

UNAVCO Annual Report

EarthScope Comprehensive SAR Archive

Report Date: August 2012

For period of performance September 2011 – August 2012

NSF Grant EAR-0952375

Submitted to:

Russell Kelz

National Science Foundation, Instrumentation and Facilities Program,
EarthScope Science Utilization

Craig Dobson

NASA Headquarters

Elizabeth Lemersal

USGS

Submitted by:

Fran Boler

UNAVCO

Eric Fielding

Jet Propulsion Laboratory

Rowena Lohman

Cornell University

1.0 Executive Summary

The EarthScope Comprehensive SAR Archive proposal set out a vision for building on the extensive archive content developed through the GeoEarthScope InSAR archive acquisitions and ongoing WInSAR data orders. Funds were requested for imagery acquired by ESA largely after October 2008, the end of the GeoEarthScope component of the EarthScope facility construction project. Funds were also requested to task the Envisat satellite until October 2010 when Envisat's InSAR capability was expected to end. At the end of 2010, ESA confirmed that costs were waived for all data ordered in 2010, and that henceforth data could be ordered and would be delivered without charge; tasking fees were similarly waived. Following this data policy change, in February 2011, NSF and NASA approved a realignment of WInSAR and EarthScope Comprehensive SAR archive funding to allow for significant investments in the SAR archive infrastructure at UNAVCO. An extensive hardware, software, database and web site upgrade is now complete.

NASA's contribution towards this award has been directed to support of the International Group on Earth Observations (GEO) Geohazard Supersites Initiative. The Group on Earth Observations is coordinating efforts to build a Global Earth Observation System of Systems, or GEOSS. A principal aim of GEO is international cooperation in data sharing to support disaster preparedness planning and disaster mitigation. The Geohazard Supersites provide access to spaceborne and in-situ geophysical data of selected sites prone to earthquake, volcano or other hazards. UNAVCO's role in the SAR area has been to provide SAR data ordering from ESA; data management for the ordered data; and web site hosting and event Supersites content management.

With the substantial completion of the SAR infrastructure upgrade project at UNAVCO, and with the completion of the WInSAR grant as of May 31, 2012, support for WInSAR activities will be the main focus of this grant during the final year prior to transition of support for WInSAR and SAR activities to a core UNAVCO Facility activity in October, 2013.

2.0 Background

In February, 2011, NSF and NASA approved a realignment of WInSAR and EarthScope Comprehensive SAR archive funding to allow for significant investments in the SAR archive infrastructure at UNAVCO. This report details the extensive upgrade that was envisioned, designed, developed, and implemented between February, 2011 and July, 2012 and included major improvements to storage, servers, database engine, database schema, metadata, SAR data organization, and more. The impetus for the infrastructure upgrade request was the reduced need for funding for data purchase due to the European Space Agency's (ESA) revised open data policy; this freed-up funding previously set aside for data tasking and purchase budgeted in the grant proposal. The WInSAR

Executive Committee (EC), as the body that drove the original proposal and that acts as a standing committee advisory to the UNAVCO Board of Directors on SAR matters, evolved a plan in conjunction with UNAVCO to address the longstanding needs within the information technology hardware and software supporting the InSAR operational activities at UNAVCO. The need for a user controllable Application Programming Interface (API) to allow automated, GUI independent search and rapid data download was identified as a high priority need by the WInSAR EC. All major elements of the upgrade have been completed and the new API, web GUI and web sites were released to the SAR community in July, 2012.

3.0 SAR Archive Infrastructure Upgrade

The goals for the UNAVCO SAR archive infrastructure upgrade were to improve the user experience while at the same time realigning the software, hardware, and database behind the archive to facilitate future maintainability and enhancement.

