# Error Aware Near Real-Time Interpolation of Air Quality Observations in GEOSS

Scott Fairgrieve[1], Christoph Stasch[2], Stefan Falke[3], Lydia Gerharz[2], Edzer Pebesma[2]

[1]Northrop Grumman Corporation, 15036 Conference Center Dr., Chantilly, VA, USA
scott.fairgrieve@ngc.com

[2]Institute for Geoinformatics, University of Muenster, Weseler Str. 253,
48151 Muenster, Germany
{staschc,gerharz,e.pebesma}@uni-muenster.de

[3]Northrop Grumman Corporation, 1010 Market St. Suite 1700, St. Louis, MO, USA
stefan.falke@ngc.com

## 1 Introduction

This paper focuses on information system interoperability and uncertainty related to integrating standards-based sensor web observation services and data with standardized, web-based geo-processing services. The work presented in this paper was conducted with the Air Quality and Health Working Group as part of the Global Earth Observation System of Systems Architecture Implementation Pilot, Phase 3 (GEOSS AIP). The goal of the GEOSS AIP is to build a system of resources, including data and data processes that are loosely coupled and made interoperable through a service architecture based on open standards. The Model Web [1] is a blueprint vision where GEOSS services can be flexibly coupled to create workflows that convert raw, low-level sensor data into aggregated information useful to various decision makers. A frequently encountered problem in building such Model Webs is that sensor data are obtained only at a limited, fixed number of spatial locations and points in time, whereas information about the attributes measured by these sensors is needed at different locations and/or time points, or for larger regions. Interpolation methods solve this problem by utilizing available sensor observations to estimate measured attributes at the locations, regions, and times where and when they are required. Since the interpolation process involves estimating new values from existing values, the results of the interpolation are made more useful when they include estimation errors derived from the number, distribution, and quality of the existing values.

Depending on the goals at hand, the estimation errors may be negligible, or they may hinder making decisions that are sufficiently supported by data. In any case, interpolation errors, and errors in general, need examination in the next chain of information processing. In this paper, we present an application for error aware near real-time interpolation of air quality observations which has been implemented within the GEOSS AIP. In this case, the phrase "error aware" refers to the fact that the interpolation process generates error values derived from the original observations. We also describe the architecture used to integrate different implementations of the Open Geospatial Consortium (OGC) Sensor Observation Service (SOS), a standardized web service for serving sensor observations [2], that provide air quality measurements for North America and Europe with the INTAMAP interpolation service, which supports executing a variety of spatial interpolation methods [3]. The interpolation is executed using different client implementations. Though the standards used increase general interoperability, we describe conceptual and technical issues encountered and propose solutions to solve these issues.

# 2 System Architecture

An air quality and health scenario was used to frame and demonstrate the system development. The scenario goal was to use available air pollution point measurement data to provide estimates of air pollutant concentrations along with uncertainty information at locations where measurements were not available. These estimates could then be used by domain scientists to correlate pollution measurements with health data (e.g. respiratory diseases) and by public health officials in assessing health risks. The specific steps of the scenario included: a scientist searches for appropriate SOS services providing air pollutant observations in the area of interest and for OGC Web Processing Services (WPS) providing spatial interpolation processes. A WPS provides a standardized web service interface for publishing, finding, and executing geospatial processes [4]. He then utilizes the discovered WPS to interpolate the point measurements retrieved from the SOS in order to estimate measurement values for a certain point/area of interest and visualizes the results on a map. Knowing that the interpolation results are estimates, he also visualizes the interpolation error to evaluate the accuracy of those estimates.

Fig. **1** shows the architecture of the interpolation system used to support the air pollution scenario. It consists of two SOSs: an SOS interface to DataFed, which provides air quality measurements for the United States, and the AirBase SOS providing air quality measurements for Europe. DataFed is a web services-based platform developed by Washington University in St. Louis that provides access to air quality and other atmospheric and environmental measurements from a variety of monitoring networks in North America and around the world [5]. For GEOSS AIP, Northrop Grumman tailored an SOS interface to the DataFed web services so that the service and its output adhered to version 1.0 of the SOS standard and its main output format, Observations and Measurements (O&M), and could be used in conjunction with the INTAMAP WPS that provides spatial interpolation of measurement data. The DataFed web services used in the scenario provide fine particulate matter concentrations (PM2.5) measured by the EPA AIRNow monitoring network [6]. AIRNow monitors are located across the United States, primarily in urban areas. For many analyses it is necessary to derive PM2.5 concentration estimates at non-monitored locations. Filling the spatial gaps in the available data is an area of research, including the use of satellite observations [7] and the spatial interpolation of point monitoring observations.

