



## SUB-SURFACE UNDERWAY DATA VALIDATION AT IEO DATA CENTER

Andrea Pannocchi, Laboratory of Experimental Oceanology and Marine Ecology, Tuscia University, [a.pannocchi@unitus.it](mailto:a.pannocchi@unitus.it)  
 Lucía Arranz, Spanish Institute of Oceanography (IEO), [lucia.arranz@md.ieo.es](mailto:lucia.arranz@md.ieo.es)  
 Alberto Hernández, Marine Technology Unit (UTM/CSIC), [ahernandez@utm.csic.es](mailto:ahernandez@utm.csic.es)  
 Elena Tel, Spanish Institute of Oceanography (IEO), [elena.tel@md.ieo.es](mailto:elena.tel@md.ieo.es)  
 Amaia Vitoria, Spanish Institute of Oceanography (IEO), [amaia.vitoria@st.ieo.es](mailto:amaia.vitoria@st.ieo.es)

### INTRODUCTION

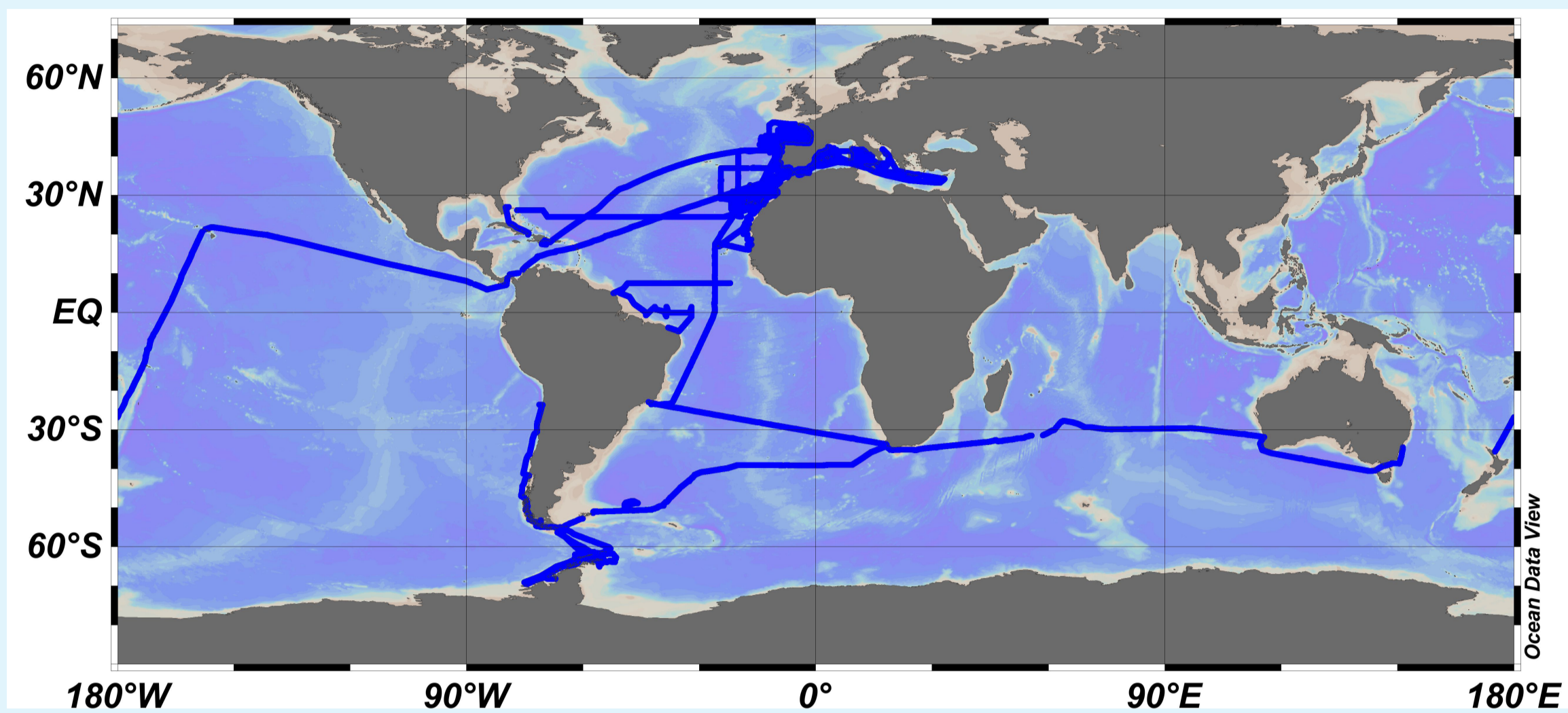
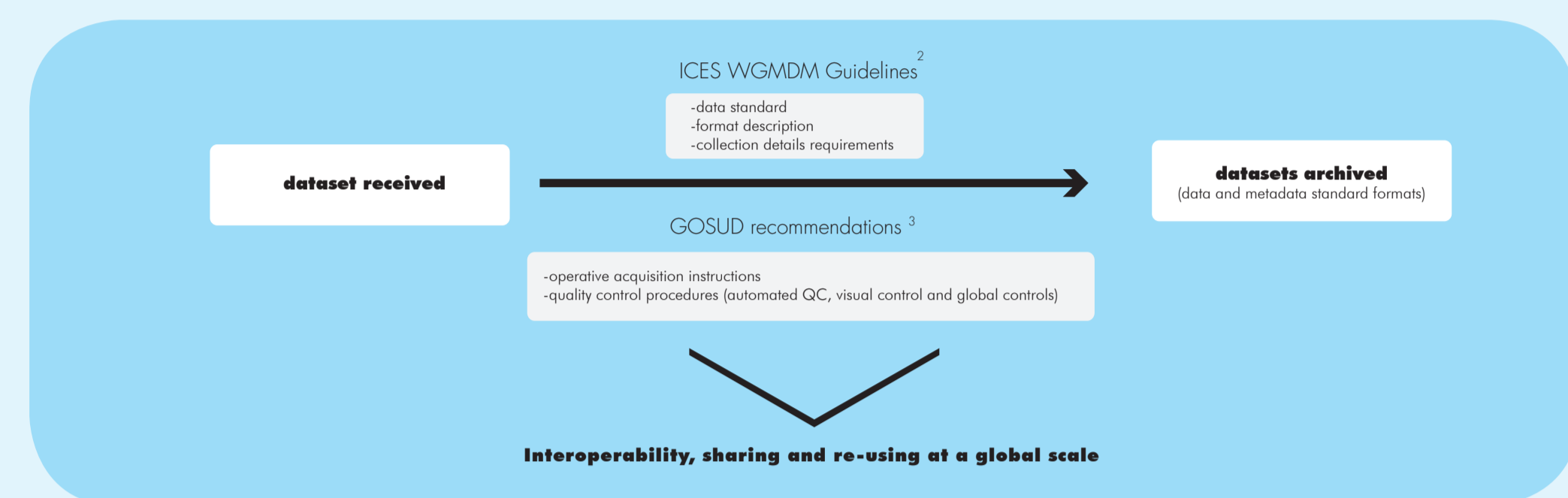


