

USING SENSOR WEB PROCESSES AND PROTOCOLS TO ASSIMILATE SATELLITE DATA INTO A FORECAST MODEL

Kathryn Regner¹, Helen Conover¹, H. Michael Goodman², Bradley Zavodsky¹, Manil Maskey¹, Gary Jedlovec², Xiang Li¹, Jessica Lu¹, Mike Botts¹, and Gregoire Berthiau¹

¹University of Alabama in Huntsville and

²National Aeronautics and Space Administration, Marshall Space Flight Center

ABSTRACT

Working closely with atmospheric scientists at the Marshall Space Flight Center, researchers at the University of Alabama in Huntsville are applying Sensor Web Enablement (SWE) technologies to the real world problem of efficiently assimilating NASA satellite data into weather forecast models in near real time. By implementing SWE protocols and services into our Data Assimilation System we expect to realize a processing framework that is distributed, interoperable and plug-and-play, thereby increasing access to scientific products in a more efficient, autonomous, and affordable way.

Index Terms— Standards, Communication Protocols, Atmospheric measurements, Geographic information systems, Weather forecasting

1. INTRODUCTION

The goal of the Sensor Management for Applied Research Technologies (SMART) On-Demand Modeling (ODM) project is to develop a prototype system that will demonstrate the readiness of the Open Geospatial Consortium (OGC) SWE capabilities to integrate both space-based Earth observations and forecast model output into new data acquisition, processing and assimilation strategies. The next two sections will explain how this project will enhance existing forecasting capabilities and why we chose SWE as the implementation framework.

1.1. Science Application

The integration of NASA satellite data into weather forecast models is a critical component of NASA's Weather Focus area. There is a need to integrate large data volumes from satellite sensors with different observational constraints and data formats into a common processing system. This can be a manually intensive and time consuming process, in that swath coverage relative to storm positions, data volume and

availability all constrain assimilation decisions. In a real-time situation, timely observational data are needed to impact the decision making processes at forecast centers. Because of the potential to reduce the loss of property and lives, appropriate data assimilation in evolving high-impact weather situations is likely a more effective use of computer time and associated manpower than forecast improvements in low-impact weather systems. The requirement to quickly process and assimilate these large data volumes from multi-sensor platforms often results in significant sub-sampling of high resolution data to accomplish the processing in a timely manner [1].

At NASA's Short-term Prediction and Research Transition (SPoRT) Center [2], Atmospheric Infrared Sounder (AIRS) data provides a key input into the regional data assimilation procedures used to produce short-term regional weather forecasts with the Weather Research & Forecasting (WRF) model [3,4]. Currently, AIRS temperature and moisture profiles are only occasionally used to support the operational forecast capabilities at the National Weather Service [5] because it is not a trivial decision on when to include the data and where spatially it will have the most effect for the forecast of day-to-day weather conditions over the United States.

The effective inclusion of AIRS data into regional forecast models is made possible through autonomous processing of model data fields, satellite orbit predictions, instrument data, and required ancillary information through sensor web capabilities and services. Researchers at SPoRT Center (located at NASA Marshall Space Flight Center) have achieved analysis and forecast improvements with the assimilation of AIRS thermodynamic profiles into the Weather Research and Forecasting (WRF) model [6]. In the assimilation process, a North American Mesoscale (NAM) forecast initializes a short-term WRF forecast that is then used as the background (i.e. first guess) for the assimilation of AIRS profiles into the Advanced Regional Prediction System (ARPS) Data Analysis System (ADAS). Because SPoRT is a transition-to-operations center [7], there is a premium on analysis and forecast speed. One hindrance of SPoRT's current system is the lengthy analysis

computation time (90+ minutes). Our contribution to reducing this time is to intelligently select only those profiles that will have the greatest impact on the analysis, such as those profiles collocated with regions of high-impact weather conditions.

1.2. Why sensor web enablement?

The OGC suite of 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 [8]. Documentation for all approved OGC standards is freely available at <http://www.opengeospatial.org/standards/>.

Sensor web enablement services implemented for the SMART prototype include Sensor Observation Services (SOS), Sensor Alert Services (SAS), and Sensor Model Language (SensorML) process chains. 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. SensorML is an eXtensible Markup Language (XML) schema for describing a functional model of a sensor system and related processes. Within the SWE framework, there is little distinction between an instrument, model, simulation, or data processing engine. They all are termed “sensor systems” and can be described in SensorML as process models or chains.

