frac{1}{2\cos 2\theta_2} & \frac{1}{2\sin 2\theta_2} \\ \frac{1}{2} & \frac{1}{2\cos 2\theta_3} & \frac{1}{2\sin 2\theta_3} \\ \frac{1}{2} & \frac{1}{2\cos 2\theta_4} & \frac{1}{2\sin 2\theta_4} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{g_1} & 0 & 0 & 0 \\ 0 & \frac{1}{g_2} & 0 & 0 \\ 0 & 0 & \frac{1}{g_3} & 0 \\ 0 & 0 & 0 & \frac{1}{g_4} \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \tag{J.2}$$

The SPOTL tidal programs [*Agnew*, 1996] will be used to predict the amplitude and phases of the M2 and O1 strain tides. The Gutenberg-Bullen Model A Green functions shall be used along with the CSR 3.0 global ocean model that has been derived from TOPEX/POSEIDON altimeter data [*Egbert and Erofeeva*, 2002]. Global ocean models often do not accurately calculate loading where there are large local tides, such as the San Francisco Bay Area. Where they exist, local ocean models will be included in the calculations.

If any finite element modeling is performed to quantify perturbations in the strain field caused by local geology and topography this information will be included when calculating the regional strain.

For the first six months of operation, it may be difficult to extract the tidal signal if the borehole relaxation and grout curing rates are high. Therefore, for the first six months the gauge data will be combined into areal and shear strain using gauge weightings of 1 and values of $c = 1.5$ and $d = 3.0$. A tidal calibration will be provided six months after the instrument has been installed.

J.2 Tidal Analysis

J.2.1 Estimation of M2 and O1 phase and amplitude from strain gauge data

At Level 2b the observed tides will be computed for the last three months of data analyzed from each gauge. The tidal analysis will be performed using the program BAYTAP-G [*Tamura et al.*, 1991] and PIASD [*Agnew*, 2004]. The program outputs a list of phases, amplitudes and associated error for a specified set of tidal frequencies. The data will have been detrended and had all offsets removed for input to the tidal analysis program.

J.2.2 Calculation of the tidal correction time series

The tidal time series corrections are computed from the tidal amplitudes and phases as measured by the strainmeter. Once the amplitude and phase of the M2 and O1 tides are known the `hartid` prediction program within the SPOTL software package will be used to generate the tidal correction tine series. The prediction program uses the amplitude of the observed M2 and O1 tide to scale and sum the deformation associated with 18 long-period, 51 diurnal and 50 semidiurnal constituents. The amplitudes of the additional constituents are found by spline interpolation of the ratio of the potential amplitude to that observed. The time series of the tidal deformation will be calculated, at any sample interval, for the areal and shear strain.

J.3 Borehole relaxation and grout curing trends

The dominant trend in strainmeter data is that of grout curing and borehole recovery. Corrections for these effects need to be generated for each gauge as the trends will vary with orientation. Factors that contribute to the trends include the compressive stress exerted by the grout expanding as it cures, dissipation of heat into the surrounding rock, and the recovery of the borehole. The general equation to describe the combined trends is

$$y = F + A_1 \exp(T_1 t) + A_2 \exp(T_2 t) + mt, \tag{J.3}$$

where F is a constant offset, m describes the linear trend which will dominate as time increases, and T_1 and T_2 are time constants. An approximate model for the borehole effects will be found through forward modeling of Equation J.3. When a good approximate model has been found, the best fitting model will be determined using a least squares procedure.

Appendix K

Bottle File Format and Archiving Structures

K.1 Bottle Format

PBO BSM dataloggers collect strain data in bottle files. Bottle files are binary files that consist of a 40-byte header followed by the strain data. If the data logger fails to make a measurement at a particular time that data point is given the value 999999. Bottle files are therefore continuous in time. All bottles collected by the data logger are in little endian format.

K.2 Archived Bottle Files

The raw bottle files downloaded from the GTSM21 data logger are archived at the IRIS Data Management Center and the Northern California Earthquake Data Center. Other than tarring the files for transfer via LDM, no changes are made to the bottle files.

Table K.1: Bottle File Header Structure

Bytes	Data Type	Description
0–1	short	magic number (15D Hex)
2–3	short	unused
4–7	integer	size of header (40 bytes)
8–15	double	start time, expressed as seconds since 1 Jan 1970
16–19	float	sample interval in seconds
20–23	integer	number of data points in the file
24–27	integer	data type: short, integer, float
28–31	integer	invalid data point, 999999
32–35	integer	unused
36–39	integer	bottle identifier

The archives store three tar files for each station for each day of the year:

- `SSSSYYDDDDay.tgz`
  10-minute strain data and diagnostic data
- `SSSSYYDDD_01.tar`
  1 sample-per-second (sps) strain data
- `SSSSYYDDD_20.tar`
  20-sps strain data,

where `SSSS` is the four-character strainmeter code, `YY` represents the last two digits of the year, and `DDD` represents the day of the year. For example, the three tar files from strainmeter B004, day 261 of 2005 would be `B00405261Day.tgz`, `B00405261_01.tar`, and `B00405261_20.tar`.

K.2.1 Day File: SSSSYYDDDDay.tgz

The Day file contains 10-minutes interval strain measurements from each of the four strain gages plus lower sample rate diagnostic data. Table K.2 lists the contents of a day file from a borehole strainmeter B004, year 2005 and day 261.

K.2.2 1-sps tar file: SSSSYYDDD_01.tar

The `SSSSYYDDD_01.tar` file contains 1-sps strain measurements from each of the four strain gages within the strainmeter. The tar file contains 24 `SSSSYYDDDHH.tgz` files where HH represents the hour. There is one file for each hour of the day. Each `SSSSYYDDDHH.tgz` file contains 4 bottle files:

- `SSSSYYDDDHHCH0`
- `SSSSYYDDDHHCH1`
- `SSSSYYDDDHHCH2`
- `SSSSYYDDDHHCH3`

The 1-sps tar file therefore contains 96 individual bottle files.

K.2.3 The 20-sps tar file: SSSSYYDDD_20.tar

The `SSSSYYDDD_20.tar` file contains 20-sps strain measurements from each of the four strain gages within the strainmeter. Within the 20-sps tar file there are 24 individual tar files, one for each hour of the day (`HH`), with the filename format `SSSSYYDDDHH_20.tar`. Each of the `SSSSYYDDDHH_20.tar` contains 60 gzipped tar files, one for each minute of the day (`MM`). The filename format for each of the minute tgz files is `SSSSYYDDDHHMM_20.tgz`. Within each `SSSSYYDDDHHMM_20.tgz` file there are four 1-minute long, 20-sps bottle files. The bottle filenames are of the format `SSSSYYDDDHHMMCHX_20`, where `X` is the channel number. The 20-sps tar file therefore contains 5760 individual bottle files.

