A (Near) Real-time Validation and Standardization System Tested for MAMBO1 Meteo-marine Fixed Station

Elena Partescano[1], Alessandra Giorgetti[1], Caterina Fanara[1], Alessandro Crise[1], Alessandro Oggioni[2], Alberto Brosich[1] and Paola Carrara[2]

[1]*OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) Borgo Grotta Gigante 42/c, 34010 Sgonico, Trieste, Italy*

[2]*CNR IREA UOS Milano, Via Bassini 15, I20133 Milano, Italy*

Keywords: Sensor Web Enhancement, Near Real-time, Meteo-oceanografic Buoy.

Abstract: The objective of this paper is to describe a new application developed to deliver validated data in (near) real-time from marine stations, together with a complete set of information. The harvesting of marine (near) real-time data with multiple data formats, the conversion in a homogeneous and standard format, the structuring in a database and, finally, the automation of the marine data validation is obtained using XML and OGC (Open Geospatial Consortium) standards for data transport and representation. The adoption of Sensor Web Enablement (SWE) specifications enables real time integration of data and metadata, related to the data processing and calibration, the data collection instruments and the data quality control. Our technological choice is led by the requirements of interoperability, as ability to cooperate and exchange information, and resilience, as ability of adaptation to new needs. The international standard SensorML has been used as a profile, adapted to our needs and results as a joint effort of the Italian RITMARE, the European SeaDataNet, Eurofleet and ODIP community.

1 INTRODUCTION

This work started from the need to archive, validate and deliver data in (near) real-time from marine stations. A series of difficulties has been addressed because the marine data are heterogeneous, being collected by different sensors and with different data formats. Furthermore, the data validation procedure should be able to apply a quality control to the single measurements, eventually by comparison with climatological mean values and assigning a data quality flag without modifying the measured data. Finally, all information needs to be stored in a database. The solution presented in this paper has been developed using open technologies and standards with the objective to develop a flexible tool, easily adaptable to the new needs.

The complete system consists of several elements: a meteo-marine fixed station called MAMBO1, the loading software called RTLoader (Real-Time Loader), the validation software called DBValidator (Database Validator) and the publishing software composed of different components namely the RTWs (Real-Time Web Service), the RTWeb (Real-Time Web interface) and the RTSOS (Real-Time Sensor Observation Service).

The paper is organized as follows: section 2 describes in detail the devices included in the system and the data collected and managed; the methodology adopted for data storage, processing and interoperable delivery is illustrated in section 3, where system's components and their functions are described in detail. The last sections close the paper describing the results and with some concluding comments and working perspectives.

2 DEVICES AND DATA

The meteo-oceanographic buoy named "MAMBO1" (Monitoraggio AMBientale Operativo), located in the Gulf of Trieste, is the first example of meteo-marine coastal station installed on a buoy in the northern Adriatic Sea, designed to acquire and transmit in (near) real-time quality measurements of key meteorological and oceanographic variables. Other stations of the same series were deployed in

Partescano E., Giorgetti A., Fanara C., Crise A., Oggioni A., Brosich A. and Carrara P..
A (Near) Real-time Validation and Standardization System Tested for MAMBO1 Meteo-marine Fixed Station.
DOI: 10.5220/0004848504150420
In *Proceedings of the 3rd International Conference on Sensor Networks* (MOEOD-2014), pages 415-420
ISBN: 978-989-758-001-7

Northern Adriatic, Ligurian Sea and Sardinian waters.

The MAMBO1 buoy has been operating since the end of 1998 and was originally designed and implemented by OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) for the environmental monitoring of the marine protected area "Miramare Marine Reserve". It is moored at the edge of the reserve (Coordinates: Lat. 45°41.86'N, Long. 13°42.50'E) at 18 m depth, about 300 m from the coast.

The model of hull, the solar panel energy system, the buoy controller, the electric wiring system and the mooring scheme has been developed with proprietary technology of OGS.