In order to integrate European Air Quality data, the Institute of Geoinformatics of the University of Muenster set up a version of the open source 52°North SOS[1] on top of the AirBase dataset provided by the European Environment Agency (EEA). Like the DataFed data for North America, the AirBase data covers entire countries of the European Union with varying spatial coverage of measurement stations, requiring interpolation or other mechanisms for filling coverage gaps. As data is provided in text files with proprietary formats, a converter was developed to read the data from the AirBase files, convert them into the O&M standard format and use the transactional interface of the 52°North SOS to insert the observations into the web service.

[1] Available at: http://52north.org/communities/sensorweb/sos/index.html

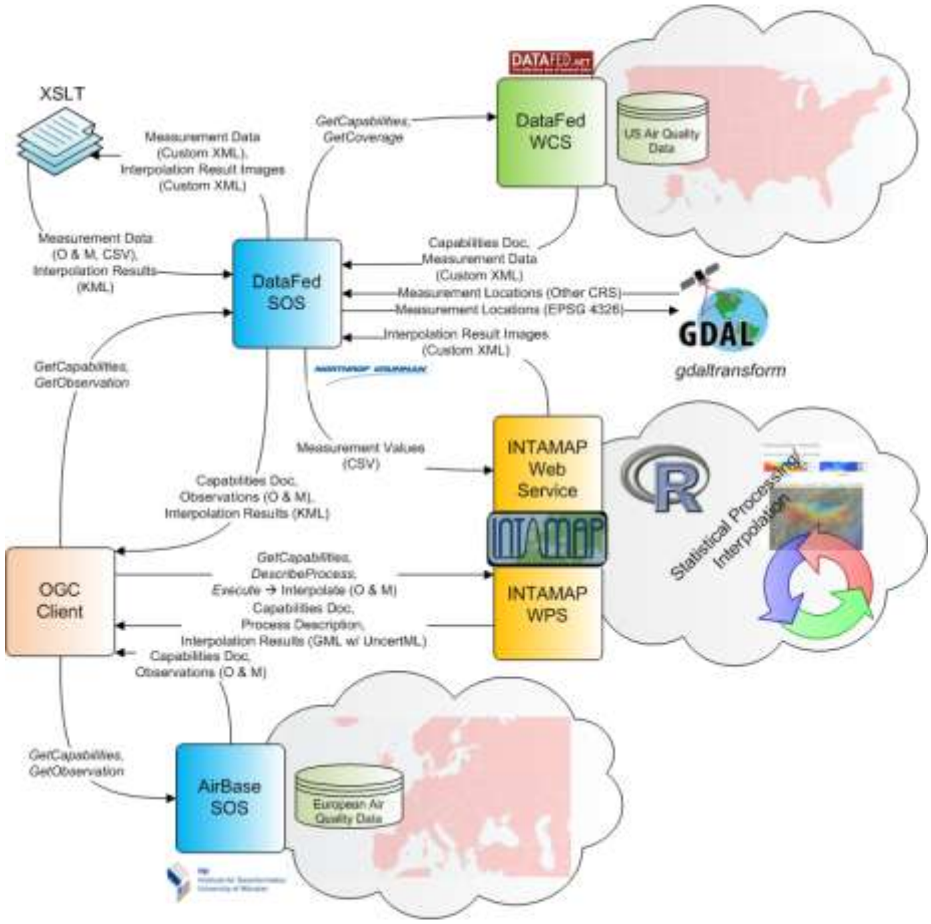



**Fig. 1.** Interpolation System Architecture

To demonstrate interoperability, two different clients were used to visualize the observations provided by both SOSes: the Northrop Grumman PULSENet™ client and the 52°North OX-Client. The clients were also used to execute the interpolation service, in this case the open source INTAMAP WPS[2], and to visualize the interpolation results. For visualization of the interpolation errors, we used the open source AGUILA client[3].

PULSENet™ is a Northrop Grumman sensor web framework comprised of architecture and software for integrating heterogeneous sensor systems using open standards [8]. As part of the PULSENet™ framework, Northrop Grumman has developed or added sensor web support to multiple client applications that utilize various mapping platforms for visualizing sensor web and corresponding geospatial data. For GEOSS AIP, the Google Maps-based PULSENet™ Web Client and Google Earth were utilized to retrieve and visualize SOS and WPS results. Fig. **2** illustrates visualization of fine particulate matter measurements and interpolation results

---

[2] Available at: http://www.intamap.org/index.php
[3] Available at: http://pcraster.sourceforge.net/Aguila/1.1.0/

for a region in the Eastern US in the PULSENet™ Web Client (map imagery provided by Google).