Table K.2: Bottle File Header Structure

File	Description	Sample interval (mins)
B00405261BatteryVolts	Battery Voltage	30
B00405261CH0	Channel 0 strain	10
B00405261CH1	Channel 1 strain	10
B00405261CH2	Channel 2 strain	10
B00405261CH3	Channel 3 strain	10
B00405261CalOffsetCH0G1	Calibration Offset Channel 0 Gain1	60
B00405261CalOffsetCH0G2	Calibration Offset Channel 0 Gain2	60
B00405261CalOffsetCH0G3	Calibration Offset Channel 0 Gain3	60
B00405261CalOffsetCH1G1	Calibration Offset Channel 1 Gain1	60
B00405261CalOffsetCH1G2	Calibration Offset Channel 1 Gain2	60
B00405261CalOffsetCH1G3	Calibration Offset Channel 1 Gain3	60
B00405261CalOffsetCH2G1	Calibration Offset Channel 2 Gain1	60
B00405261CalOffsetCH2G2	Calibration Offset Channel 2 Gain2	60
B00405261CalOffsetCH2G3	Calibration Offset Channel 2 Gain3	60
B00405261CalOffsetCH3G1	Calibration Offset Channel 3 Gain1	60
B00405261CalOffsetCH3G2	Calibration Offset Channel 3 Gain2	60
B00405261CalOffsetCH3G3	Calibration Offset Channel 3 Gain3	60
B00405261CalStepCH0G1	Calibration Step Channel 0 Gain1	60
B00405261CalStepCH0G2	Calibration Step Channel 0 Gain2	60
B00405261CalStepCH0G3	Calibration Step Channel 0 Gain2	60
B00405261CalStepCH1G1	Calibration Step Channel 1 Gain1	60
B00405261CalStepCH1G2	Calibration Step Channel 1 Gain2	60
B00405261CalStepCH1G3	Calibration Step Channel 1 Gain2	60
B00405261CalStepCH2G1	Calibration Step Channel 2 Gain1	60
B00405261CalStepCH2G2	Calibration Step Channel 2 Gain2	60
B00405261CalStepCH2G3	Calibration Step Channel 2 Gain2	60
B00405261CalStepCH3G1	Calibration Step Channel 3 Gain1	60
B00405261CalStepCH3G2	Calibration Step Channel 3 Gain2	60
B00405261CalStepCH3G3	Calibration Step Channel 3 Gain2	60
B00405261DownholeDegC	Down hole temperature	30
B00405261LoggerDegC	Logger Temperature	30
B00405261PowerBoxDegC	Power Box Temperature	30
B00405261PressureKPa	Atmospheric Pressure	30
B00405261RTSettingCH0	Ratio Transformer Channel 0 Calibration	60
B00405261RTSettingCH1	Ratio Transformer Channel 1 Calibration	60
B00405261RTSettingCH2	Ratio Transformer Channel 2 Calibration	60
B00405261RTSettingCH3	Ratio Transformer Channel 3 Calibration	60
B00405261Rainfallmm	Rainfall mm	30
B00405261SolarAmps	Solar Amps	30
B00405261SystemAmps	System Amps	30

Appendix L

Ice-9 Data File Naming and Format Description

This chapter is an adaptation of the general Ice-9 format document provided by Larry Beck (Ice-9 Software) and Duncan C. Agnew (UC San Diego) and describes the naming convention and internal format structure for the Ice-9 raw data files provided by PBO laser strainmeter stations.

L.1 Ice-9 File Name Convention

The names for PBO Ice-9 files have the following structure:

- `YYYYDDD.SSSS.1s.gz`
  A gzipped file of binary data from station `SSSS`, recorded at 1-second intervals, and sampled every second, containing data for year `YYYY` and day of year `DDD`. For example, `2005362.DHL2.1s.gz` contains data from station `DHL2`, sampled once per second, for day `362` of year `2005`.

- `YYYYDDD.CCCC.5m.gz`
  A gzipped file of binary data from station `SSSS`, recorded at 1-second intervals, and sampled every five minutes, containing data for year `YYYY` and day of year `DDD`. For example, `2005362.DHL2.5m.gz` contains data from station `DHL2`, sampled once every five minutes, for day `362` of year `2005`.

L.2 Ice-9 File Format

L.2.1 Header Records

PBO Ice-9 files begin with a 128-byte descriptive header, containing file-specific information, labels to allow cross-checking of the source of the file, and version and system identification information. All header information is stored in Little Endian (PC) byte order. Table L.1 describes the header format.

Table L.1: PBO Ice-9 Header Format

Byte Offset		Stored Information
Start	Stop	
1	4	Binary file version number, which identifies the binary file format.
5	8	Header size in bytes.
9	12	End-of-file marker; data at or past this location is not valid.
13	16	Last sample marker, gives the position in bytes of the start of the last valid sample record.
17	20	4-character system ID, used to identify the source of the data.
21	24	Sample interval in seconds.
25	128	Padding bytes.

Table L.2: PBO Ice-9 Data Record Header Format

Byte Offset		Stored Information
Start	Stop	
1	4	UNIX UTC time (seconds since 1 January 1970 00:00 UTC, ignoring leap seconds).
5	6	Year
7	8	Day of year (1-365, or 366 in leap years).
9	10	Hour
11	12	Minute
13	14	Second
15	16	Number of channels of data in record

L.2.2 Data Records

Each data record consists of a 16-byte record header followed by the data. The record header information and data are stored in Big Endian (Sun/UNIX) byte order.

Table L.2 describes the format of the data record header.

The header is followed by N 2-byte integer values, where N is the number of channels in the record header. The data are stored in counts, with the range ± 10 Volts being -32768 to 32767 counts; the values are stored as 2's-complement 16-bit integers.

Appendix M

PBO SEED Codes and Conventions

As described in Chapter 6, PBO strain data are made available in both the raw logger format (bottle or Ice-9) and SEED format. This chapter outlines the SEED codes used to store and access PBO strainmeter data in SEED format. Armed with these codes, a user may extract PBO strainmeter data using the standard SEED tools available from either the IRIS Data Management Center or Northern California Earthquake Center.

The PBO Network Code for all strainmeter and seismic data is **PB**.

The SEED Codes for the first five PBO strainmeters are given in Table M.1.

Table M.1: PBO Site Codes

Site Code	16-character station name	Location
	Borehole Strainmeter Stations	
B001	Golbeck01BWA2005	Sequim, Washington
B004	HokoFallsBWA2005	Hoko Falls, Washington
P403	FloeQuaryBWA2005	Snider, Washington
	Laser Strainmeter Stations	
DHL2	DurmidHillCS2005	Durmid Hill, California
GVS1	GlendallSMCS2005	Glendale, California

The channel and location codes are given in Tables M.2 and M.3 for borehole and strainmeter stations, respectively.