The MAMBO1 buoy is equipped with a meteorological station (R.M. YOUNG Wind Monitor-MA) mounted on the tripod that determines the following parameters: air temperature and humidity, wind speed and direction, barometric pressure and solar radiation.

A multi-parametric probe (SBE 16plus V2 SeaCAT Recorder) is installed at 10 m depth, for the study of the main physical-chemical parameters (pressure, water temperature, conductivity, dissolved oxygen, fluorescence, pH, chlorophyll &turbidity and radiation). Temperature, conductivity and pH sensors are routinely calibrated at the OGS oceanographic calibration laboratory following procedures, developed by the Calibrations & Testing Operations group (CTO group), that are compliant with the international standards of excellence.

During 2012, a second probe (CT SBE 37) has been installed inside a cage at 15 m depth, together with a pCO2 sensor (ProOceanus PSI CO2-ProTM) and a pH sensor (Sunburst SAMI2-pH). The system has been tested and, after improvements, it should be redeployed within the end of 2013.

Since the beginning, the MAMBO1 buoy has undergone several changes through the implementation of instruments, also thanks to the contribution of several international projects (JERICO, http://www.jerico-fp7.eu/ and EUROSITES, http://www.eurosites.info/).

The data are acquired twice per hour and instantaneously transferred via GSM modem to the shore-based receiving station. They are archived at the OGS' National Oceanographic Data Centre (NODC-OGS) and can be accessed at the web page http://nettuno.ogs.trieste.it/mambo/. The historical time series records are available since 1999 while the bio-geochemical data records started in 2012.

Meteorological data are also accessible through the portal of the European initiative "EMODNET Physical Parameters": http://www.emodnet-physics.eu.

From January 2013, the MAMBO1 buoy has been included among the infrastructures for the "Service and Data Access" activity, within the EU FP7 project JERICO. Under this core activity, OGS will provide free access to the observations and well referenced metadata coming from the MAMBO1 buoy, for a two years period, through MyOcean INS TAC Portal for the Mediterranean Sea (http://www.myocean.eu/web/69-myocean-interactive-catalogue.php). Depending on the specific usage, data are provided to users in real time or in delayed mode, following a data assembly process that is targeted to be compliant with SeaDataNet standards and MyOcean requirements. The format in use for the data delivery is NetCDF, i.e. OceanSites *de-facto* standard.

3 METHODOLOGY

The working flow developed for the data management in (near) real time at shore is based on five different elements (Brosich et al., 2013): RTLoader (Real-Time Loader), DBValidator (Database Validator), the RTWs (Real-Time Web Service), the RTWeb (Real-Time Web) and the RTSOS (Real-Time Sensor Observation Service) using 52°North implementation (http://52north.org/) version 3.2.

RTLoader (Fig.1) has the task to store in a database real-time heterogeneous data, coming from different kind of instruments and with different formats.

DBValidator checks the quality of the data, applying some different algorithms.

RTWs is the RESTful Web Service used to extract data from the database.

RTWeb is the web interface that allows querying the database using the Web Service RTWs. It extracts data into a downloadable file, satisfying the conditions selected by the users.

Finally, RTSOS is a OGC (Open Geospatial Consortium) SOS service that enables to integrate real-time observations of heterogeneous sensors into a Spatial Data Infrastructure. It is fed by data coming from RTLoader; specifically, the conversion of the input data (included into the RTLoader) is made by an open source java library "ServingXML" that allows to read and translate using the directives inside the XML files (one for each input file format). This conversion generates a new XML file following the "Observations and Measurements" (O&M) OGC