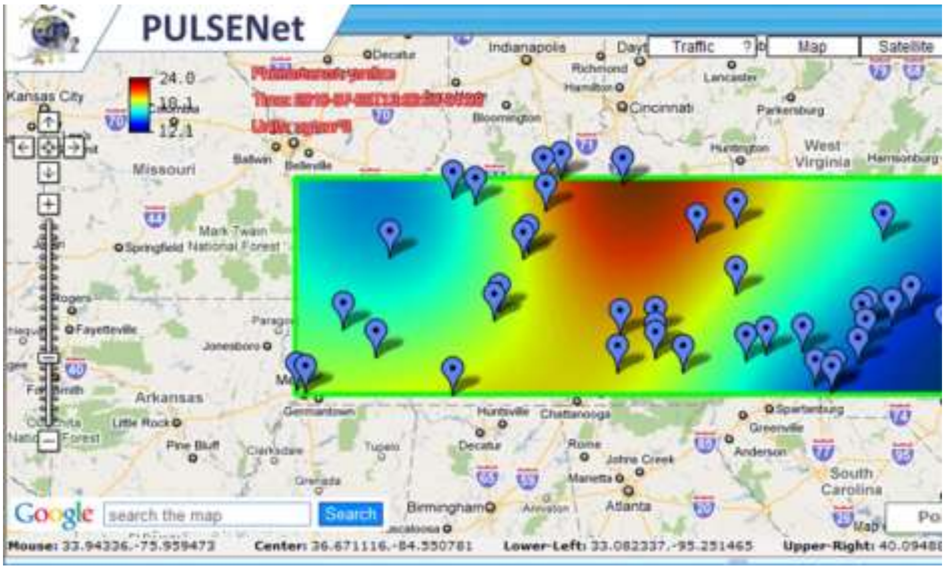
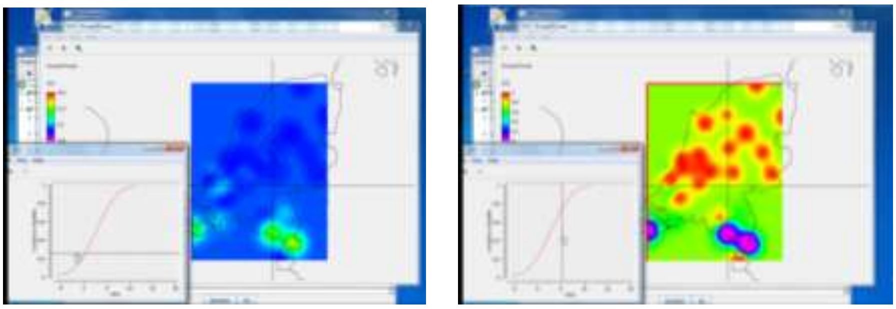



**Fig. 2.** Visualizing Eastern US Fine Particulate Matter Observations and Interpolation Results in the PULSENet™ Web Client



(A) (B)

**Fig. 3.** Viewing INTAMAP WPS Spatial Interpolation Output in the AGUILA Client showing the first quartile (A) and the cumulative probability (B)

The open source 52°North OX-Framework[4] [9] provides a customizable and extendable framework of cooperating classes which supply a reusable design applicable for the implementation of software clients integrating different spatial web services. The OX-Client utilizes the OX-Framework and provides a thick-client application that enables visualizing common geospatial data like map images, coverages or vector data along with observation data. In order to execute the interpolation, the OX-Client was extended by a WPS connector. While

[4] Available at: http://52north.org/communities/sensorweb/oxf/index.html

the interpolation results could be visualized directly using the OX-Client, the AGUILA client provides a more sophisticated means to visualize the interpolation results along with their associated estimation errors [10].
Fig**. 3** illustrates how the AGUILA client can be used to visualize the first quartile of the nitrogen dioxide ($NO_2$) concentration and the cumulative probability.

# 3 Discussion

While the standardized interfaces and formats eased the integration of the different components, some interoperability issues were encountered. The INTAMAP WPS assumes certain pre-conditions on the observations provided by the SOS: the observations must be provided in the O&M Measurement type, and the spatial information within these observations needs to be provided in a projected Coordinate Reference System (CRS). These constraints and preconditions are specified in the INTAMAP documentation, but the INTAMAP WPS Capabilities document and process description for the interpolate process do not provide parameters or properties that specify these constraints. In order to maximize interoperability in the future, a WPS that performs interpolation of observations in O&M format either needs to explicitly state input constraints in the process description for the interpolation process, or the WPS needs to be more flexible in supporting additional O&M types and CRSes.

Both the WPS standard and the SOS standard are quite generic regarding the supported types and metadata, making it difficult for a user to discover an SOS that can be used in conjunction with the INTAMAP WPS. In theory, any SOS that provides numerical measurement data from multiple sensors distributed over a geographic area should be a suitable candidate for use with the INTAMAP WPS. Due to the aforementioned restrictions regarding specific O&M formats and CRSes, finding an appropriate SOS to use is currently difficult or not possible. Additionally, flexibility in the SOS specification leads to variations across SOS implementations (e.g. different metadata in the service capabilities and different grouping of observations into offerings), which requires SOS clients to be flexible in order to be interoperable.