Table M.2: PBO Borehole Strainmeter SEED Channel and Location Codes

SEED Codes		Sample Rate	Description
Channel	Location	(samples/sec)	
BS1	T0	20	GTSM Strain Channel 1, sampled at 20 samples/sec
LS1	T0	1	GTSM Strain Channel 1, sampled at 1 samples/sec
RS1	T0	0.001667	GTSM Strain Channel 1, sampled once every 600 sec
BS2	T0	20	GTSM Strain Channel 2, sampled at 20 samples/sec
LS2	T0	1	GTSM Strain Channel 2, sampled at 1 samples/sec
RS2	T0	0.001667	GTSM Strain Channel 2, sampled once every 600 sec
BS3	T0	20	GTSM Strain Channel 3, sampled at 20 samples/sec
LS3	T0	1	GTSM Strain Channel 3, sampled at 1 samples/sec
RS3	T0	0.001667	GTSM Strain Channel 3, sampled once every 600 sec
BS4	T0	20	GTSM Strain Channel 4, sampled at 20 samples/sec
LS4	T0	1	GTSM Strain Channel 4, sampled at 1 samples/sec
RS4	T0	0.001667	GTSM Strain Channel 4, sampled once every 600 sec
VDD	TP	0.1	Downhole pore pressure
VKD	TP	0.1	Downhole temperature recorded at the pore pressure sensor, approximately 100 meters deep in the borehole
LDO	TS	1	Atmospheric pressure
RDO	TS	0.0005556	Atmospheric pressure
RRO	TS	0.0005556	Rainfall
RK1	T0	0.0005556	Logger temperature
RKD	T0	0.0005556	Downhole temperature, measured by a thermistor inside the GTSM sonde
REO	T0	0.0005556	Solar amps
RE1	T0	0.0005556	Battery voltage
RK2	T0	0.0005556	Power box temp
RE2	T0	0.0005556	System amps

Note: There are also 24 calibration channels available under channel codes RCA, RCB, RCC, and RCD with location codes T1 through T6. Please contact Greg Anderson, anderson@unavco.org, or Kathleen Hodgkinson, hodgkinson@unavco.org, for more details.

Table M.3: PBO Laser Strainmeter SEED Channel and Location Codes

SEED Codes		Sample Rate	Description
Channel	Location	(samples/sec)	
LDV	LI	1	Vacuum Pressure
RDV	LI	0.00333	Vacuum Pressure
LS1	LM	1	Laser strain
RS1	LI	0.00333	Laser strain
RS1	LM	0.00333	Laser strain
LX1	LI	1	Correction series from optical anchor at interferometer end
RX1	LI	0.00333	Correction series from optical anchor at interferometer end
LX2	LR	1	Correction series from optical anchor at retroreflector end
RX2	LR	0.00333	Correction series from optical anchor at retroreflector end
LE1	LI	1	Voltage reference channel 1
RE1	LI	0.00333	Voltage reference channel 1
LKI	LI	1	Room temperature at interferometer end
RKI	LI	0.00333	Room temperature at interferometer end
LK2	LI	1	Box temperature at interferometer end
RK2	LI	0.00333	Box temperature at interferometer end
LKI	LR	1	Room temperature at retroreflector end
RKI	LR	0.00333	Room temperature at retroreflector end
LK3	LR	1	Box temperature at retroreflector end
RK3	LR	0.00333	Box temperature at retroreflector end
LKO	LV	1	Air temperature
RKO	LV	0.00333	Air temperature
LKD	LV	1	Ground temperature
RKD	LV	0.00333	Ground temperature
LDO	LV	1	Barometric pressure
RDO	LV	0.00333	Barometric pressure
LUO	LV	1	Light intensity
RUO	LV	0.00333	Light intensity
LRO	LI	1	Rainfall
RRO	LI	0.00333	Rainfall
LE2	LI	1	Voltage reference channel 2
RE2	LI	0.00333	Voltage reference channel 2
LX3	LI	1	Correction series from backup optical anchor at interferometer end
RX3	LI	0.00333	Correction series from backup optical anchor at interferometer end
LX4	LR	1	Correction series from backup optical anchor at retroreflector end
RX4	LR	0.00333	Correction series from backup optical anchor at retroreflector end

Appendix N

PBO XML Format For Strain Data

This chapter describes PBO's eXtensible Markup Language (XML) structure for storing Level 2 strainmeter data products, format version 1.0.

N.1 Overview

Processed PBO borehole strainmeter (BSM) and laser strainmeter (LSM) data are available in eXtensible Markup Language (XML) format. The strainmeter XML format consists of two distinct sections, an instrument information section and a data section. The instrument information section contains the following sub-sections: station information, sensor information and processing information. The data section contains the strain measurements and time series corrections. The structure can be summarized as follows.

```
<strain_xml >
   <inst_info>
      <station_information>
      </station_information>
      <sensor_information>
      </sensor_information>
      <processing>
      </processing>
   </inst_info>
   <data>
   </data>
</strain_xml>
```

N.2 Instrument Information

N.2.1 Station Information

The station information section contains basic information about the site that is not expected to change. A user can view this part of the instrument information section to see quickly if it is the data set they are interested in. The following is an example of station information for a BSM at Hoko Falls. The elements are summarized in Table N.1.

```
<station_information>
   <site_name>HokoFalls</site_name>
   <station_long>hokofalls_b004_makah_fishhatchery_sekiu_washington_2005_bsm</station_long>
   <station_short>HokoFallsBWA2005</station_short>
   <dot_number>B004</dot_number>
   <geocode>E4T1U2F6J0Q9</geocode>
   <itype>BSM_</itype>
   <model>GTSM21</model>
   <institution>PBO_UNAVCO</institution>
   <region>WA</region>
   <install_date>2005-06-15</install_date>
   <coordinate kind="station" ellipsoid="WGS84">
      <lat>48.201925</lat>
      <long>-124.427006</long>
      <height units="m">30</height>
   </coordinate>
   <minicluster_stations> </minicluster_stations>
   <time_zone>UTC</time_zone>
   <SEED>
      <site>B004</site>
      <network>PB</network>
   </SEED>
</station_information>
```

Table N.1: XML Station Information Elements

Element	Description
<site_name>	Site descriptive geographic name.
<station_long>	Station long name, up to 60 characters long.
<station_short>	Station's short name, a 16-character code.
<dot_number>	PBO's four-character code that identifies the strainmeter.
<geocode>	Level-6 12-character GHAM code.
<itype>	4-character station type code: LSM_ (Laser Strainmeter) or BSM_ (Borehole Strainmeter)
<model>	Strainmeter model e.g., GTSM
<institution>	Institution responsible for instrument e.g., PBO
<region>	PBO region 2-charcter code
<install_date>	Instrument installation date (YYYY-MM-DD).
<coordinate>	There are four kinds of coordinates: station — center point of borehole strainmeter lsm_midpoint — mid-point of laser strainmeter lsm_retro — laser retro reflector coordinate lsm_interf — laser interferometer coordinate The ellipsoid used is WGS84.
<lat>	Latitude (decimal degrees, North positive).
<long>	Longitude (decimal degrees, East positive).
<height units="m">	Height (m)
<minicluster_stations>	Contains the 4 character codes of other strainmeters in a minicluster.
<time_zone>	Time zone of dates and time in file.
<SEED>	Strainmeter SEED code (http://www.iris.edu/manuals/SEED_chpt1.htm.)
<site>	Four-character SEED site code, same as the PBO dot number.
<network>	Two-character SEED network code (PB for all PBO products).

N.2.2 Sensor Information

The sensor information section contains information about the sensors from which measurements have been used to create the processed data. For the BSMs, there is at least one sensor response element for each gage and the barometric pressure sensor. For the laser strainmeters there is at least one sensor response element for the strain measurement and for each of the optical anchors. Any time the sensor responses change a new sensor response element is created. The spans of time for which the responses are valid are given by the start and end time within the element. The sensor response element contents are summarized in Table N.2.

