A Distributed Organizational Structure For PBO Network Operation And Data Analyses

J. L. Davis, R. A. Bennett, and G. Blewitt

Motivation

GPS network management, including data retrieval, archiving, and analysis (from phase data through velocity vectors), and site maintenance, should be an integral part of the planning, implementation, and operation of PBO. Furthermore, the experience of our community is that the complicated process of GPS data analysis benefits greatly from overlap and redundancy, and from the direct interest of the analyst in the science under investigation. The PBO White Paper places data analysis mainly in the operational stage, but we argue that taking a more fundamental view of these activities will contribute significantly to the overall success of PBO. Specifically, our motivation is to achieve the following benefits to the community:

(1) *GPS network implementation*: Our experience has been that the network implementation benefits greatly from having a data management system in place as soon as the site becomes operational, mainly to work out communication problems. In the case of a PI-based project, this management system is fairly straightforward and involves a minimum of equipment and personnel. In the case of PBO, it is more complicated, with responsibilities for the monitoring and the data flow for hundreds of sites having to be worked out beforehand.

(2) *GPS network operation*: Again, our experience is that the project benefits greatly from having a method by which the data analysts have a tight link to those responsible for physical site maintenance. If the analysis is performed in a timely manner, with specific responsibilities previously defined, problems can be identified fairly quickly. If an agreed upon system is in place for responding to problems (for example, in the BARGEN project we begin with a "911" e-mail from the analysis group), the problems can be rectified in a timely *and cost effective* manner.

(3) *Science*. The goals of PBO are scientific, and everything we do should be aimed at maximizing the scientific benefit of PBO. Having standards and practices for data analysis previously decided upon will enable us to focus on the science as well as insure more accurate data products obtained in a more timely manner. The goal is to make the flow from observable to data product well defined, efficient, and cost effective while insuring the highest quality.

Below, we highlight the data analysis aspects of the PBO operation. We do not discuss other aspects of implementation and operation that will be very important, such as determining standards for equipment, purchases, science goals, etc.

Approach

Our proposed approach is to adapt for use by PBO a model that has already been well developed and tested. The International GPS Service (IGS) is an organization that has been operating for some years, and has contributed much toward standardization of data management and analysis procedures. The IGS forms a much looser group than would the investigators of the PBO, however, a fact that we could use to our benefit to adopt improved procedures. There is not always a tight feedback, for example, between IGS analysis centers and those responsible for maintaining sites.

The IGS is composed of a number of "centers," each with a specific responsibility. Each of the centers is in reality an individual or a group that performs a specific task. These centers can dedicate themselves to performing these tasks effectively and with high quality. The number of committees is kept to a minimum. The Central Bureau is instrumental in coordinating the centers and imposing standards, but is charged with performing no data management tasks. The projects and working groups are specific to IGS, and would be replaced in PBO mainly by the science investigators. Although the IGS organization may seem complex, it is actually well streamlined and designed to avoid bureaucracy, since everyone concentrates on performing their task well, and committees do not interfere with operations.

Regional vs. Centralized Structure

One of the features of the IGS organizational structure that works extremely well is its regional structure. Many of the GPS sites around the globe are grouped into regional networks that are analyzed by a regional center. Most of these regional centers are operated by groups that have a specific interest in these GPS data as part of a scientific project. (As a result, they have a stake in the quality of the data products they pass on.) Since these groups would be analyzing these data as a part of their projects anyway, it makes sense for these centers to act as IGS analysis centers.

The question is, do we want to adopt a distributed regional structure for the management of the PBO, or a highly centralized structure? The PBO White Paper describes a distributed management structure for the operation of PBO, and we now propose that the data analysis be organically integrated into this structure. A distributed management structure has many advantages. The most important may be that, as with the IGS structure, *regional site management may be overseen by groups that have a specific interest in the GPS data from these regions.* Many groups currently have ongoing GPS studies in regions within the PBO area, and it is logical and reasonable to expect that their main interest may continue to be focused within these regions. A group that oversees 50–75 GPS sites in which they have a clear scientific interest will be much more motivated to insure that these GPS sites are functioning properly and are attended to in a timely manner when they are not.