This paper will focus on SensorML, first describing the two SensorML process chains being implemented for the SMART project, then the remaining SWE components that comprise the entire Data Assimilation System.

2. SENSOR MODEL LANGUAGE PROCESS CHAINS

SensorML is used 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. That is, a SensorML process chain is a logical group of inter-connected software sub-processes or hardware components (each of which may be single processes or process chains) and the connections that link their inputs and outputs. Within the SMART Data Assimilation System we have defined two process chains: one that identifies weather events in user-specified regions and time windows, and one that preprocesses sensor data

for assimilation into models. The inner workings of these process chains are described next.

2.1. Event Identification Process Chain

The event identification process chain, shown in the green box in Figure 1, is composed of functions for retrieving North American Mesoscale (NAM) model data using an SOS Client, detecting weather events or phenomena using the Phenomenon Extraction Algorithm (PEA) [9], and event filtering. PEA is a data mining process that examines a series of NAM forecasts and identifies regions of significant weather. For SMART, the PEA is configured to detect low pressure systems, but in the future may be tuned to detect fronts and/or areas of likely convective initiation. Because the PEA algorithm treats both low pressure systems and high pressure systems as anomalous regions, the Event Filter module is required to filter out high pressure systems.

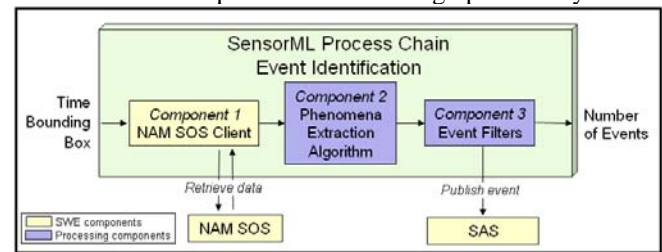


Figure 1. Event Identification Process Chain

Given a particular date, time and region, this process chain first fetches the NAM mean sea-level pressure field from the NAM SOS, then performs general event detection using PEA, and finally determines the low pressure system events of interest using the event filter function. Any events identified are published to the Sensor Alert Service (SAS), which sends an alert message with the spatial domain of the detected event as represented by a bounding box of latitude and longitude, and the temporal range encompassing one hour before and after the NAM forecast time. This alert is used to query the availability of AIRS data coincident with the time period and the spatial area of the event identified. If AIRS profiles are available, another alert is published through the SAS system.

2.2. Data Assimilation for Models Process Chain

The data assimilation for models process chain, shown in the blue box in Figure 2, is triggered by a “Phenomena/AIRS Intersection” alert from the event identification SAS. Within the process chain, the SAS Listener receives an alert, upon which an SOS client retrieves the AIRS data, which are then preprocessed and formatted into the ASCII format required by the ADAS assimilation system that is already running on a compute cluster at the SPoRT Center. ADAS produces the analyses used to initialize SPoRT’s regional WRF model runs.

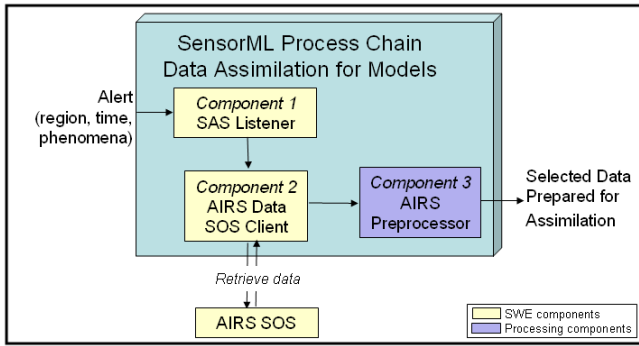


Figure 2. Data Assimilation for Models Process Chain

2.3. Implementing a Process Chain

To create process models for these functions, the authors followed examples in the white paper “Creation of Specific SensorML Process Models” [10]. The component software modules, written in java, needed minimal modification to accept inputs from the process chain and to pass outputs to the process chain. *SensorML Process Models* were defined for each of the software modules, and then these process models were connected in a *Process Chain*. A separate *Process Map* description was created to map each process model to its actual implementation. When the system is

invoked, the *Process Chain Execution Engine* parses the process chain and the various process models it contains together with the process map. It runs the component software modules in sequence and makes the data flow from one module to the next as a stream.