The following is an example of a sensor response element for one gage within a Gladwin Tensor Strainmeter (GTSM) BSM.

```
<sensor_response>
   <sensor_start_date>2005-06-15T00:00:00</sensor_start_date>
   <sensor_end_date>Present</sensor_end_date>
   <sensor_type>Gladwin_BSM_component_1_</sensor_type>
   <sensor_code>E4T1U2F6J0Q9+BSM_+Gladwin_TSM_component_1_</sensor_code>
   <channel_code>Gladwin_BSM_component_1_+00086400</channel_code>
   <depth units="m">166.116</depth>
   <orientation direction="east_of_north" units="degrees">168.2</orientation>
   <sensor_volts_per_unit>0.0</sensor_volts_per_unit>
   <digitizer_counts_per_volt>0</digitizer_counts_per_volt>
   <scalefactor_units_per_count unit="nstrain"
        kind="manufacturer">0.1</scalefactor_units_per_count>
   <voltage_input_to_logger></voltage_input_to_logger>
   <assigned_logger_bits>0</assigned_logger_bits>
   <datalogger_manufacturer>GTSM Technologies</datalogger_manufacturer>
   <datalogger_serial_number>PBO-05-000</datalogger_serial_number>
   <datalogger_model>GTSM21</datalogger_model>
   <sensor_serial_number>PBO-05-000</sensor_serial_number>
   <response>
         <real_poles> </real_poles>
         <imag_poles> </imag_poles>
         <real_zeros> </real_zeros>
         <imag_zeroes> </imag_zeroes>
   </response>
</sensor_response>
```

Table N.2: XML Sensor Response Elements

Element	Description
<sensor_start_date>	Date at which the sensor responses became valid (YYYY-MM-DDTHH:MM:SS).
<sensor_end_date>	Date at which the sensor responses ceased to be valid. (YYYY-MM-DDTHH:MM:SS). The term ”present” indicates the information is still valid.
<sensor_type>	24-character descriptive sensor code.
<sensor_code>	Station location geocode plus station type plus the descriptive sensor code.
<channel_code>	The sensor type plus the number of samples in a 24-hour period.
<depth units="m">	Depth of instrument (m).
<orientation direction="east_of_north" units="degrees">	Orientation (degrees East of North).
<sensor_volts_per_unit>	Number of volts per unit count.
<digitizer_counts_per_volt>	Number of counts per volt.
<scalefactor_units_per_count unit="nstrain" kind="manufacturer">	Scale factor supplied by the manufacturer. Determined from lab calibrations.
<voltage_input_to_logger>	Voltage range input to the datalogger.
<assigned_logger_bits>	Number of assigned logger bits.
<datalogger_manufacturer>	Datalogger manufacturer.
<datalogger_serial_number>	Datalogger serial number.
<datalogger_model>	Datalogger model.
<sensor_serial_number>	Sensor serial number.
<response>	Description of the poles and zeroes
<real_poles>	Real poles.
<imag_poles>	Imaginary poles.
<real_zeros>	Real zeroes.
<imag_zeroes>	Imaginary zeroes.

N.2.3 Processing Information

The information required to derive the areal and shear strain from the gage data for the BSMs and linear strain from the LSMs are given in the processing section. Because different steps are taken to process LSM and BSM data this section of the XML file has different elements for each instrument type.

The processing information section is divided into blocks of processing history information describing how the data were processed between start and end dates. There may be more than one processing history element within the processing section. For example, if between 1st June 2005 and 1st July 2006 the BSM data were processed using a certain set of scale factors, and then from 1 July 2006 onwards the same data were reprocessed with a different set of scale factors, there would be two processing history elements; one describing the first set of scale factors and the second describing the second set.

Within each processing history element there are `lsm_processing` or `bsm_processing` elements. These elements describe how the data where processed between a start and end time within the time series. There may be one or more of these processing elements within one processing history element. The following is an example of the processing section for a borehole strainmeter.

```
<processing>
   <bsm_processing_history>
      <processing_history_start>2005-06-15T00:00:00</processing_history_start>
      <processing_history_end>2006-12-15T00:00:00</processing_history_end>
      <bsm_processing>
         <timeseries_start_date>2005-06-15T00:00:00</timeseries_start_date>
         <timeseries_end_date>2006-03-15T00:00:00</timeseries_end_date>
         (processing information)
      </bsm_processing>
      <bsm_processing>
         <timeseries_start_date>2006-03-15T00:00:00</timeseries_start_date>
         <timeseries_end_date>2006-12-15T00:00:00t</timeseries_end_date>
         (processing information)
      </bsm_processing>
   </bsm_processing_history>
   <bsm_processing_history>
      <processing_history_start>2006-12-15T00:00:00</processing_history_start>
      <processing_history_end>present</processing_history_end>
      <bsm_processing>
         <timeseries_start_date>2005-06-15T00:00:00</timeseries_start_date>
         <timeseries_end_date>2006-03-15T00:00:00</timeseries_end_date>
         (processing information)
      </bsm_processing>
      <bsm_processing>
         <timeseries_start_date>2006-03-15T00:00:00</timeseries_start_date>
         <timeseries_end_date>present</timeseries_end_date>
         (processing information)
      </bsm_processing>
   </bsm_processing_history>
```

```
</processing>
```

