Sensors & Transducers, Vol. 194, Issue 11, November 2015, pp. 123-126

www.sensorsportal.com

Sensors & Transducers

http://www.sensorsportal.com

Near Real-Time Oceanographic Data Management: Latest Developments

Elena PARTESCANO, Alberto BROSICH and Alessandra GIORGETTI

OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale)
Borgo Grotta Gigante 42/C
34010, Sgonico (TS), Italy
E-mail: epartescano@ogs.trieste.it

Received: 31 August 2015 /Accepted: 15 October 2015 /Published: 30 November 2015

Abstract: Data sharing, interoperability and quality check are only a short part of the activities connected to the data management. Create a procedure to harmonize and disseminate heterogeneous data, collected by different meteo-oceanographic buoys in near real-time, is a challenge. In this paper, we describe a system for data validation, conversion in a homogeneous and standard format and dissemination adopting Sensor Web Enablement (SWE) using XML (eXtensible Markup Language) and OGC's (Open Geospatial Consortium) standards. To meet the needs of different scientific communities as RITMARE (La Ricerca ITaliana per il MARE), Jerico (Towards a joint European research infrastructure network for coastal observatories), Copernicus, ODIP (Ocean Data Interoperability Platform) and FixO3 (Fixed-point Open Ocean Observatories), we decided to adopt SWE standard allowing interoperability between data in near real-time, using Sensor Model Language (SensorML) and Observations and Measurements(O&M) standards in a Sensor Observation Service (SOS).

Keywords: Sensor web, Real-time data, Marine observations.

1. Introduction

In this paper, we describe the data management, in the Italian National Oceanographic Data Center, with a short analysis of the importance of role of connection between data providers and users and the relevance of the interoperability that guarantee the success of the system at European level. After this, we illustrate in detail the system developed for managing near real-time data, using seven different steps, starting from the data loader to data visualization, passing through data validation. This system applies an innovative method, using the OGC's Sensor Web Enablement (SWE) standards, using SensorML and O&M standards in a Sensor Observation Service (SOS), that guarantee data sharing, maintaining interoperability and resilience. At the same time, it provides new collected data to national and international communities, as SeaDataNet, RITMARE, JERICO, Copernicus and FixO3.

2. Data Management

2.1. Meaning of Data Management

The Oceanographic Data Management in OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) is managing by the Italian National

Oceanographic Data Center. The data management consists in several procedures to store, validate and disseminate oceanographic data. These procedures follow European Directives and the adoption of common standards and adapted technologies guarantee the interoperability. In this paper we referred at physical data and associated metadata, acquired in near real-time by two meteo-oceanographic buoys.

2.2. Connection Users to Data

The data management is the element between data provider and users.

An important aspect of the data management is the dissemination of data; to realize a system for a near real-time data management is complex. The aim is to allow users to be able to access and use data in real-time. That means to have a system working in time.

In this moment the National Oceanographic Data Center disseminates real-time data using two different systems:

NetCDF data format, by an online repository [1], adopting OceanSites standards and a Sensor Web Enablement (SWE) OGC standards [2].

2.3. Interoperability

Interoperability includes *technical* interoperability (regarding data format and transport format for example) and *semantic* interoperability (for example metadata standards and parameter's vocabularies).

Interoperability is the key to data management system success (SeaDataNet Project):

- Using common vocabularies: this consists of lists of standardized terms that cover several domains. Using standardized expressions resolves the problem of uncertainties associated with data.
- Adopting metadata standard for all metadata: in near real-time data management we use OGC's standards to collect metadata as SensorML and O&M.
- Using harmonized Data Transport Formats for data sets delivery.
- Using common quality control protocols and flag scale: the quality control flag, in this case, is a number associated to each measurement field, whose value grows according to the importance of the failure (0=not controlled, 1=correct, 2=suspect, 3=dubious, 4=wrong, 5=changed, 9=missing).

The quality control procedure includes the following series of automatic checks:

- Missing data and data format completeness;
- Impossible timestamp and measuring position;
- Duplicate vertical profiles or measures;
- Spikes by testing data for large differences between adjacent values;
- Invalid values by comparison with min & max values fixed for each parameter archived.

3. Components of near Real-Time Data Management System

To manage the near real-time oceanographic data, OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) developed a system as an answer for the need to find a standard procedure to share (near) real-time data collected by the meteo-oceanographic MAMBO1 buoy (in the North Adriatic Sea) and the observatory site E2M3A (in the South Adriatic Sea). This procedure has seven different elements [3].