3. SMART'S SENSOR WEB ENABLED SATELLITE DATA ASSIMILATION DATA FLOW

Figure 3 shows the two process chains in context of the full SMART data assimilation prototype. The prototype is organized into three subsystems: data access, event identification, and data assimilation. Feeding the two process chains described above are two SOSs that provide a pertinent subset of data from the NAM weather model and archived data from the Aqua AIRS instrument. (Note that future versions will be able to process near real time as well as archived data.) The Event Identification Process Chain mines NAM forecasts for low pressure centers. A Sensor Alert Service will query for AIRS observations that are coincident with these weather events and issue a notification identifying the relevant AIRS overpasses. In response to such an alert, the Data Assimilation Process Chain will acquire AIRS observations of the phenomena detected, and generate input files for the ADAS analysis. When the improved ADAS analysis with assimilated AIRS

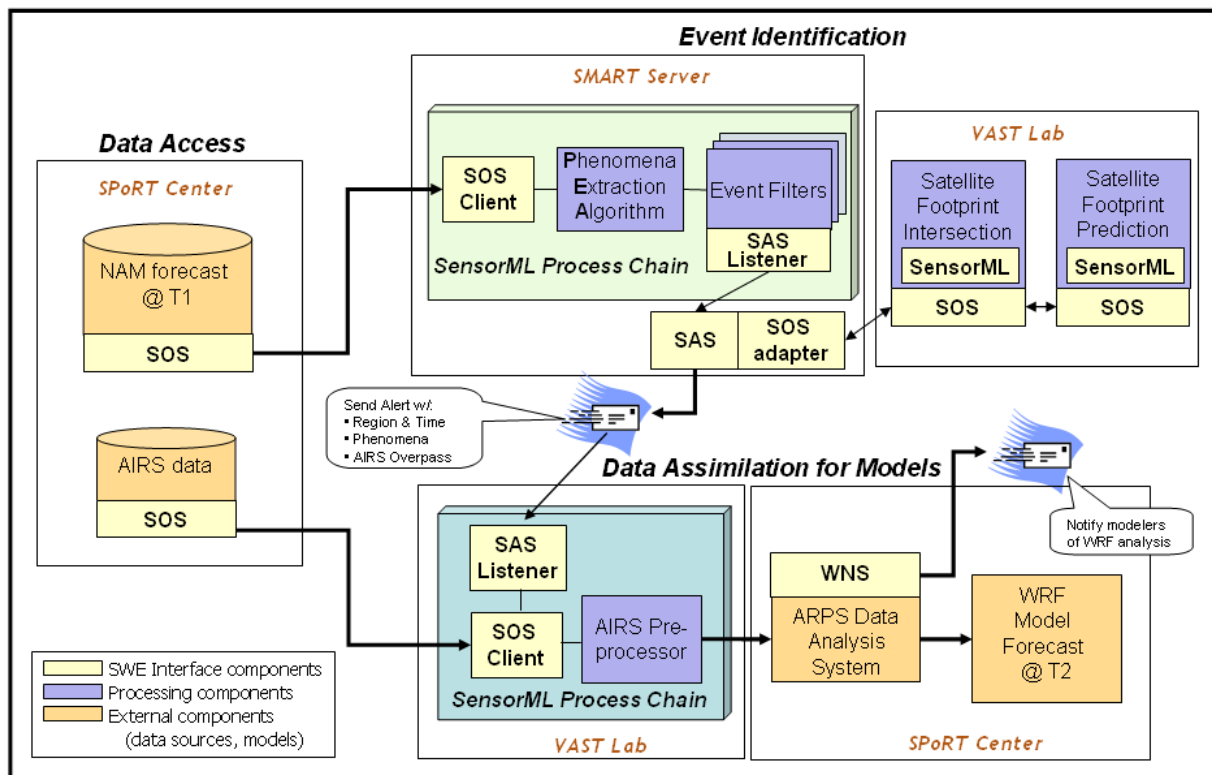


Figure 3. SMART Assimilation of AIRS Data into a Weather Forecast Model: A Sensor Web Enabled Distributed Processing System

observations is complete, a Web Notification Service associated with the data assimilation process will issue an alert to WRF modelers that a new set of initial conditions for the WRF model is available.