N.2.3.1 BSM Processing Information

The following is an example of the processing information contained within one `bsm_processing` element. The elements are summarized in Table N.3.


```
<bsm_processing>
   <timeseries_start_date>2005-06-15T00:00:00</timeseries_start_date>
   <timeseries_end_date>present</timeseries_end_date>
   <linearization>
      <linear_date>2005-06-29T00:00:00</linear_date>
      <g0_value>0.48391551</g0_value>
      <g1_value>0.49872537</g1_value>
      <g2_value>0.49840455</g2_value>
      <g3_value>0.49541471</g3_value>
   </linearization>
   <gage_weightings>
      <gw0>1</gw0>
      <gw1>1</gw1>
      <gw2>1</gw2>
      <gw3>1</gw3>
   </gage_weightings>
   <orientation_matrix>
      <o11>0.5</o11>
      <o12>0.45818</o12>
      <o13>-0.200177</o13>
      <o21> 0.5 </o21>
      <o22>-0.402449</o22>
      <o23>-0.296706</o23>
      <o31>0.5</o31>
      <o32>-0.0557338</o32>
      <o33>0.496884</o33>
      <o41>0.5 </o41>
      <o42>0.401409</o42>
      <o43>0.298112</o43>
   </orientation_matrix>
   <areal_and_shear_factors>
      <c_areal>1.5</c_areal>
      <d_shear>3.0</d_shear>
   </areal_and_shear_factors>
   <topography></topography>
   <atm_pressure unit="microstrain_per_millibar">
      <apc_g0>-0.00187</apc_g0>
      <apc_g1>-0.00188</apc_g1>
```

```
    <apc_g2>-0.00265</apc_g2>
    <apc_g3>-0.00211</apc_g3>
  </atm_pressure>
  <tidal_parameters>
    <units phase="degrees" potential="local" lag="negative" amp="nanostrain"/>
    <tide name="M2" period_hours="12.42" doodson="2 0 0 0 0 0">
      <phz kind="gage0">117.0</phz>
      <amp kind="gage0">10.9</amp
    </tide>
    <tide name="O1" period_hours="25.82" doodson="1 -1 0 0 0 0 ">
      <phz kind="gage0">50.6</phz>
      <amp kind="gage0">8.6</amp>
    </tide>
    ...(repeat for each of the four gages)...
  </tidal_parameters>
  <detrend_g0 units="microstrain" model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">
    <F>-4.4</F>
    <A1>-6.0034</A1>
    <T1>-.54857</T1>
    <M>-0.42909</M>
    <A2>8.6</A2>
    <T2>.0115</T2>
  </detrend_g0>
  <detrend_g1 units="microstrain" model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">
    <F>-5.28055</F>
    <A1>-4.673</A1>
    <T1>-.409</T1>
    <M>-0.020709</M>
    <A2>6.9725</A2>
    <T2>-0.017637</T2>
  </detrend_g1>
  <detrend_g2 units="microstrain"   model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">
    <F>4.64</F>
    <A1>-7.6548</A1>
    <T1>-0.40383</T1>
    <M>0.059574</M>
    <A2>-7.6533</A2>
    <T2>-0.027437</T2>
  </detrend_g2>
  <detrend_g3 units="microstrain"  model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">
    <F>-1.0278</F>
    <A1>3.3555</A1>
    <T1>-0.060154</T1>
    <M>-0.037414</M>
    <A2>0</A2>
```

```
      <T2>0</T2>
   </detrend_g3>
   <decimate_by_two_filter kind="acausal" number="30">
      0.0983262 0.2977611 0.4086973 0.3138961 0.0494246
      -0.1507778 -0.1123764 0.0376576 0.0996838 0.0154992
      -0.0666489 -0.0346632 0.0322767 0.0399294 -0.0097461
      -0.0341585 -0.0039241 0.0246776 0.0099725 -0.0157879
      -0.0099098 0.0078510 0.0081126 -0.0026986 -0.0061424
      0.0007108 0.0039659 -0.0006209 -0.0017117 0.0007240
   </decimate_by_two_filter>
   <decimate_by_three_filter kind="acausal" number="23">
      0.0373766 0.1165151 0.2385729 0.3083302 0.2887327
      0.1597948 0.0058244 -0.0973639 -0.1051034 -0.0358455
      0.0359044 0.0632477 0.0302351 -0.0168856 -0.0356758
      -0.0190635 0.0126188 0.0159705 0.0082144 -0.0087978
      -0.0037289 -0.0017068 0.0028335
   </decimate_by_three_filter>
   <decimate_by_five_filter kind="acausal" number="34">
      0.0218528 0.0458359 0.0908603 0.1359777 0.1830881
      0.1993418 0.1957624 0.1561194 0.0994146 0.0346412
      -0.0236544 -0.0580081 -0.0703257 -0.0555546
      -0.0287709 0.0032613 0.0267938 0.0358952 0.0311186
      0.0134283 -0.0028524 -0.0170042 -0.0176765
      -0.0123123 -0.0036798 0.0057730 0.0059817 0.0083501
      0.0000581 0.0005724 -0.0033127 0.0004411 -0.0030766
      0.0016604
   </decimate_by_five_filter>
</bsm_processing>
```

Table N.3: XML BSM Processing Elements

Element	Description
<timeseries_start_date>	Date within the time series at which the processing information becomes valid. (YYYY-MM-DDTHH:MM:SS)
<timeseries_end_date>	Date within the time series at which the processing information ceases to be valid (YYYY-MM-DDTHH:MM:SS). The term ”present” indicates the information is still valid.
<linearization>	Information on how the gage readings are converted to linear strain.
<linear_date>	Date of reference point for linearization.
<g0_value>	Gage-0 value used to linearize gage 0 data.
<g1_value>	Gage-1 value used to linearize gage 1 data.
<g2_value>	Gage-2 value used to linearize gage 2 data.
<g3_value>	Gage-3 value used to linearize gage 3 data.
<gage_weightings> <gw0> <gw1> <gw2> <gw3>	Gage weights used to combine gage data into areal and shear strain. The gage weighting matrix is: $$\begin{bmatrix} g_{w0} & 0 & 0 & 0 \\ 0 & g_{w1} & 0 & 0 \\ 0 & 0 & g_{w2} & 0 \\ 0 & 0 & 0 & g_{w3} \end{bmatrix}$$
<orientation_matrix> <o11> <o12> <o13> <o21> <o22> <o23> <o31> <o32> <o33> <o41> <o42> <o43>	The gage orientation matrix used to combine gage data into areal and shear strain. The orientation matrix is: $$\begin{bmatrix} o_{11} & o_{12} & o_{13} \\ o_{21} & o_{22} & o_{23} \\ o_{31} & o_{32} & o_{33} \\ o_{41} & o_{42} & o_{43} \end{bmatrix}$$
<areal_and_shear_factors> <c_areal> <d_areal>	Scale factors used to create areal and shear strain. The values form the coupling matrix: $$\begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix},$$ where c is the `c_areal` value and d is the `d_shear` value.

XML BSM Processing Elements (cont'd)

Element	Description
`<topography>`	Matrix describing the topographic effects.
`<atm_pressure unit="microstrain_per_millibar">`	Atmospheric Pressure Response coefficients. The atmospheric pressure correction for each gage is calculated by multiplying the pressure by each gage factor.
`<apc_g0>`	Response coefficient for gage0.
`<apc_g1>`	Response coefficient for gage1.
`<apc_g2>`	Response coefficient for gage2.
`<apc_g3>`	Response coefficient for gage3.
`<tidal_parameters>`	Parameters used to generate the tidal correction.
`<units phase="degrees" potential="local" lag="negative" amp="nanostrain"/>`	Describes the units and convention of the tidal calculations.
`<tide name="M2" period_hours="12.42" doodson="2 0 0 0 0 0">`	Contains the name, period in hours and Doodson numbers for each tide described. There is no upper limit on the number of tides described.
`<phz kind="gage0">`	The phase of the tide with respect to the local potential. There are four kinds for the BSMs: gage0, gage1, gage2, gage3.
`<amp kind="gage0">`	The amplitude of the tide. There are four kinds for the BSMs: gage0, gage1, gage2, gage3.
`<detrend_g0 units="microstrain" model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">`	Borehole curing and detrending information for gage0. These elements contain the parameters used to generate the borehole effects. The model is the equation used to generate the correction. The variable "t" is time in days since installation.
`<F>`	F in the borehole model equation.
`<A1>`	A1 in the borehole model equation.
`<T1>`	T1 in the borehole model equation.
`<M>`	M in the borehole model equation.
`<A2>`	A2 in the borehole model equation.
`<T2>`	T2 in the borehole model equation.
`<detrend_g1 units="microstrain" model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">`	Borehole curing and detrending information for gage1.
`<detrend_g2 units="microstrain" model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">`	Borehole curing and detrending information for gage2.

