CNR-ISMAR in situ observations network: new approaches for an interactive, high performance, interoperable system

Stefano Menegon, stefano.menegon@ismar.cnr.it
Pierluigi Penna, pierluigi.penna@ismar.cnr.it
Mauro Bastianini, mauro.bastianini@ismar.cnr.it
Giuseppe Stanghellini, giuseppe.stanghellini@ismar.cnr.it
Francesco Riminucci, francesco.riminucci@ismar.cnr.it
Tiziano Minuzzo, tiziano.minuzzo@ismar.cnr.it
Alessandro Sarretta, alessandro.sarretta@ismar.cnr.it
CNR-ISMAR (Italy)

This paper describes the activities in ISMAR (Institute of Marine Sciences[1] of the Italian National Research Council) for the creation of an in-situ observational network able to: gather data from a distributed network; allow an efficient visualization of time-series data; access and share harmonised data through interoperable and standard services.

Figure 1: Web interface of the portal

Context and objectives

During the last years, the need for a unique visible access point for the ISMAR observational network emerged from various research projects and, in particular, in the projects RITMARE, LTER and JERICO. This is also related to the request, by various European-funded projects, that data should be made discoverable, accessible and sharable using standards formats and services (NetCDF, ODV, SOS…).

More in general, the ISMAR observational network aims to ensure archiving and storage of the historical time series data of the Institute (useful to long-term research and climate changes research) and to provide an efficient and convenient support for the operational model (through real-time data streams). These goals are achieved through harmonization of data management, using standard formats and interoperable services and creating a unique access point (web portal) for all ISMAR data related to in-situ observations, where to visualize and access real-time and archived data; in addition

1 http://www.ismar.cnr.it/index_html-1?set_language=en&cl=en

to that, a clear data policy and standard open licences are adopted to facilitate dissemination and reuse of data.

ISMAR manages various marine/atmospheric measurement sites that have had different implementation paths (funds, specialised personnel, IT infrastructures) and that now need to be commonly organised and managed to allow a better use and dissemination of the data collected.

Currently, 5 stations for marine/meteorological observations (Acqua Alta, E1, GB, Telesenigallia, Gargano) and 3 meteorological stations (ISMAR-Venezia, ISMAR-Ancona, ISMAR-Lesina) are active, while 4 more are in plan to be included (buoys and moorings), with a clear focus on the Adriatic Sea region. In addition to standard meteorological parameters, the system collects data about the water column (i.e. temperature, salinity, sea level, oxygen, waves, currents and other water quality parameters).

Other platforms or acquisition systems might be added in the future to extend the geographical scope and the availability of environmental parameters observed.

Architecture and IT solutions

The adopted solution aims to combine approaches based on interoperability standards (i.e. SensorML) with systems that are optimized for handling time series data.

The core of the system is a centralized database implemented with InfluxDB, an open source database created specifically to handle time series data with high availability and high performance requirements. InfluxDB[2] has several interesting features like a simple, high performing write and query HTTP(S) API, a flexible and dynamic data model and is part of a stack of integrated components and tools that allow to implement a complete solution for collecting, storing, visualizing and alerting on time-series data. The Graphical User Interface (GUI) is based on Grafana[3], an open source, feature rich metrics dashboard and graph editor.

Data are collected from the measurement stations and stored without further modification in a “raw data” database. At the same time, a data processing engine computes an automatic data quality control in order to clean the data, underlying statistical anomalies and, if necessary, triggers alerts. The “cleaned” data are stored in a “higher quality” database.

InfluxDB and Grafana are commonly used as part of the ICT network and infrastructure monitoring system and, if on the one hand they are simple to use and very efficient, on the other they do not offer some basic features for an environmental and scientific monitoring network (i. e. Interoperability services like THREDDS and SOS, interchange formats like NetCDF, rigorous description of metadata).

In order to handle such issues we have, first of all, adopted already existing conventions and best practices (i.e. the use of SeaDataNet Agreed Parameter Groups[4]); then, we have extended the platform with specific functionalities and tools, for instance a plugin to export data in NetCDF format (i.e. SeaDataNet NetCDF - CF) and a plugin to harvest and import data from SOS (Sensor Observation Services) servers.

2 https://influxdata.com/time-series-platform/influxdb/

3 http://grafana.org/

4 http://vocab.nerc.ac.uk/collection/P03/current/