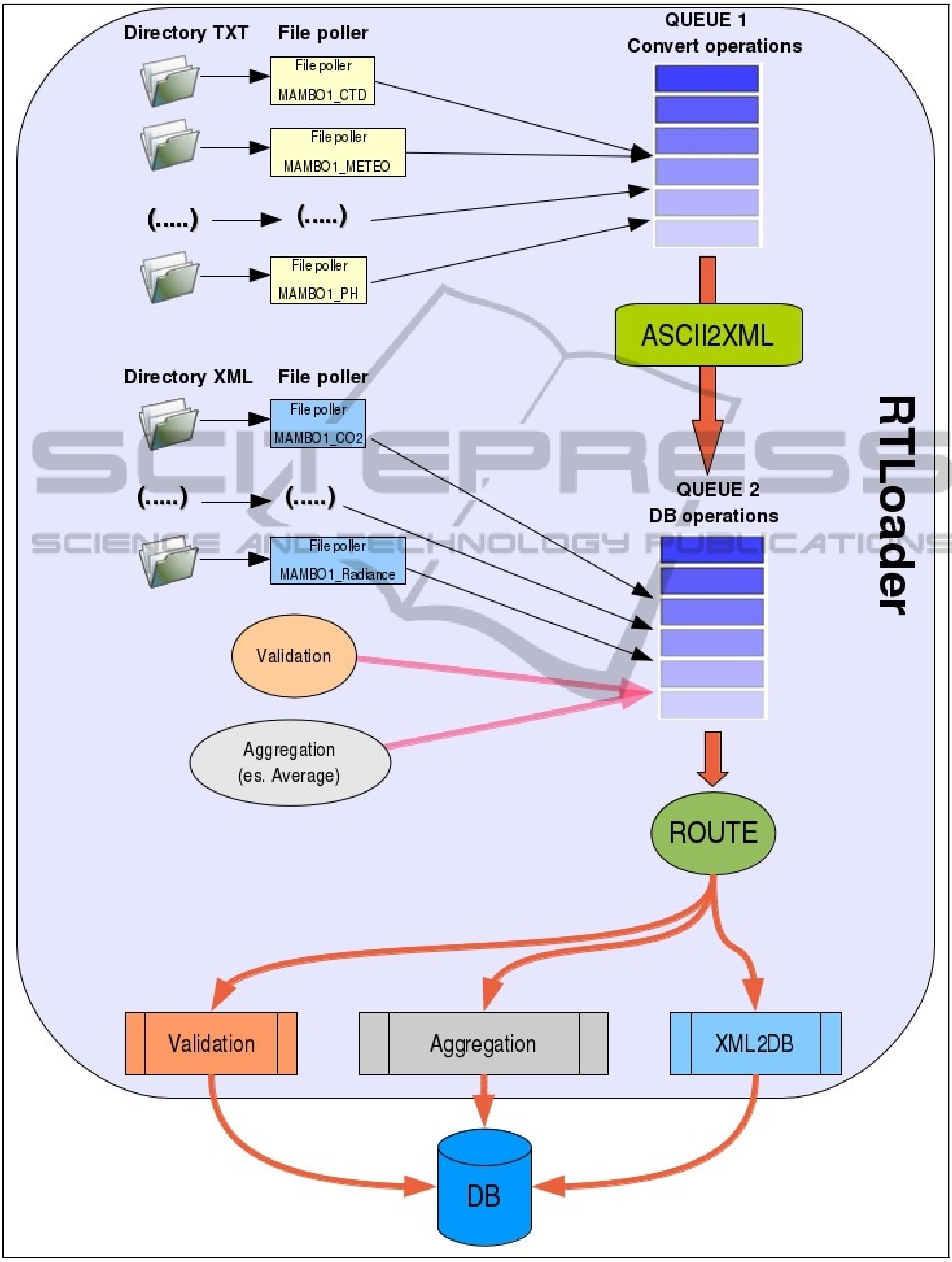

Figure 1: Figure showing the internal workflow of the RTLoader component.

schema. We have decided to adopt these schemata because they allow to fully describe data from all sensors of our interest and to fulfil standardization requirements enabling interoperability among different data providers.

The use of Java language guarantees the platform independence.

The resilience of RTLoader application is obtained adopting a new approach using "Apache Camel" as rule-based routing and mediation engine operating in an event driven way, decoupling software services, in a Java environment. This framework is responsible for the quality of the service features such as message persistence, guaranteed delivery, failure handling, and transaction support.