To improve the user experience, the following goals and tasks were identified as requirements:

- provide an Application Programming Interface to facilitate user search and improve download speed
- provide a scriptable client to the API
- reduce the complexity for accessing data by providing a unified search across the WInSAR, EarthScope, and Supersites collections
- allow the user to identify scenes to be ordered in the same interface that they see scenes already in the UNAVCO SAR and Supersites collections
- improve the search speed
- provide for polygon-based spatial search
- provide a single login to access collections available to user
- convert data in STF format to CEOS
- cut swath data into frames
- provide for TerraSAR-X data to be incorporated in collections limiting access to selected users
- redesign the web site to simplify user access to WInSAR Consortium information and SAR data

To support the user experience requirements and to improve future maintainability, these goals and tasks were identified:

- address aging web server hardware
- provide for current and future storage needs
- replace the database engine with a spatially-enabled database
- secure all data behind the firewall
- refactor ingestion software

- merge and migrate all metadata into a single database and schema for all collections (WInSAR, EarthScope, other)

This extensive set of requirements and tasks was used to guide the overall upgrade. To achieve a such wide-ranging upgrade such as this involving hardware, software, database, data management, and web site improvements in a little over one year while drawing upon numerous individuals for short periods of focused effort requires careful attention to project management throughout the process. Technical management for this upgrade project was undertaken by Dr. Lou Estey, long-time UNAVCO employee with extensive knowledge with the parallel GNSS archive IT infrastructure at UNAVCO. Dr. Estey coordinated the specification and configuration of new hardware, the allocation of tasks to individuals with appropriate expertise, and establishing timeframes for completion of tasks and the overall upgrade.

Scott Baker, University of Miami PhD student working with Falk Amelung on InSAR processing and analysis with application to volcano deformation, was hired in a part-time temporary basis to provide InSAR knowledge to the upgrade project. (Note: Scott completed his PhD and has been hired full-time at UNAVCO.)

In addition to Estey and Baker, several other UNAVCO staff with expertise in database development, Java, Python, and Perl coding, Systems Administration, and data management all shared the approximately 1 FTE available for technical developments. Some of the individuals provided just a few days of effort at the right time, with the key API developer devoting the most total hours to the project.

3.1 Pre-Upgrade SAR Archive IT Infrastructure

To appreciate the scope of the upgrade project, it is helpful to review where the infrastructure technology stood at the start of 2011. The WInSAR Archive software system was obtained from Stanford and installed at UNAVCO in 2005. The system included the functions of providing an archive for preservation and distribution of the raw data, a database for metadata storage and retrieval, a second database for membership and data orders management, and a website with both informational content and dynamic tools to show the WInSAR data catalog and allow for data access. UNAVCO contributed one server for ingestion processing, running the database, and hosting the website, and about 4 Tb of storage RAID for holding the raw files. This was older hardware that had been retired from other UNAVCO activities.

At the time, the support for WInSAR through the NSF, NASA and the USGS was dedicated solely for data acquisition through purchase from the European Space Agency. The WInSAR grant included minimal budget for personnel, and no budget for hardware. The GeoEarthScope SAR Archive was instituted in 2007, and \$93,000 was spent for about 48 Tb of RAID storage hardware. The storage came with processing servers for archiving and maintaining the GeoEarthScope collection including the database. Because it was simpler to host the GeoEarthScope SAR web site alongside the WInSAR website, the WInSAR hardware was used for both web sites. As the GeoEarthScope SAR archive

was being built, the need to upgrade the WInSAR database was recognized and effort was put in to make that a reality and to institute the ESA - requested minicat method of approving data access. This was sufficiently different from the WInSAR model that significant schema differences were needed. Meanwhile, the WInSAR production system and all of its elements needed to continue as they were to continue to serve the WInSAR community. Thus the software, hardware, and database supporting the infrastructure had bifurcated as of the initiation of GeoEarthScope.

Since the end of the construction phase of PBO, the GeoEarthScope infrastructure has remained relatively maintenance free. However, with the changed data policy at ESA, UNAVCO has more freedom to serve the EarthScope SAR collection in new ways that are more useful for data access.