XML BSM Processing Elements (cont'd)

Element	Description
`<detrend_g3 units="microstrain" model="F+A1*exp(T1*t)+M*t+A2*exp(T2*t)">`	Borehole curing and detrending information for gage3.
`<decimate_by_two_filter kind="acausal" number="30">`	Filter weights for the decimate by two filter used to reduce the data from 1 sps to 300 second interval.
`<decimate_by_three_filter kind="acausal" number="23">`	Filter weights for the decimate by three filter used to reduce the data from 1 sps to 300 second interval.
`<decimate_by_five_filter kind="acausal" number="34">`	Filter weights for the decimate by five filter used to reduce the data from 1 sps to 300 second interval.

N.2.3.2 LSM Processing Information

The following is an example of the processing information contained within one `lsm_processing` element. Elements not described in Table N.3 are summarized in Table N.4.

```
<lsm_processing>
   <timeseries_start_date>2005-06-01T00:00:00</timeseries_start_date>
   <timeseries_end_date>Present</timeseries_end_date>
   <linear_strain factor="strain_per_count">1.95500E-10</linear_strain>
   <tidal_parameters>
      <units phase="degrees" potential="local" lag="negative" amp="nanostrain"/>
      <tide name="M2" period_hours="12.42" doodson="2 0 0 0 0 0">
         <phz kind="lsm">0.0</phz>
         <amp kind="lsm">0.0</amp>
      </tide>
      <tide name="O1" period_hours="25.82" doodson="1 -1 0 0 0 0 ">
         <phz kind="lsm">0.0</phz>
         <amp kind="lsm">0.0</amp>
      </tide>
   </tidal_parameters>
   <optical_anchor_retroreflector_scalefactor>
      1.0
   </optical_anchor_retroreflector_scalefactor>
   <optical_anchor_interferometer_scalefactor>
      1.0
   </optical_anchor_interferometer_scalefactor>
   <laser_frequency_correction_scalefactor>
      1.0
```

```
   </laser_frequency_correction_scalefactor>
   <decimate_by_two_filter kind="acausal" number="30">
      0.0983262 0.2977611 0.4086973 0.3138961 0.0494246
      -0.1507778 -0.1123764 0.0376576 0.0996838 0.0154992
      -0.0666489 -0.0346632 0.0322767 0.0399294 -0.0097461
      -0.0341585 -0.0039241 0.0246776 0.0099725 -0.0157879
      -0.0099098 0.0078510 0.0081126 -0.0026986 -0.0061424
      0.0007108 0.0039659 -0.0006209 -0.0017117 0.0007240
   </decimate_by_two_filter>
   <decimate_by_three_filter kind="acausal" number="23">
      0.0373766 0.1165151 0.2385729 0.3083302 0.2887327
      0.1597948 0.0058244 -0.0973639 -0.1051034 -0.0358455
      0.0359044 0.0632477 0.0302351 -0.0168856 -0.0356758
      -0.0190635 0.0126188 0.0159705 0.0082144 -0.0087978
      -0.0037289 -0.0017068 0.0028335
   </decimate_by_three_filter>
   <decimate_by_five_filter kind="acausal" number="34">
      0.0218528 0.0458359 0.0908603 0.1359777 0.1830881
      0.1993418 0.1957624 0.1561194 0.0994146 0.0346412
      -0.0236544 -0.0580081 -0.0703257 -0.0555546
      -0.0287709 0.0032613 0.0267938 0.0358952 0.0311186
      0.0134283 -0.0028524 -0.0170042 -0.0176765
      -0.0123123 -0.0036798 0.0057730 0.0059817 0.0083501
      0.0000581 0.0005724 -0.0033127 0.0004411 -0.0030766
      0.0016604
   </decimate_by_five_filter>
</lsm_processing>
```

Table N.4: XML LSM Processing Elements

Element	Description
`<phz kind="lsm">`	The phase of the tide with respect to the local potential. There is one kind for the LSMs, `lsm`.
`<amp kind="lsm">`	The amplitude of the tide. There is one kind for the LSMs, `lsm`.
`<linear_strain factor="strain_per_count">`	Strain per unit count.
`<optical_anchor_retroreflector_scalefactor>`	Scale factor applied to the retro reflector optical anchor measurements to create the optical anchor correction.
`<optical_anchor_interferometer_scalefactor>`	Scale factor applied to the interferometer optical anchor measurements to create the optical anchor correction.
`<laser_frequency_correction_scalefactor>`	Laser frequency correction scale factor.

N.3 Data

The data are contained in observation (<obs>) elements. In the LSM XML files there is one observation element for each sample point. For the BSM XML files there are seven observation elements for each sample point, one for each of: gage0, gage1, gage2, gage3, Eee+Enn, Eee-Enn and 2Ene.

N.3.1 BSM Observation Elements

An example of borehole strainmeter observation element follows. The elements are described in Table N.5.

```
<obs strain="gage2">
   <date>2005-09-25T23:50:00</date>
   <doy>267</doy>
   <MJD>53638.993056<MJD>
   <s>     10.2632</s>
   <s_offset> -1.23</s_offset>
   <s_q>g</s_q>
   <tc>   0.0185</tc>
   <dtc>  10.2500</dtc>
   <apc>  -0.0091</apc>
   <apc_q>i</apc_q>
   <v>2005270201116</v>
   <level>2a</level>
</obs>
```

Table N.5: XML BSM Observation Elements

Element	Description
<obs strain="gage2">	Observation element. There are seven kinds for the BSM: gage0 Linearized strain data from gage0. gage1 Linearized strain data from gage1. gage2 Linearized strain data from gage2. gage3 Linearized strain data from gage3. Eee+Enn Areal strain computed using data from the 4 gages. Eee-Enn γ_1 shear strain computed using same data. 2Ene γ_2 shear strain computed using same data.
<date>	Date of observation (YYYY-MM-DDTHH:MM:SS).
<doy>	Day of year (1 .. 366).
<MJD>	Modified Julian Date.
<s>	Strain (microstrain, extension positive).
<s_offset>	Offset in strain time series (microstrain). The offset is not applied to the strain measurement. The user can choose to apply the offsets when extracting data from the XML files using the rdstrain.pl program. This element was introduced 10 Jan 2006. XML files written before that do not have <s_offset> elements.
<s_q>	Quality of strain value. The possible quality flags are: g good Observation is of good quality. m missing There is no strain measurement for the sample time. b bad Observation is thought to not be a true measure of the strain (e.g., it is known that engineers where working at the strainmeter at that time.) i interpolated Observation value has been determined through linear interpolation. p imputed Observation value has been generated using a combination of a linearly interpolated value and a predicted value (e.g., linearly interpolated strain + predicted earth tides).
<tc>	Tidal correction (microstrain).
<dtc>	Detrend correction (microstrain).
<apc>	Atmospheric pressure correction (microstrain).
<apc_q>	Atmospheric pressure correction quality flag. The same flags are used as for the strain quality element.
<v>	Version. A version number that uniquely identifies the XML file. It reflects the date and time of the file creation (YYYYDDDHHMMSS).
<level>	Processing level, 2a or 2b.