As described below, the products from the analysis of GPS data within the different regions should be combined in a statistically rigorous manner to produce an internally consistent velocity field for the entire PBO area. However, to insure the highest accuracy for studies that depend on a PBO-wide network, such as the propagation of strain, we

also propose that the PBO Backbone Array be analyzed *in toto* by one or more groups. In essence, the Backbone Array is treated as a separate "region" for data analysis purposes, although sites within the Backbone Array may also be incorporated into regional analysis. This method is one of the ways that redundancy is incorporated in our proposed organizational structure. *Redundant data reduction should not be viewed as inefficiency, but rather as a necessity for the assessment of accuracy in a region where many of the geophysically important signals may, after all, be small and time-dependent.*

Required PBO Facilities

The organizational structure for PBO network operation and data analyses we propose requires the following facilities. Some of these are discussed in the PBO White Paper, but we include them for completeness and to demonstrate how they interact with the new facilities we propose. Each of the facilities below is in reality an investigator-based group. Groups may, of course, have overlapping or multiple responsibilities. The overall structure is depicted in Figure 1.

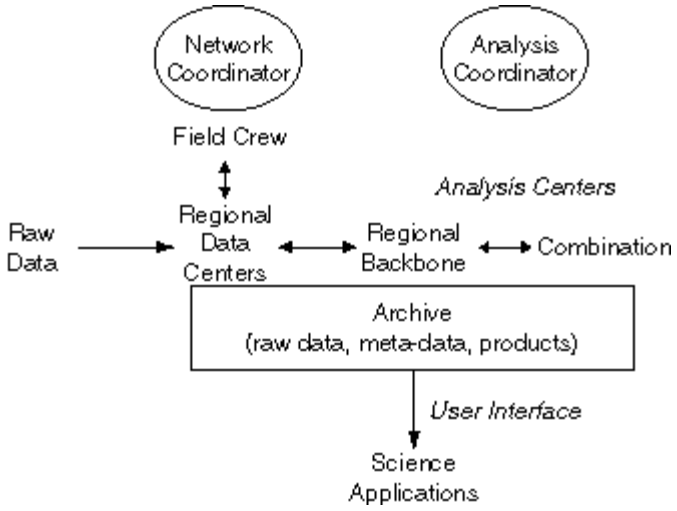

Figure 1. Proposed organization for PBO analysis.

- *PBO Archive:* The Archive will house raw data (e.g., RINEX), meta-data (e.g., site logs), and geodetic and geophysical data products (e.g., SINEX files). The archive should be accessible via the Internet. It should be distributed physically, but seamless to the user. The archive will serve as the primary source of data and data products to the PBO community including PBO Analysis Centers as well as the general end product user community. The UNAVCO Seamless Archive would be the obvious choice here.

- *Regional Data Centers:* Regional Data Centers are responsible for the daily interaction with a subset of the sites in the PBO network, including data collection, site operations, and overall fidelity of sites. All stations in the PBO should be associated with a subnet managed by a single Regional Data Center. Regional Data Centers will perform (typically daily) data downloads and translate (to RINEX) data from the subnet, as well as

maintain the site logs for the stations in the subnet. The Regional Data Centers be responsible for providing data and meta-data to the archive but should also make these data available to regional centers and the network coordinator, in effect serving as subnet archive in order to reduce the load on the primary archive when appropriate.

- *Equipment and Field Crew Facilities:* We include the equipment and Field Crew Facilities in this document since they receive feedback on site status from the Regional Data Centers. One or more facilities with dedicated engineers and surveyors to maintain the PBO equipment pool, install and repair permanent sites and associated communications equipment, build monuments, and perform campaign surveys.