3.2 Hardware Upgrades

Early in the upgrade project, a critical need to secure additional robust, maintainable storage was identified. The WInSAR and GeoEarthScope systems had together utilized the available 52 Tb of RAID storage from multiple storage systems. There was little scratch space for data management or testing, and no room for expansion.

In addition to requirements for storage of new data orders, the community has identified the mission and geographically sorted structure (relative orbit and frame) on an ftp server provided for the WInSAR collections as a goal for the EarthScope collections. With the change in ESA data policy, there is no longer a strict requirement of segregation of these two collections. Providing the EarthScope collection with access analogous to the WInSAR system and allowing for download via links (after authentication) simplifies some data delivery aspects of the API development. The user community also expressed a preference for having the Radarsat data reformatted to CEOS. To provide this geographic sort copy of the ESA data and the CEOS formatted Radarsat data while retaining the long-term storage and preservation archiving requirements, translated to the need for several tens of terabytes of additional storage to be used both for long term preservation and for data management operations.

With all of the above needs, plus anticipated future needs for storage, UNAVCO elected to devote the bulk of the hardware funds available through the budget realignment to upgrading the existing storage systems by replacing 92 500 Gb drives with 1 Tb drives, effectively adding 60 Tb of usable storage space at a cost of \$30,000. These drives were ordered in 2011 and implementation occurred this year.

Servers that were integral to the storage RAID systems were determined to have sufficient processor capacity to provide the necessary database, data delivery, data ingestion, and ancillary software needs for the SAR system. Implementation of these servers for these functions became part of the upgrade.

Some reconfiguration of systems to provide for data to be kept behind the UNAVCO firewall was incorporated into the final systems design.

3.4 Software, Database, and API

Requirements for the API development were received from the WInSAR EC and others in 2011. These requirements and others stemming from the need to redefine multiple separate SAR collections were the basis for a complete overhaul of the SAR database schema. At the same time, the decision was made to migrate the schema and all metadata into a Postgres/PostGIS database engine to reap the speed benefits of the spatial queries this allows. The database schema design is tailored for SAR searches and for identifying potential InSAR pairs. The schema also allows for control of user access to data, which is done on a per collection basis. A “pseudo-collection” consisting of orderable scenes that are not yet in available collections functions as a simple way for users to request ordering of scenes.

