

Intelligent Assimilation of Satellite Data into a Forecast Model Using Sensor Web Processes and Protocols

*Helen Conover¹, H. Michael Goodman², Bradley Zavodsky¹, Kathryn Regner¹, Manil Maskey¹,
Jessica Lu¹, Xiang Li¹, Mike Botts¹, and Gregoire Berthiau¹*

¹University of Alabama in Huntsville

²National Aeronautics and Space Administration, Marshall Space Flight Center

Abstract—The goal of the Sensor Management for Applied Research Technologies (SMART) On-Demand Modeling project is to develop and demonstrate the readiness of the Open Geospatial Consortium Sensor Web Enablement capabilities to integrate both space-based Earth observations and forecast model output into new data acquisition and assimilation strategies. Working with NASA’s Short-term Prediction Research and Transition Center, the SMART team is developing a sensor web-enabled processing workflow to intelligently assimilate Atmospheric Infrared Sounder (AIRS) satellite temperature and moisture retrievals into a regional configuration of the Weather Research and Forecast model over the southeastern United States. The SMART workflow involves mining North American Mesoscale forecasts for interesting weather phenomena, then determining whether AIRS observations are coincident with the detected weather events. The assumption is that assimilating AIRS observations of anomalous weather conditions will improve the forecast and is worth the extra computational cost; whereas, if there is no significant weather there is no advantage to assimilating additional data and not worth the extra computational cost. A variety of SWE protocols are used for data access and alert services, and for process chain definition.

I. INTRODUCTION

Advancement of both Earth science and technology is the cornerstone of the NASA Advanced Information Systems Technology program within the Earth Science Technology Office. Two of the primary objectives of the program are to identify, develop, and demonstrate advanced information system technologies that (1) enable new science observation measurements and information products and (2) increase the accessibility and utility of science data. The SMART project is achieving these objectives by focusing on developing and demonstrating the use of Sensor Web Enablement protocols and processes in the assimilation of satellite sensor observations into a numerical weather forecast model. Modelers and information technology (IT) experts on the team have worked closely together to determine how the sensor web-enabled data access and analysis tools can best facilitate data assimilation decisions while minimizing any required changes to forecast models. Currently, the Sensor Management for Applied Research Technologies – On Demand Modeling (SMART-ODM) prototype system is

under development. This paper will describe progress and lessons learned to date.

II. ENABLING TECHNOLOGIES

The Open Geospatial Consortium (OGC) suite of Sensor Web Enablement (SWE) specifications, some newly released and others under development, provides standards for data and information acquisition from sensor systems and data repositories. The OGC is an international consortium of industry, academic and government organizations using a voluntary consensus process to collaboratively develop open standards for geospatial data and information services. The OGC SWE standards framework provides specifications for interfaces, protocols and encodings that are designed to enable implementation of interoperable, service-oriented networks of sensors and applications [1]. Providing such standard interfaces to sensor data can minimize the custom software required for management, visualization and analysis of different types of sensors and observations. Sensor Web Enablement services implemented for this project include Sensor Observation Services (SOS) and Sensor Alert Services (SAS) [2]. SOS provides a web service interface for requesting, filtering and retrieving sensor system information and observations, while SAS provides a web service interface for advertising, publishing and subscribing to alerts from sensors. Other SWE protocols used in SMART-ODM include the Observations and Measurements Schema, an XML schema for encoding sensor data objects, and Sensor Model Language (SensorML), an XML schema for describing a functional model of a sensor system and related processes. Multiple processes can be combined with SensorML to form an executable process chain. Documentation for all approved OGC standards is freely available at <http://www.opengeospatial.org/standards/>.

III. SCIENCE DRIVERS

The science focus of the project is the assimilation of NASA satellite data for use in numerical models for weather forecasting. At NASA’s Short-term Prediction Research and Transition (SPoRT) Center [3], a North American Mesoscale (NAM) forecast is used as the background field for an Advanced Regional Prediction System (ARPS) Data

Analysis System (ADAS) [4] analysis, which is used to initialize the regional Weather Research and Forecasting (WRF) [5] model runs. The addition of current weather observations – such as those from the Atmospheric Infrared Sounder (AIRS) [6] used in this project – can improve the accuracy of a WRF forecast [7]. However, assimilating voluminous satellite observations into the initial conditions is computationally expensive and thus it must be done judiciously and efficiently. Modelers and IT experts on the SMART team have worked together closely to determine how sensor web-enabled data access and analysis tools can best facilitate data assimilation decisions. SMART meteorologists determined that the system should initially be configured to assimilate AIRS observations coincident with low pressure systems detected in NAM forecast data. Low pressure systems were selected because these phenomena should be relatively easy to detect in NAM forecasts, as well as valuable in deciding whether to assimilate AIRS data. A NAM forecast at any time step contains more than 600 parameter fields, of which 16 parameter fields were initially identified as related to low pressure systems. These fields include mean sea level pressure; 2m relative humidity; 6h precipitation; 2m skin temperature; convective available potential energy; convective inhibition; 10m and 200 mb east-west and north-south wind components; 500 hPa geopotential height, temperature and vorticity; and 850 hPa height, temperature and vorticity.