The data structure coming from the storage in a relational database allows wide data-warehouse and analysis (data-mining) opportunity, granting capability to reach a wide end-user-needs spectrum.

Inspired by its European working experience the Italian NODC is building, in collaboration with other OGS groups, this technological infrastructure. At the moment the work is following the needs coming from the management of the already existing meteo-marine monitoring network of the Regional Civil Protection. The data comes from a wide range of instruments like: meteo-oceanographic buoys with a meteorological station and a CTD profiler, directional waverider buoys with a satellite positioning systems. We aim to increase the data types covered. Within this perspective, interoperability of the infrastructure is a pillar for the system development.

3.1 Database Structure

The relational database used to store observation from the buoy contains several tables (more than 30).The structure of the database was branched in three parts:

1. The sector dedicated to describe the instruments, their characteristics and their position (DEPLOYMENT);
2. The section used to store data and related metadata (MEASURES);
3. The part reserved to collect the vocabularies used to standardize data and metadata and the quality control (COMMON VOCABULARIES and QUALITY CONTROL).

3.2 RTLoader

The Java framework "Apache Camel" guarantees the possibility to manage the sequences, the queues and the end of the process.

Starting from XML standard, defined by the "JAXB" (Java Architecture for XML Binding) application that convert the XML files in Java classes, we have the opportunity to move toward an object language.

This permits to move from a multitude of files formats (ASCII) into a unique common format by defining an XML file for each file format. The framework "ServingXML" is used to re-build the data into a standardized XML file using an appropriate schema.

This XML file, that represents the first step for inserting data, is converted by "JAXB" into an object language. The framework JAXB transforms the XML input file into java objects.

Therefore, the software "XML2DB" converts XML input into java objects and then inserts the measurements and the associated metadata into the database. Before storing data, it obtains some information from the database about the position and the instrument used.

This workflow needs to be checked by a controller that permits to manage automatically the input files, as soon as they are created or updated. In this system, this role is carried out by "Apache Camel", which checks constantly the presence of new files (or eventually the presence of updated files) and launches the conversion and the insertion of data into the database.

3.3 DBValidator

Once the data have been included in the database, a validation procedure (data quality control procedure) is applied to the information to qualify the data values (Giorgetti et al., 2007a), (Giorgetti et al., 2007b). The procedure has been developed following the European protocols (SeaDataNet, 2010), (UNESCO, 2010) eventually tuned to the regional statistics (Manca et al., 2004). 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 points. The quality control flag (UNESCO, 1990) 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 implemented for Mambo meteo-oceanographic data includes the following series of automatic checks:

- checks for missing data and data format completeness;

- check of the date/time and of the measuring position;
- check of duplicate vertical profiles or measures;
- check for spikes by testing data for large differences between adjacent values,
- check for invalid values by comparison with min & max values fixed for each parameter archived.

These checks are implemented at fixed time intervals to series of three data points and aim at highlighting:

- the presence of casual errors, that may be due to unhappy manual operations, to voltage drop during measurements, to data transmission problems, to sensor calibration problems, to ordinary or extraordinary maintenance operations;
- the presence of systematic errors, that may be due to changing of measuring routine operations, changing of instrumentation, changing of the measuring site or changing of environmental conditions at the site.

3.4 RTWs and RTWeb

RTWs is a RESTful Web Service that accepts simple requests to extract the data from the database. These requests can be parameterized with temporal range and the output format. The Web Service interface RTWeb can be accessed at the URL "http://nodc.ogs.trieste.it/rtws/application.wadl". It is written using Java and open source libraries like Spring and Jersey.

The Web interface allows the selection of data parameters. It is mandatory to specify the temporal interval. The Web interface collects the criteria specified by the user, and calls the Web Service RTWs to create the result with the data matching the user' criteria.

3.5 RTSOS

Recently, real-time observations of few sensors have been published by an OGC SOS and descriptions of different sensors can be discovered or compared using OGC SensorML standard requests (e.g. DescribeSensor()). Observations are stored in a PostgreSQL/PostGIS database, they can be obtained by standard requests (GetObservation()) and geo-located by GetFeatureOfInterset(). Finally, in order to access observations,an SDI client displays on a Web interface the sensors position, their observed properties and long term trends of observations. It has been implemented using JavaScript toolkits (OpenLayers, GeoExt and ExtJS).