3.1. Methodology

The workflow developed for the data management (Fig. 1) in (near) real-time is based on seven different elements: two meteo-marine stations, RT-Loader (Real-Time Loader), DB-Validator (Database Validator), the RT-Observations (Real-Time Observation), the RT-SOS (Real-Time Sensor Observation Service) using 52°North implementation version 4.2, the RT-Web (Real-Time Web) and the Sensor Web Client [4] .

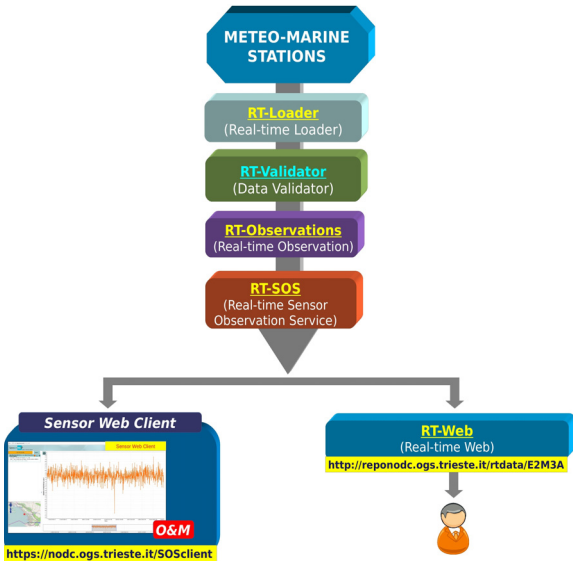

Fig. 1. Workflow.

3.1.1. Devices

Currently the OGS manages two different buoys that acquire data in (near) real-time: the meteo-marine buoy Monitoraggio AMBientale Operativo (MAMBO1), placed in the Gulf of Trieste, equipped with a meteorological station and two multi-parametric probes, and the E2-M3A, situated in the South Adriatic Sea, hosting a meteo station including a radiometer aimed, to collect air-sea interaction measurements, and a mooring with sensors for physical and biochemical parameters.

3.1.2. RT-*Loader*

Real-Time Loader permits to store into a database real-time heterogeneous data, coming from different kind of instruments and with different formats. This is possible by using three main steps:

- Converting heterogeneous ASCII files formats in a unique XML common format,
- Converting XML entities into objects, using Java Architecture for XML Binding (JAXB);
- Inserting of measurements and the associated metadata into a relational database (PostgreSQL).

This workflow needs a supervisor to handle automatically the input files, as soon as they are created or updated. In this system, this role is carried out by a "Apache Camel" application, which checks constantly the presence of new files.

3.1.3. DB-*Validator*

Real-Time Validator is a validation procedure (data quality control procedure) applied to the information to qualify the data values [5], [6], once the data have been included in the database. The procedure has been developed following the European protocols [7] eventually tuned to the regional statistics. As a result of the validation process, a quality flag is defined for all checked information (in the data and in the meta-data) without changing or eliminating any data point.

3.1.4. RT-*SOS*

Real-Time SOS is an OGC's Sensor Observation Service that enables to integrate real-time observations of heterogeneous sensors into a Spatial Data Infrastructure. It works using standard requests (e.g., DescribeSensor()). The descriptions of sensors and observations are stored in a PostgreSQL/PostGIS database using standard metadata format (respectively SensorML and O&M) and standard requests (InsertSensor() and InsertObservation()). Then, the data can be obtained by request GetObservation() and geo-located by GetFeatureOfInterset().

Also a Web interface was used, in order to visualize observations, the sensors position, their observed properties and long term trends of observations. It has been implemented using JavaScript toolkits (OpenLayers, GeoExt and ExtJS).

3.1.5. RT-*Observations*

Real-Time Observations is a real-time automatic procedure that permits to insert Observations into the Sensor Observation Service. In the latest development, this new element can load in near real-time data into a SOS, using an OGC standard format O&M. This is possible applying a Batch() operation and a POST request using JSON as payload.

3.1.6. RT -*Interfaces*

Real-Time Interfaces are the tools used to allow users to access and use data in real-time. Currently, two different systems are used to visualize and download data: an online repository [1], where it is possible to download data in Netcdf data format adopting OceanSites standards, and a Web Client (52° North) (Fig. 2) through it is possible visualize and download data [2].