IV. SYSTEM COMPONENTS

Figure 1 shows the high-level architecture and information flow for SMART-ODM. The *Data Access Subsystem* consists of two independent Sensor Observation Services (SOS) for NAM and AIRS data, respectively. In the *Event Identification Subsystem*, the NAM forecasts are mined for interesting weather phenomena (Fig. 1, step 1) and an SAS will issue an alert whenever AIRS observations are

coincident with these weather events (Fig. 1, step 2). In response to such an alert, the *Data Assimilation for Models Subsystem* will acquire and preprocess AIRS observations of the phenomena detected (Fig. 1, step 3), and will insert them into the data assimilation process. A Web Notification Service (WNS) associated with the data assimilation process will issue an alert when data assimilation is complete and a new set of initial conditions for the WRF model is available (Fig. 1, step 4). The next three sections describe these system components in detail.

A. Data Access Subsystem

Operational forecasts from the NAM and observations from the AIRS instrument on the Aqua satellite are acquired as they are generated several times per day, and made available via the publicly-accessible *Sensor Observation Services* of the Data Access Subsystem. These services provide a standard web service interface to the data products, allowing the user or calling program to select any of the variables available and to subset the data by spatial and temporal range, thereby significantly reducing data volume and transfer time. The data products are also converted from their native formats to the common OGC SWE Observations and Measurements format, for easier integration with each other and with SWE-enabled visualization and analysis tools.

B. Event Identification Subsystem

The Event Identification Subsystem consists of a number of software components as shown in Figure 1, which communicate by standard SWE interface protocols (yellow boxes). Tighter integration of consecutively-invoked modules can be accomplished with a SensorML process chain (white box).

A key component of this subsystem is the *Phenomena Extraction Algorithm (PEA)*, developed at the University of Alabama in Huntsville (UAH) in previous data mining research. The PEA algorithm is a generic algorithm that performs statistical analysis of a given parameter field and detects regions with strong gradients, anomalous conditions, and/or high variance [8]. For this project, PEA was implemented in Java to facilitate integration into the Event Identification Subsystem, and an *Event Filter* module for low pressure systems was developed. After a thorough examination of phenomena identified via PEA analysis of the 16 selected NAM fields related to low pressure systems, we have found the mean sea level pressure field best for detecting these systems. Because the PEA algorithm treats both low pressure

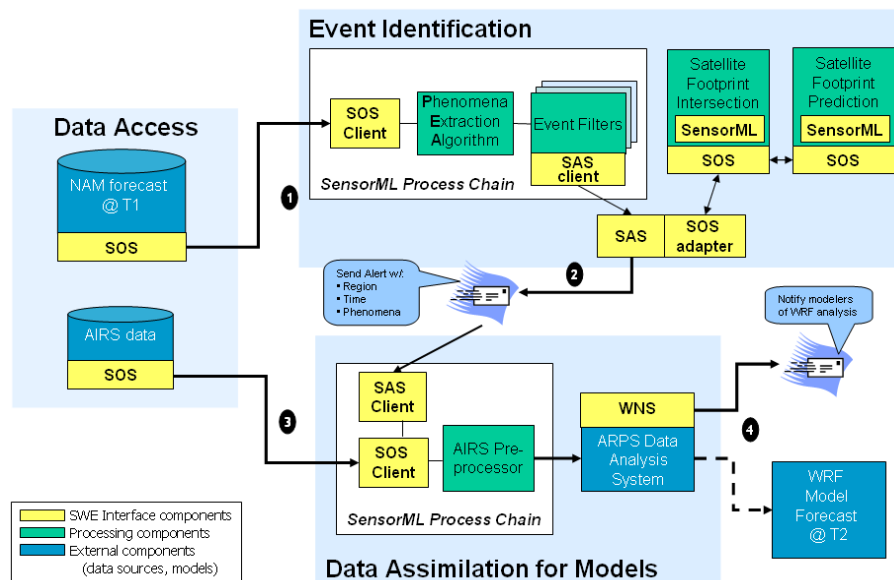


Fig. 1. Sensor-web enabled intelligent data assimilation

systems and high pressure systems as anomalous regions, the Event Filter module is required to filter out high pressure systems. Further study indicates that tracking a given low pressure system across consecutive time steps from a given forecast can make event identification more robust. The current implementation of the analysis modules of this subsystem includes PEA and low pressure filter modules which examine three consecutive forecast time steps.