Fig.1. Global trajectories of marine Spanish research vessels

Almost the entire Spanish Research Vessels fleet, together with VOSS (Voluntary Observing Ship) and SOOPs (Ship of Opportunity Program), are equipped with automatic sampling devices able to collect physical, chemical and biological data along the survey routes. On board thermosalinographs (TSG) measure the sub-surface seawater temperature and salinity. As

the international scientific community welcome such instrumentation on board<sup>1</sup>, data quality control is essential to ensure an international acceptance and to detect deviations and/or errors in the sensor operations, missing information and duplicates. At this work there are presented the procedures to check and validate.

### VISUAL QUALITY CONTROL PROCEDURES



**IEO DATASETS.** Quality control is realised using the MATLAB tool procesaTSG<sup>4</sup> developed at the IEO following GOSUD's (Global Ocean Surface Underway Data) protocols that is continuously improved. These routines include:

- automated quality control of dates
- global and regional positions
- temperature and salinity ranges duplicate control
- detection of constant values
- velocity test<sup>5</sup>

Daily data are published online for operational purposes. The high facilities of visualization are used to:

- validate the data by checking the vessels tracks
- eliminate stabilization time periods and spikes
- check gradient variations
- compare with average values

**UTM DATASETS.** Due to the information provided by UTM (metadata and ancillary data which includes cruise events, navigation, track map, etc), datasets were checked for validation following this procedure:

- NEMO<sup>6</sup> software: to obtain a dataset coupled with metadata and for parameter code matching with SDN common vocabularies.
- ODV<sup>7</sup> software: for pre- and post-control procedures, data representation and quality flag management. ODVs visualization gives a good understanding of the overall trends of the sampled parameter along each track, revealing their latitudinal and circadian gradients.
- Navigation file: for vessel speed checking (ad-hoc Matlab routine), in order to delete all that data corresponding to a vessel velocity less than 1 kn.
- NOAA World Ocean Database: to verify doubtful data against climatological reference values.

The scientific validation is supported by information provided by the water profiles obtained in the IEO monitoring programs and oceanographic cruises.