N.3.2 LSM Observation Elements

An example of an LSM observation element follows. Elements not already described in Table N.5 are described in Table 6.

```
<obs strain="lsm">
   <date>2005-05-30T00:00:00</date>
   <doy> 150</doy>
   <MJD>53520.000000</MJD>
   <s> 26.0</s>
   <s_q>g</s_q>
   <tc> 62.7</tc>
   <oaic> 43.0</oaic>
   <oaic_q>g</oaic_q>
   <oarc> 314.0</oarc>
   <oarc_q>g</oarc_q>
   <v>20050829174853</v>
   <level>2a</level>
</obs>
```

Table N.6: XML LSM Observation Elements

Element	Description
<obs strain="lsm">	Observation element. There is one kind for the LSM: lsm, linear strain.
<oaic>	Optical anchor interferometer-end correction.
<oaic_q>	Quality flag for the optical anchor interferometer-end correction. Possible values are as listed for strain in Table N.5.
<oadc>	Optical anchor retro reflector-end correction.
<oarc_q>	Quality flag for the optical anchor retro reflector-end correction. Possible values are as listed for strain in Table N.5.

Appendix O

PBO Network Buildout and Data Rates

Table O.1: PBO Network Buildout

PBO Year	Cumulative GPS Stations			Cumulative Strainmeters	
	Continuous[1]	Survey-mode	High-rate[2]	Laser	Borehole
1	50	50	13	1	1
2	475	100	63	3	10
3	725	100	125	5	41
4	975	100	188	5	72
5	1100	100	245	5	103
6	1100	100	245	5	103
7	1100	100	245	5	103
8	1100	100	245	5	103
9	1100	100	245	5	103
10	1100	100	245	5	103

[1] Stations that record 15-second data. Total includes 209 PBO Nucleus stations and sixteen stations at USArray sites in the eastern United States.

[2] Stations that record both 15-second and 5-sps data. All high-rate GPS stations are also continuous GPS stations, so the total number of new PBO GPS stations is not over 875. Assumes that no more than 100 of the 875 GPS stations in the final PBO network will return any high-rate data in any given year, and that the ratio of 100:875 is maintained throughout the network buildout.

Table O.2: PBO Network Data Volumes (TB)

PBO	Calendar	Annual Volume					Cumulative Volume				
Year	Calendar	LSM	BSM	GPS	Seismic	Total	LSM	BSM	GPS	Seismic	Total
1	2003–04	0.000	0.030	0.000	0.000	0.030	0.000	0.030	0.000	0.000	0.030
2	2004–05	0.001	0.224	0.023	0.083	0.331	0.001	0.254	0.023	0.083	0.361
3	2005–06	0.002	0.529	0.116	0.413	1.060	0.003	0.783	0.139	0.496	1.420
4	2006–07	0.003	0.811	0.259	0.926	1.999	0.006	1.594	0.398	1.422	3.420
5	2007–08	0.003	1.058	0.403	1.438	2.902	0.009	2.652	0.801	2.860	6.322
6	2008–09	0.003	1.165	0.477	1.703	3.348	0.013	3.817	1.278	4.563	9.670
7	2009–10	0.003	1.165	0.477	1.703	3.348	0.016	4.982	1.754	6.265	13.017
8	2010–11	0.003	1.165	0.477	1.703	3.348	0.019	6.147	2.231	7.968	16.365
9	2011–12	0.003	1.165	0.477	1.703	3.348	0.022	7.312	2.708	9.671	19.713
10	2012–13	0.003	1.165	0.477	1.703	3.348	0.026	8.477	3.185	11.374	23.061

LSM: Laser strainmeter; BSM: Borehole Strainmeter; GPS: Total GPS; Seismic: Borehole seismic

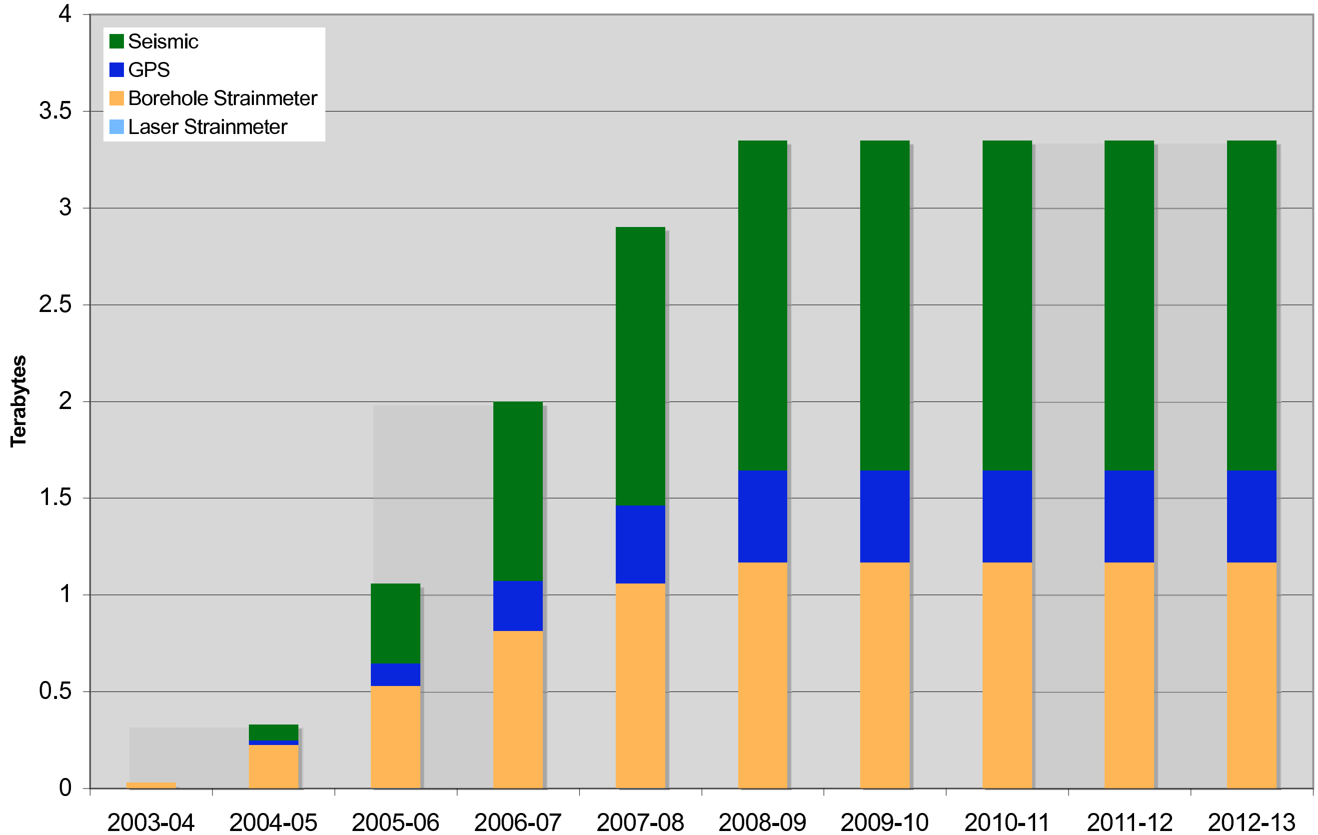

Figure O.1: Annual data volume for PBO network, by data type. Values in terabytes. Maximum annual data volume is approximately 3.3 TB/yr. The laser strainmeter volume is so small that it is not visible in this plot.