Low pressure weather systems identified by the PEA and the event filters are published to a *Sensor Alert Service* (SAS). The SAS sends two types of alerts: 1) “Phenomena” alerts when a weather event (e.g., low pressure system) is found within the NAM forecast data and 2) “Phenomena/AIRS intersection” alerts when the AIRS data intersects the region of the meteorological phenomena.

A Sensor Alert Service is not an event notification system, rather, it is a registry that cross-references the different types of alerts available from a given sensor system with the consumers subscribing to these alerts. Sensor systems can advertise their capabilities and publish alerts, to which users can subscribe. Users send subscription requests to the SAS, which returns a communication endpoint (e.g., multi-user chat space, or MUC) for the alert subscribed to. The user must then open a connection to the communication endpoint to receive alerts from the sensor system. Fig. 2 shows the various SAS interfaces.

SMART is using the SAS software package from 52°North, a company which promotes the development and application of free open source geo-software for research, education, training and practical use [9]. The SAS specification is not yet an approved OGC standard, but still under community review. However, the authors were able to obtain the latest “snapshot” release of the SAS package from 52°North, which follows the current OGC SAS specification. This SAS implementation uses Extensible Messaging and Presence Protocol (XMPP), an open XML technology for presence and real time communication used in applications such as instant messaging [10].

Because the 52°North SAS package supports a much different set of sensors, several modifications and additions to the SAS software were needed for use with our prototype

sensor systems, including a “DescribeAlert” operation to provide the message structure of an alert, and a modification of the alert message schema to include geographical information (i.e., bounding box). Using existing protocols as models, we have also developed a number of SWE-enabled related modules to complete the SMART process flow:

- An *SAS Client* which provides a “publishAlert” operation for sensors to call to publish alerts, and a “postRequest” operation for users to advertise and subscribe to alerts programmatically rather than manually.
- An *SAS-SOS Adapter* which will query the Satellite Footprint Intersection SOS for satellite overpasses coincident with identified weather events.
- An *SAS Client Listener* which will listen to “Phenomena/AIRS intersection” alerts and retrieve the alerts.

SMART leverages the *Satellite Footprint Prediction SOS*, developed previously by UAH, which runs a SensorML process chain to determine satellite locations and instrument footprints at any given time. The newly-developed *Satellite Footprint Intersection SOS* determines whether a given instrument footprint intersects a specified spatio-temporal region of interest. For this data assimilation scenario, these two SOSs are used together to determine whether there is an AIRS overpass coincident with a low pressure system identified by PEA.

C. Data Assimilation Subsystem

The Data Assimilation Subsystem contains a *SensorML Process Chain* which is triggered by a “Phenomena/AIRS Intersection” alert from the Event Identification SAS. Within the process chain, an *SOS Client* retrieves the AIRS data, which are preprocessed and formatted into the ASCII format ADAS requires. The preprocessed AIRS data are then sent to the ADAS, which is already running on a compute cluster at the SPoRT Center at NASA Marshall Space Flight Center (MSFC). The *AIRS Preprocessor* is being integrated from several pieces of code the SPoRT team developed and is presently running in a single SensorML process model. This integration is in progress at this writing.

Finally, the improved ADAS analysis with assimilated AIRS observations will be made available to the local WRF model run at the SPoRT Center. Ideally, SMART-ODM will supply these analyses to other researchers around the country who are also running WRF. A *Web Notification Service (WNS)* can be used to broadcast the availability of each new ADAS analysis to the interested community. We are currently investigating the 52°North WNS package for this purpose.

V. LESSONS LEARNED

As a direct result of integrating evolving OGC SWE technologies into the SMART subsystems, we have learned that it is difficult and time consuming to balance the desire to

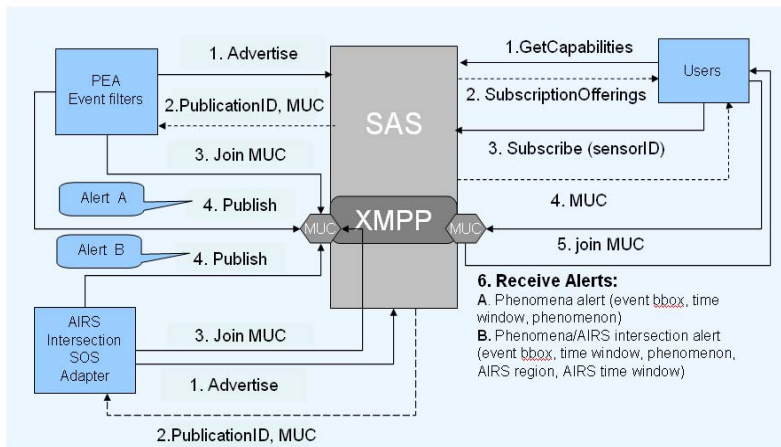


Fig. 2. SAS interfaces

explore new technology developments against the need to build stable demonstration applications. That is, upgrading services to the latest standard needs to be carefully coordinated, or distributed applications will break. To solve this problem, the team is implementing a configuration management and software version control process that has been proven to streamline integration and testing. Basically the process involves configuring three separate but nearly identical computing environments, for development, integration testing, and production and moving software from one environment to the next as certain test criteria are met. The goal is to always have a verified working version of code in production and captured in the software repository. Since development activities and integration testing are separated, the integrity of the production system is preserved.