- *Analysis Centers:* Analysis centers are responsible for reducing the raw data collected from the PBO network to scientifically useful data products. Types of analysis centers are described below. Organizationally it is advantageous to have Regional Analysis Centers in close association with particular Regional Data Centers, with subnets determined based on scientific objectives.

- *Network Coordinator:* Individual responsible for the state of the entire GPS network. The Network Coordinator is the one individual who can be counted on for information on the current status of the network, and who can coordinate problem solving.

Type of Analysis Centers

As described above, it will be necessary to have several different types of analysis centers, each with a different focus, clearly defined and coordinated.

- *Regional Analysis Centers:* Regional Analysis Centers retrieve data from the Archive (or directly from the Regional Data Centers). Regional Analysis Centers are responsible for the production of data products, primarily network solutions, for subsets of the total PBO network, including both campaign and continuous GPS sites, and providing geodetic data products (e.g., SINEX files) to the archive. Ideally there would be several analysis centers for each region, with each analysis center using different software packages. It might be more flexible, and even more preferable, to have overlapping network solutions.

- *Backbone Analysis Centers:* Backbone Analysis Centers retrieve data from the Archive. Backbone Analysis Centers are responsible for the production of data products, primarily network solutions, for the backbone component of the total PBO network. Backbone Analysis Centers would provide geodetic data products (e.g. SINEX files) to the archive. Ideally, there should be multiple backbone analysis centers each using different software packages. Backbone Analysis Centers should ensure that the PBO backbone is adequately tied to the IGS, and should make use of precise orbits and EOP provided by the IGS.

- *Combination Analysis Centers:* Combination Analysis Centers retrieve data products from the PBO Archive and from a global analysis center. Combination Analysis Centers are responsible for the synthesis of the data products from backbone and regional analysis centers to realize geophysical products (e.g., the velocity field). All geophysical data products are sent to the Archive. There should be multiple Combination Analysis Centers, each using different software packages.

- *Analysis Coordinator:* The Analysis Coordinator systematically cross-checks solutions from different analysis centers in an attempt to identify and mitigate errors and identify biases etc. This role has proved very important in IGS. Not all of the Combination Analysis Centers perform at the same level. The Analysis Coordinator is also responsible for producing the final (geophysical) PBO products (e.g., North America velocity).

Standards and Performance Committee

Comprised of the managers of the data and analysis centers, the network coordinator, the EFCF manager, and the archive manager, this committee should be formed to establish site and network performance criteria. They should adopt standards for analysis procedures. Consistency with IGS and IERS standards should be maintained whenever possible. They should meet regularly to reassess the standards. The techniques used by the analysis centers should conform to these standards.

Required Effort and Recommendations

Again, because of its importance in operations, we recommend that data analysis be viewed as an integral part of the PBO operations. To that end, *we recommend strongly that the support for data analysis be obtained within the PBO operations budget, as implied by the PBO White Paper.*

We assume that the PBO outside the Canadian region is divided into four overlapping regions, and the Backbone is divided into two overlapping regions. (The overlap will enable us to have some redundancy in a cost-effective manner.) The number of sites per region will be unequal. For example, it probably makes sense for the sites within SCIGN to remain as a distinct "region." Aside from the PBO archives and Equipment and Field Crew Facilities, we conservatively assume that each center or coordinator requires 1.5 FTE. Using two Combination Analysis Centers for minimum redundancy, our proposed organization requires 21 FTEs for analysis. This organization requires approximately $2.5 million per year, or approximately $2K per site per year. This amount could be more or less depending on the level of redundancy required, the number of regions, and the amount of overlap between centers.

The PBO White Paper presents a proposed budget that does, in fact, include support for data analysis. It is mostly factored into the "Operation" line item that includes "technical staff for field maintenance, data downloading, calibration, reduction, and archiving at $6K/site per year." Our calculation is that for a fully deployed PBO and for the IGS-like analysis structure we propose, the "reduction" would require only one-third of the "operation" resources, but it may not scale as assumed in the White Paper.