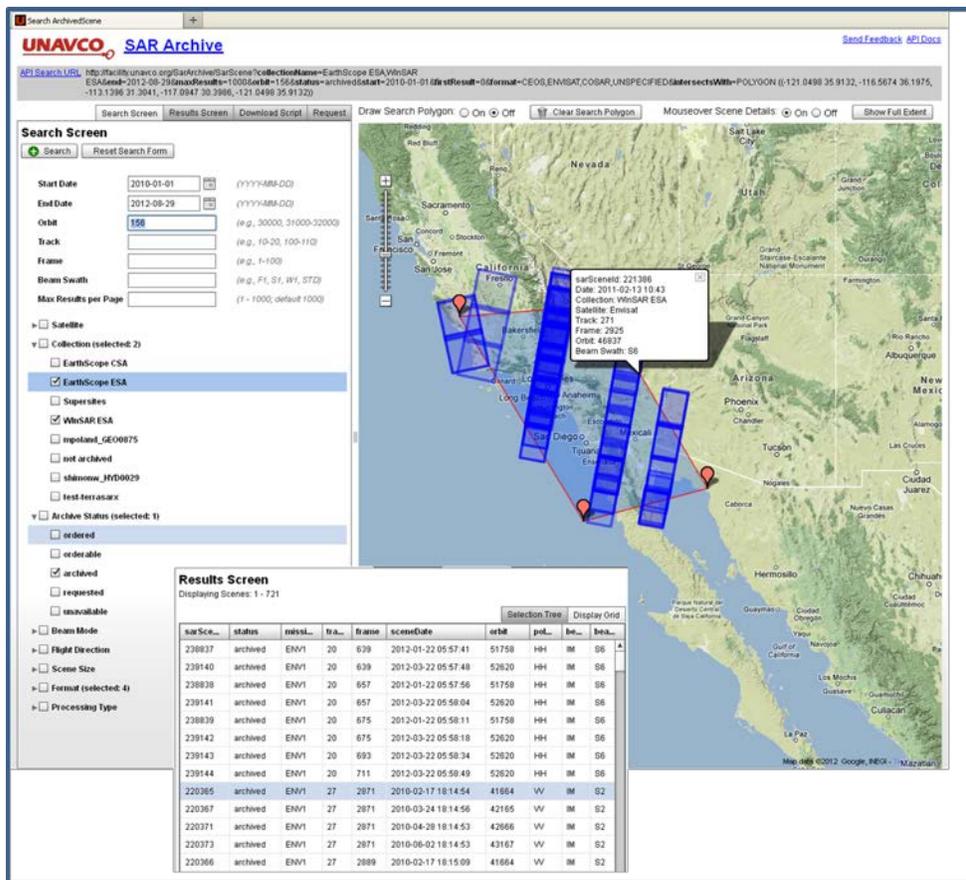


Figure 1. The new SAR search GUI accesses the database through the API middleware.

The API that was developed provides a set of web services for searching the archive, for retrieving metadata for SAR holdings, and for returning links to data. The API web services are formulated as a RESTful interface since queries are sent as HTTP GET or HTTP POST requests for resources based upon API parameters supplied in the request. The API forms the middleware between the database and user interfaces including

UNAVCO's web GUI (Figure 1). A query URL is constructed using a base access point, a resource name, and an optional parameter list. The API returns the search result in the open, text-based data exchange format JSON. These mechanisms help make the API operate in a manner familiar to many programmers. Parameters that are queryable include collection name, satellite, spatial polygon or line segment designations, orbit, beam mode, orbit, polarization, and others. UNAVCO also provides a Python command line client that many users find convenient for scripting. Users can incorporate access to the API in their own programs, and the API provides an access point for interoperability with other SAR archives or processing work flows.

For robustness of SAR services, the software that supports archiving and the API is mirrored between multiple servers. A test/development environment including processing server and a test instance of the database has also been set up to simplify the software development for SAR services.

3.5 Web Sites

The SAR information accessible from the UNAVCO web site was reorganized as part of the upgrade process. The web site content has been condensed and the look of the site has been improved. Input to the web site redesign was gathered through a user survey conducted in 2011. All of the WInSAR Consortium information is now found at winsar.unavco.org (Figure 2). The application for membership and for data user has been improved. WInSAR Consortium documents and activities, such as tasking and ISCE Software information are readily accessible.