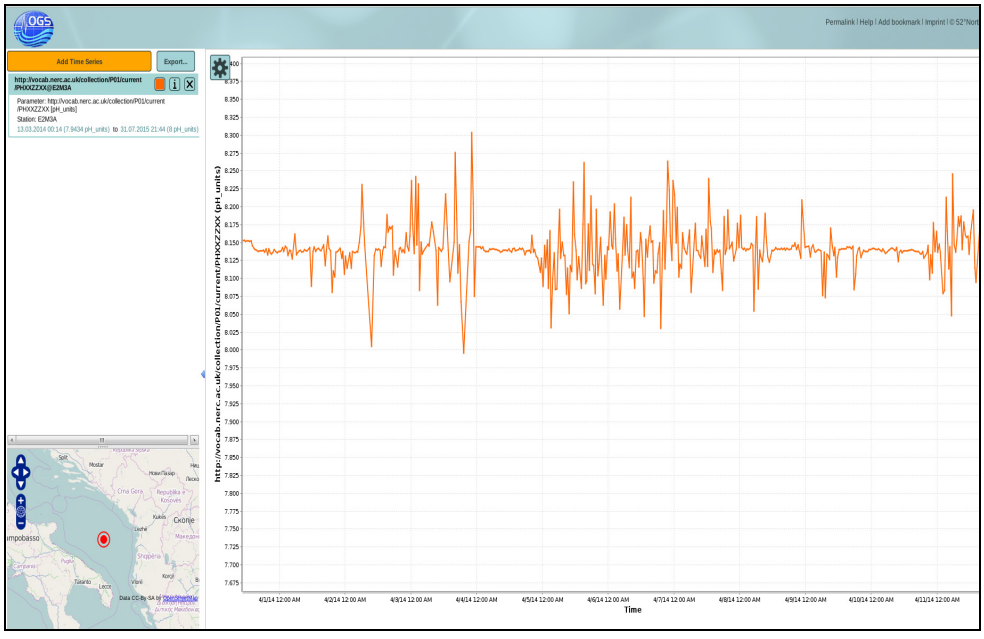

Fig. 2. SOS Web Client.

4. Conclusions

Usually a user expects to obtain immediately the data selected from a web interface. The objective of a data manager is to realize a working system capable to answer to the user needs. In the Italian National Oceanographic Data Center we developed a procedure to harmonize and disseminate heterogeneous data, collected by different meteo-oceanographic buoys in near real-time. We decided to adopt SWE (Sensor Web Enablement) standards allowing interoperability between data in near real-time, using SensorML (Sensor Model Language) and O&M (Observations and Measurements) standards in a Sensor Observation Service.

Acknowledgements

This work was supported by the Ministry of Education, Universities and Research (MIUR) under RITMARE Project. We also acknowledge funding from the European Programme (JERICO, FixO3 and SeaDataNet II). The authors are thankful to Paolo Mansutti and Giuseppe Siena for technical support in SensorML developing.

References

[1]. Online repository: OGS's Online data repository: (http://reponodc.ogs.trieste.it/rtdata/E2M3A/).

[2]. RT-WEB: OGS's Web interface: (http://nodc.ogs.trieste.it/SOSclient/)

[3]. M. Ravaioli, *et al*., The RITMARE Italian Fixed-point Observatory Network (IFON) for marine environmental monitoring: a case study, in *Proceedings of the 3rd Italian GNOO Conference on Operational Oceanography, Innovative Technologies and Applications*, *Journal of Operational Oceanography*, Special Issue, 2015 (Submitted).

[4]. E. Partescano, *et al*., Near Real-Time Oceanographic Data Management through Sensor Web Enablement (SWE) Standards, in *Proceedings of the 9th International Conference on Sensor Technologies and Applications (SENSORCOMM'15),* Venice, Italy, 23-28 August 2015, pp. 103-104.

[5]. A. Giorgetti, *et al*., OGS oceanographic data archiving and validation system: the IOC/National Oceanographic Data Centre, *Boll. Geof. Teor. Appl.*, Vol. 48, Issue 3, 2007, pp. 359-369.

[6]. A. Giorgetti, *et al*., Maritime data standardization in the Archimede Project, *Boll. Geof. Teor. Appl.*, Vol. 48, Issue 3, 2007, pp. 305-317.

[7]. UNESCO: GTSPP Real-Time Quality Control Manual, First Revised Edition, IOC Manuals and Guides, No. 22, Revised version 1.0 Ostend, 3 December 2010, (IOC/2010/MG/22Rev.).