A similar approach has been adopted by other Italian research organizations managing marine observation from fixed stations and proved to be feasible and promising (Oggioni et al., 2013).

4 RESULTS AND CONCLUSIONS

MAMBO1 meteo-marine fixed station, managed by OGS, is operating since 1998. It has been recently included into the joint European research infrastructure network for coastal observatories, giving a strong boost to the harmonization of the real-time validation and standardization procedures. The main objective of the approach presented in this paper is data sharing. The added value is the complete description of the operations carried out on the sensors and the data themselves (both in the acquisition phase and in the subsequent management of information with quality control). Our technological choice is led by the requirements of interoperability and resilience. A software package including the Sensor Observation Service (SOS) installed for selected sensors of the (near) real-time monitoring system has been developed within the RITMARE Italian Flagship and installed at OGS premises. The SensorML used is compliant with the international standards and developed results as a joint effort of the Italian RITMARE, the European SeaDataNet, Eurofleet and ODIP community. The adoption of SWE (Sensor Web Enablement) standards allow as a direct consequence the integration of platforms to be effectively managed in a multi-channel and multi-platform way. The resulting data set is publicly available and can be progressively integrated into a distributed infrastructure, capable to manage real-time data acquired by different observation systems (from fixed sites for sea observation, multiplatform observatories, autonomous tools, to eventually remote sensing data).

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Education, Universities and Research (MIUR) under RITMARE Flagship Project. We also acknowledge funding from the European Union Seventh Framework Programme (grant agreement n° 123456 JERICO and grant agreement n° 283607 SeaDataNet II) and from the Friuli Venezia Giulia Region – Regional Civil Protection (contract n° 17/CD2/2009

PALME4). Operations at sea are coordinated and constantly performed by OGS TecDev Research Unit, while the maintenance of the instruments is followed by the Calibrations & Testing Operations Group (OGS CTO Group).

REFERENCES

Brosich A., Altenburger A., Giorgetti A.: *Software RTLoader - Manuale Utente.* Rel. 2013/43 Sez. OCE 21 NODC.

Giorgetti A., Brosich A. and Mosetti R.: OGS oceanographic data archiving and validation system: the IOC/National Oceanographic Data Centre, Boll. *Geof. Teor. Appl*., 48 (3) (2007), pp. 359-369.

Giorgetti A., Inghilesi R., Orasi A., Morucci S., Bruschi A.: Maritime data standardization in the Archimede Project, Boll. *Geof. Teor*. Appl., 48 (3) (2007), pp. 305-317.

Manca B., Burca M., Giorgetti A., Coatanoan C., Garcia M.J. and Iona A.: Physical and biochemical vertical profiles in the Mediterranean regions: an important tool to trace the climatology of water masses and to validate incoming data from operational oceanography. *J. Mar. Sys*., 48 (2004), pp. 83-116.

Oggioni A., M. Bastianini, P. Carrara, F. Pavesi, "Condivisione di dati real time tra sensori di Istituti del CNR: un approccio possibile", *17a Conf. Naz. ASITA*, Riva del Garda, 5-7 novembre 2013, pp. 1047-1054.

SeaDataNet: *Data Quality Control Procedures,* Version 2.0, May 2010, 78 pp.

UNESCO: GTSPP Real-Time Quality Control Manual, IOC Manuals and Guides n. 22. In*: Manual of Quality Control Procedures for Validation of Oceanographic Data.* CEC: DG XII, MAST and IOC: IODE Manual and Guides n. 26 (1990), pp. 305-432.

UNESCO: GTSPP Real-Time Quality Control Manual, First Revised Edition. *IOC Manuals and Guides No. 22, Revised Edition.* (IOC/2010/MG/22Rev.) (2010), 145 pp.