The INTAMAP WPS output is a coverage consisting of a standard Geography Markup Language (GML) RectifiedGrid along with Uncertainty Markup Language (UncertML) mean and variance values that map to cells in the GML RectifiedGrid. This information is suitable for plotting results on a map or producing a georeferenced image of the interpolation results that can be displayed on a map, but clients need to have specialized knowledge of how to utilize the results given that the GML RectifiedGrid/UncertML combination is currently not a standard practice.

A common profile for the web-based interpolation of observations provided by the SOS would address some of the noted interoperability challenges. Due to the generic and flexible nature of the SOS and related standards, profiles for common formats and processes are needed to ease the integration of services across several clients. An interpolation profile should define the supported observation types, the service metadata needed to support discovery and integration of SOS instances with WPS instances providing interpolation, and a standardized output format, such as the O&M format including UncertML. In addition to the need for defining interpolation profiles, accounting for the uncertainty of the original measurements used in the interpolation also needs to be addressed. The UncertWeb[5] project is helping to

[5] Information available at: http://www.uncertweb.org

address this aspect by investigating uncertainty assessment in workflows consisting of several services.

# 4 Summary

During the GEOSS AIP, the Air Quality and Health Working Group worked to advance information system interoperability and the utilization of uncertainty information by developing an error aware near real-time interpolation system that integrates standards-based sensor web observation services with standardized, web-based geo-processing services. The use of OGC SOS and WPS standards enabled an interoperable framework among multiple data services, processing services and visualization clients that could enhance current systems used by air quality researchers and public health decision makers. The effort exposed challenges in using and implementing standards for achieving interoperability between sensor web and geo-processing services and presented next steps for future GEOSS AIP pilots or other interoperability initiatives.

# References


1. Geller, G. N., Melton, F.: "Looking Forward: Applying an Ecological Model Web to assess impacts of climate change". In: Biodiversity, vol. 9, no. 3 & 4, pp. 79-83 (2008)
2. Na, A., Priest, M.: "OpenGIS Sensor Observation Service Implementation Specification". OGC 06-009r6 (2007)
3. Pebesma, E. J., Cornford, D., Dubois, G., Heuvelink, G. B. M., Hristopoulos, D., Pilz, J., Stöhlker, U., Morin, G., Skøien, J. O.: "INTAMAP: the design and implementation of an interoperable automated interpolation web service". In: Computers & Geosciences, vol. 37, issue 3, March 2011, pp. 343-352 (2011)
4. Schut, P.: "OpenGIS Web Processing Service Implementation Specification". OGC 05-007r7 (2007)
5. Husar, R. B., Höijärvi, K., Falke, S. R., Robinson, E. M., Percivall, G. S.: "DataFed: An Architecture for Federating Atmospheric Data for GEOSS". In: IEEE Systems Journal, vol. 2, issue 3, pp. 366-373 (2008)
6. Wayland, R. A., Dye, T. S.: "AIRNOW: America's Resource for Real-Time and Forecasted Air Quality Information." In: Environmental Manager, September 2005, pp. 19-24. Air and Waste Management Association (2005)
7. Al-Saadi, J., Szykman, J., Pierce, R. B., Kittaka, C., Neil, D., Chu, D. A., Remer, L., Gumley, L., Prins, E., Weinstock, L., MacDonald, C., Wayland, R., Dimmick, F., Fishman, J.: "Improving National Air Quality Forecasts with Satellite Aerosol Observations". In: Bulletin of the American Meteorological Society, vol. 86, issue 9, pp. 1249–1261 (2005)
8. Fairgrieve, S. M., Makuch, J. A., Falke, S. R.: "PULSENet$^{TM}$: An Implementation of Sensor Web standards." In: Proceedings of the 2009 International Symposium on Collaborative Technologies and Systems (CTS '09), pp. 64-75. IEEE Press, New Jersey (2009)
9. Bröring, A., Jürrens, E. H, Jirka, S., Stasch, C.: "Development of Sensor Web Applications with Open Source Software". In: First Open Source GIS UK Conference (OSGIS 2009), 22 June 2009, Nottingham, UK (2009)
10. Pebesma, E. J., de Jong, K., Briggs, D. J.: "Visualising uncertain spatial and spatio-temporal data under different scenarios: an air quality example". In: International Journal of Geographical Information Science, vol. 21, issue 5, pp. 515-527. Taylor & Francis, Inc., Bristol (2007)