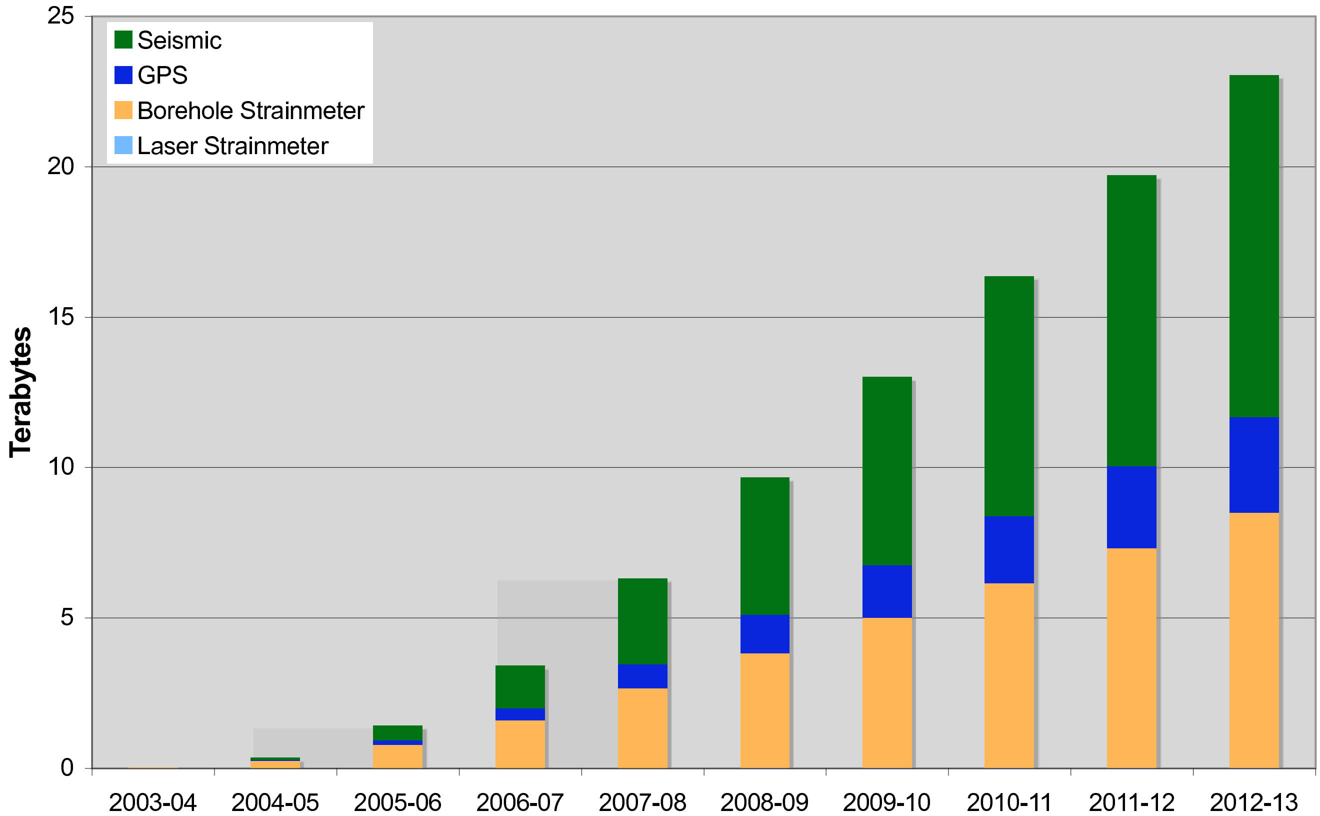

Figure O.2: Cumulative data volume for PBO network, by data type. Values in terabytes. Overall 10-year cumulative volume is approximately 23 terabytes. The laser strainmeter volume is so small that it is not visible in this plot.

Bibliography

Agnew, D. C., SPOTL: Some Programs for Ocean-Tide Loading, *SIO Reference Series 96-8*, Scripps Institution of Oceanography, 1996.

Agnew, D. C., PIASD: Programs for the Interactive Analysis of Strainmeter Data, 2004, http://jacinto.ucsd.edu/piasd.intro.pdf.

Agnew, D. C., GHAM: a compact global geocode suitable for sorting, *Comput. Geosci.*, *31*, 1042–1047, 2005, Unpublished manuscript.

Ash, M. E., Determination of Earth satellite orbits, *Tech. Note 1972-5*, Lincoln Laboratory, MIT, 1972.

Bar-Sever, Y. E., A new module for GPS yaw attitude, in *Proceedings of the IGS Workshop: Special Topics and Directions*, edited by G. Gendt and G. Dick, pp. 128–140, 1996.

Beutler, G., E. Brockmann, W. Gurtner, U. Hugentobler, L. Mervart, and M. Rothacher, Extended orbit modeling techniques at the CODE Processing Center of the International GPS Service for Geodynamics (IGS): theory and initial results, *manuscripta geodaetica*, *19*, 367–386, 1994.

Dong, D., and Y. Bock, Global Positioning System network analysis with phase ambiguity resolution applied to crustal deformation sudies in California,, *J. Geophys. Res.*, *94*, 3949–3966, 1989.

Dong, D., T. A. Herring, and R. W. King, Estimating regional deformation from a combination of space and terrestrial geodetic data, *J. Geodesy*, *72*, 200–214, 1998.

Egbert, G. D., and S. Y. Erofeeva, Efficient inverse modeling of barotropic ocean tides, *Journal of Atmospheric and Oceanic Technology*, *19*, 183–204, 2002.

Gladwin, M. T., and R. Hart, Design parameters for borehole strain instrumentation, *Pure Appl. Geophys.*, *123*, 59–88, 1985.

Hart, R. H. G., M. T. Gladwin, R. L. Gwyther, D. C. Agnew, and F. K. Wyatt, Tidal calibration of borehole strainmeters; removing the effects of small-scale inhomogeneity, *J. Geophys. Res.*, *101*, 25,553–25,571, 1996.

Niell, A. E., Global mapping functions for the atmospheric delay, *J. Geophys. Res.*, *101*, 3227–3246, 1996.

Ray, R. D., IERS standards, 1995, ftp://maia.usno.navy.mil/conventions/chapter8/ray.f.

Schaffrin, B., and Y. Bock, A unified scheme for processing GPS phase observations, *Bulletin Geodesique*, *62*, 142–160, 1988.

Scherbaum, F., and M. P. Bouin, FIR filter effects and nucleation phases, *Geophys. J. Int.*, *130*, 1997.

Springer, T. A., G. Beutler, and M. Rothacher, A new solar radiation pressure model for the GPS satellites, in *IGS Analysis Center Workshop*, 1998.

Tamura, Y., T. Sato, M. Ooe, and M. Ishiguro, A procedure for tidal analysis with a Bayesian information criterion, *Geophys. J. Int.*, *104*, 507–516, 1991.

Wu, J. T., S. C. Wu, G. A. Hajj, W. I. Bertiger, and S. M. Lichten, Effects of antenna orientation on GPS carrier phase, *manuscripta geodaetica*, *18*, 91–98, 1993.

Zumberge, J. F., M. B. Heflin, D. C. Jefferson, M. M. Watkins, and F. H. Webb, Precise point positioning for the efficient and robust analysis of GPS data from large networks, *JĠeophys. Res.*, *B102*, 5005–5017, 1997.