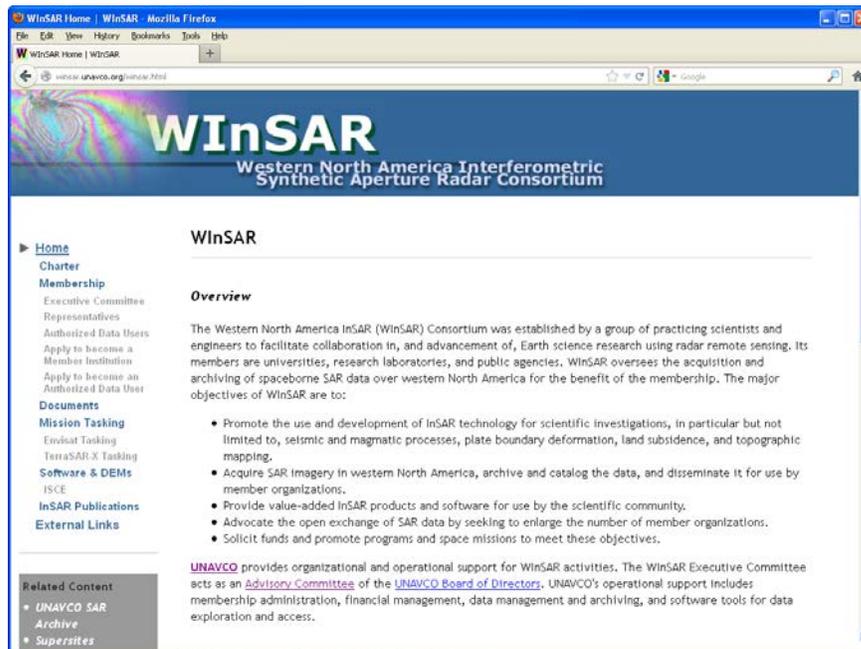


Figure 2. The new WInSAR web site - all WInSAR information regarding membership and activities of the consortium are collected on this site.

Data holdings including WInSAR and EarthScope collections are found at the new UNAVCO SAR archive site (Figure 3). Previously, there were two sites where users needed to search to gain access to UNAVCO’s SAR holdings, corresponding to the WInSAR and GeoEarthScope collections. Under the new configuration, there are multiple SAR collections (Table 1), and these are all part of the UNAVCO SAR Archive site. Collections can be added as needed, and user access to data is controlled at the collection level. It is expected that there will be multiple TerraSAR-X collections, each with a corresponding small set of users with access rights. The SAR Archive site guides users in accessing data through the web GUI, through the API, and using the command-line client that utilizes the API.

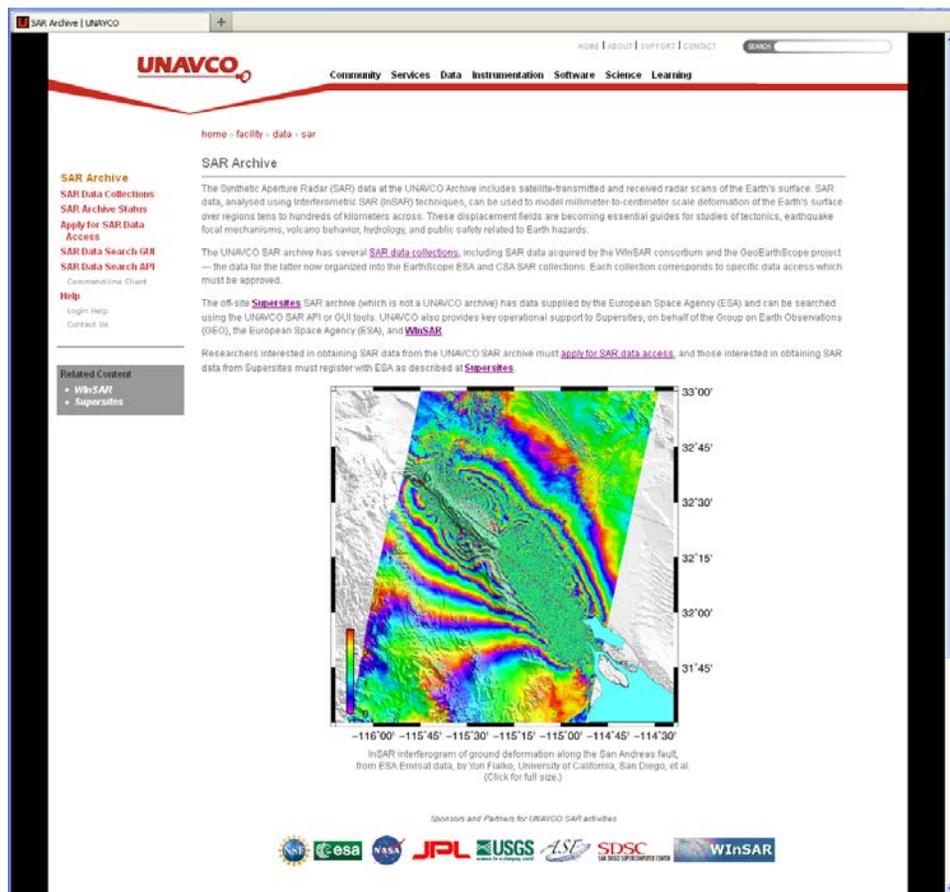


Figure 3. The new UNAVCO SAR Archive site (<http://facility.unavco.org/data/sar/sar.html>). Data collections, data access, the API, documentation, and client software are accessed from this site.