VI. CONCLUSION

The success of SWE in applied science systems will only be achieved through the proliferation of SWE technologies within the science community. By applying SWE protocols to a real world science exercise we are demonstrating the utility and promise of a more timely and efficient satellite data assimilation process. The SMART team comprises members of the atmospheric science and meteorology disciplines as well as software and systems engineers. This collaboration of scientists and information technologists seems to be essential to the implementation of viable, real world results using the advanced technologies described here, which will in turn result in better exploitation of sensor observations for not only researchers and decision makers but for the benefit of casual end users as well.

ACKNOWLEDGMENTS

The SMART team consists of Rich Blakeslee, Gary Jedlovec, Michael Goodman and Robbie Hood of NASA/MSFC, and Gregoire Berthiau, Mike Botts, Helen Conover, Tony Cook, Xiang Li, Jessica Lu, Manil Maskey, Kathryn Regner, and Brad Zavodsky of UAH. The authors would also like to thank the NASA/MSFC SPoRT Center for assistance in designing the intelligent data assimilation prototype, and Johannes Echterhoff and Jan Torben Heuer of 52°North for insight into the 52°North approach to SWE services and access to the latest SAS software.

REFERENCES

- [1] M. Botts, G. Percival, C. Reed, and J. Davidson, "OGC Sensor Web Enablement: Overview and High Level Architecture," Open Geospatial Consortium Inc., OpenGIS White Paper OGC 07-165, 2007-12-28 2007.
- [2] H. Conover, K. Regner, M. Maskey, J. Lu, X. Li, and M. Goodman, "Evolving Sensor Web Protocols for Sensor Data Management," presented at Environmental Information Management 2008: Managing Sensor Data in Near Real Time, Albuquerque, NM, 2008.
- [3] S.J. Goodman, W.M. Lapenta, G.J. Jedlovec, J.C. Dodge, and T. Bradshaw, "The NASA Short-term Prediction Research and Transition (SPoRT) Center: A collaborative model for accelerating research into operations," presented at 20th Conference on Interactive Information Processing Systems (IIPS), at AMS Annual Meeting,

- Seattle, WA, 2004.
- [4] K. Brewster, "Implementation of a Bratseth analysis scheme including Doppler radar," presented at 15th Conference on Weather Analysis and Forecasting at AMS Annual Meeting, Boston, MA, 1996.
- [5] W.C. Skamarock, J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, W. Wang, and J.G. Powers, "A Description of the Advanced Research WRF Version 2," NCAR Tech Note, NCAR/TN-468+STR, 88 pp. [Available from UCAR Communications, P.O. Box 3000, Boulder, CO, 80307; on-line at: http://box.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf], 2005.
- [6] H.H. Aumann, M.T. Chahine, C. Gautier, M.D. Goldberg, E. Kalnay, L.M. McMillin, H. Revercomb, P.W. Rosenkranz, W.L. Smith, D.H. Staelin, L.L. Strow, and J. Susskind, "AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products, and Processing Systems," IEEE Transactions on Geoscience and Remote Sensing, vol. 41, pp. 253-264, 2003.
- [7] B.T. Zavodsky, S.-H. Chou, G. Jedlovec, and W. Lapenta, "The Impact of Atmospheric Infrared Sounder (AIRS) Profiles on Short-term Weather Forecasts," Proceedings of SPIE 65651J, 2007.
- [8] R. Ramachandran, X. Li, S. Movva, and S. Graves, "A Simple and Efficient Feature Extraction Algorithm for Geophysical Phenomena," presented at IEEE International Geoscience and Remote Sensing Symposium and 27th Canadian Symposium Remote Sensing, Denver, CO, 2006.
- [9] I. Simonis, A. Wytzisk, and J. Echterhoff, "Sensor Web Enablement: The 52°North Suite," presented at Free and Open Source Software for Geospatial (FOSS4G), Lausanne, Switzerland; on line at: http://52north.org/joomla/components/com_publications/publications/52N_SWE_FOSS4G.pdf, 2006.
- [10] P. Saint-Andre, "Streaming XML with Jabber/XMPP," IEEE Internet Computing, vol. 9, pp. 82-89, 2005.