Note that this is a distributed, service-oriented architecture, with different data and processes on different servers. The SMART server is dedicated to this project, while the VisAnalysis Systems Technologies (VAST) Lab hosts many OGC services developed for this project and others. The SPoRT Center's cluster computers host the AIRS data, as well as routine modeling and assimilation runs, which SMART is intended to augment.

4. CONCLUSION

This SMART-ODM prototype exercises a variety of SWE components in a useful application that automates a currently manual process, that of deciding when to assimilate AIRS data into a forecast model. The initial demonstration implementation is accessing historical NAM forecast runs and archived AIRS data. Subsequent versions will culminate with the system prototype of near real-time satellite assimilation into the WRF model.

The success of SWE in applied science systems will only be achieved through the proliferation of SWE technologies within the science community. Through this project, SMART team members have gained experience with SensorML and other OGC SWE technologies, learning how to describe a standalone executable program with a SensorML process model, and how to write a SensorML description of a chain of such processes. Though the initial learning curve was steep, we expect creation of additional process chains to be much easier after the first one. As these emerging standards come into wider use, we expect user documentation and tools to improve, greatly facilitating sensor web development and process chain composition.

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. This project demonstrates the applicability of SWE technologies far beyond communication with the sensors themselves. Furthermore, the use of standard SWE protocols and formats supports the interoperability of any of these components, facilitating their reuse in a variety of process chains.

5. REFERENCES

[1] J. LeMarshall, F. Weng, S. Lord, L-P Riishojgaard, and P. Phoebus, "The Joint Center for Satellite Data Assimilation," *Preprints, 14th Conference on Satellite*

Meteorology and Oceanography, Amer. Meteor. Soc., CD-ROM, Atlanta, GA, 2006.

[2] 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," *Preprints, 20th Conference on Interactive Information Processing Systems (IIPS)*, Amer. Meteor. Soc., CD-ROM, Seattle, WA, P1.34, 2004.

[3] J. Michalakes, S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, "Development of a next-generation regional weather research and forecast model," *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology*, W. Zwiefnofer and N. Kreitz, Eds., World Scientific, pp. 269–276, Reading, U.K., 2001.

[4] W. C. Skamarock, J. B. Klemp, and J. Dudhia, "Prototypes for the WRF (Weather Research and Forecasting) model," *Preprints, Ninth Conf. on Mesoscale Processes*, Amer. Meteor. Soc., Fort Lauderdale, FL, pp. J11–J15, 2001.

[5] S-H. Chou, B. Zavadsky, G. J. Jedlovec, and W. Lapenta, "Assimilation of Atmospheric Infrared Sounder (AIRS) data in a regional forecast model," *Preprints 14th Conference on Satellite Meteorology and Oceanography*, Amer. Meteor. Soc., CD-ROM, Atlanta, GA, pp. 5.12, 2006.

[6] B.T. Zavadsky, S-H. Chou, G. Jedlovec, and W. Lapenta, "The Impact of Atmospheric Infrared Sounder (AIRS) Profiles on Short-term Weather Forecasts," *Preprints, Proceedings of SPIE*, Orlando, FL, pp. 65651J, 2007.

[7] C. Darden, B. Carroll, S. Goodman, G. Jedlovec, W. Lapenta, "Bridging the Gap Between Research and Operations in the National Weather Service: Collaborative Activities Among the Huntsville Meteorological Community," *NWS Technical Memorandum*, NWS-SR-222, Fort Worth, TX, 2002.

[8] M. Botts, G. Percival, C. Reed, and J. Davidson, "OGC Sensor Web Enablement: Overview and High Level Architecture," *OpenGIS White Paper OGC 07-165*, Open Geospatial Consortium Inc., San Diego, CA, 2007.

[9] R. Ramachandran, X. Li, S. Movva, and S. Graves, "A Simple and Efficient Feature Extraction Algorithm for Geophysical Phenomena," *IEEE Inter. Conference on Geoscience and Remote Sensing Symposium*, Denver, CO, 2006.

[10] A. Robin and M. Botts, "Creation of Specific SensorML Process Models (Draft)," http://vast.uah.edu/downloads/documents/Creating_SensorML_Process_Model_preV1.0.pdf, 2006.