Table 1. SAR collections accessible through the UNAVCO SAR Archive site.

SAR Collection	Scenes/frames	Size, Gigabytes
WInSAR ESA	19,011	3,132
EarthScope ESA	34,975	11,257
EarthScope CSA	64,456	23,455
TerraSAR-X Hawaii	TBD	TBD
Supersites (metadata only)	N/A	N/A

The web sites have been migrated to UNAVCO’s main web server, which will provide robust service and allow leveraging of UNAVCO look and feel.

4.0 Additional Activities in Support of SAR Data Management, Access, and Distribution at UNAVCO

4.1 TerraSAR-X Background Mission Acquisition Requests

UNAVCO is supporting the WInSAR community in tasking of acquisitions by the TerraSAR-X satellite background mission. DLR has provided 1,000 acquisitions at no cost, and may increase this number if needed. Through the WInSAR web site, WInSAR scientists can submit their requests, which are prioritized by the WInSAR EC. Twenty targets are being regularly tasked.

4.2 Geohazard Supersites Support

NASA had identified continued operational support of GEO's Geohazard Supersites initiative as a key priority for UNAVCO. This support includes data ordering, management of upload of data into the Supersites archive, tasking of the TerraSAR-X satellite, and other related tasks.

Data ordering (from the European Space Agency), along with data management (download and repackaging) of the orders received, and upload of the orders to the ESA supported cloud storage (Level 3 and Terradue storage services) is the major activity. During the past year, UNAVCO has placed orders for and received nearly 6,100 scenes from ESA during the report period in locations including Turkey, Tokyo, and Southern California; files are repackaged by UNAVCO and uploaded to the Supersites cloud storage servers. In addition, data previously ordered by US scientists for their research in global Supersites locations are being contributed to the Supersites initiative; data management for more than 2,500 scenes has been provided.

A new activity this year is placing tasking/acquisition orders for TerraSAR-X data for designated Geohazards Supersites. DLR is providing tasking quota to Supersites without charge. In December, 2011, UNAVCO began tasking for thirteen targets. The tasking plan, orders, and response from DLR are published on the Supersites web page.

In addition to this data management operational activity, UNAVCO provides web site content management and hosting (Figure 4) for the Supersite web site overall, and for Supersite Event pages, the most recent case being the Ahar Iran Earthquakes (August, 2012). Table 2 shows official and other event Supersites.

Table 2. Event Supersites web pages supported by UNAVCO.

Event Supersites			
Location	Date	Category	Type
Wenchuan	May, 2008	Event	Earthquake
L'Aquila	April, 2009	Other	Earthquake
Haiti	January, 2010	Event	Earthquake
Chile	February, 2010	Event	Earthquake
Eyjafjallajokull	April, 2010	Other	Volcanic eruption
Sierra El Mayor	April, 2010	Other	Earthquake
Yushu (China)	April, 2010	Other	Earthquake
Tohoku-Oki	March, 2011	Event	Earthquake
Van	October, 2011	Other	Earthquake
Emiliano Po Plain	May, 2012	Other	Earthquake sequence
Ahar (Iran)	August, 2012	Other	Earthquakes



Figure 4. The Geohazards Supersites web site is supported by UNAVCO.

4.3 Future Activities

The activities in this grant through August 31, 2013 will include support to the WInSAR Consortium through (1) operation of the SAR Archive WInSAR and EarthScope collections and SAR API, web GUI, and API clients; (2) Executive Committee administrative support, telecons, and annual meeting; (3) member support; (4) WInSAR scientist travel to InSAR Meetings to represent WInSAR activities; (5) ordering and tasking of data with possible limited data purchase; (6) Geohazards Supersites data management and web site support; and (7) ISCE software distribution and user support after training of UNAVCO staff.