### DATASETS

**IEO DATASETS** started in 2002 and cover the Spanish coasts and adjacent seas. Five research vessels: Cornide de Saavedra, J.M. Navaz, J. Rioja, R. Margalef and A. Alvaríño, collect the data from a TSG device. **UTM DATASETS** coming from surveys activities of the Sarmiento de Gamboa and Hespérides research vessels spanning from year 2008 to 2011, undertaken around the Spanish coasts, and along ocean routes (North and South Atlantic, Pacific and Caribbean Sea).



Fig.2. Spanish research vessels. Above: Cornide de Saavedra. Below: Sarmiento de Gamboa.

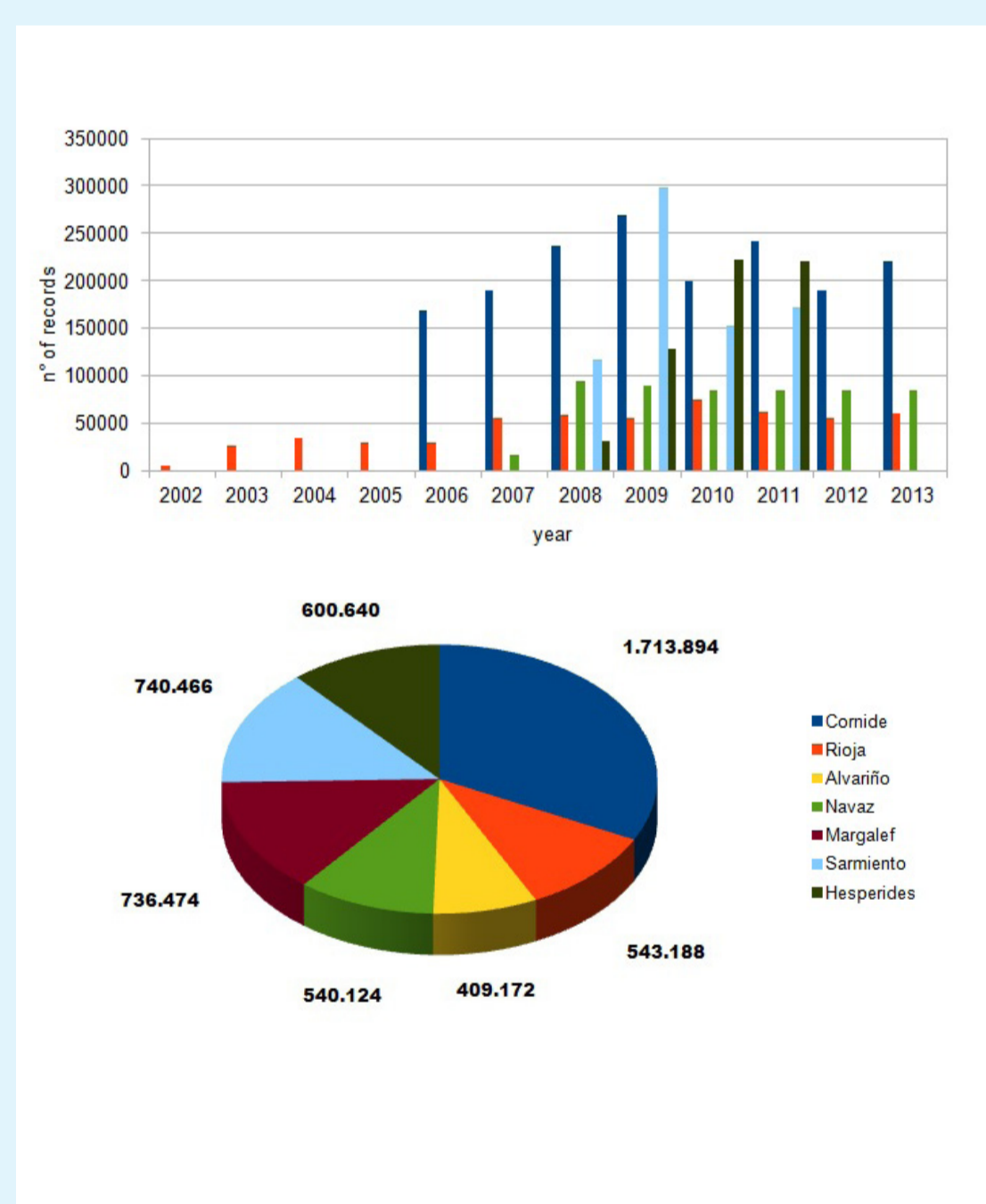


Fig.3. Total amount of data collected by Spanish research vessels. Above: records distribution per research vessel. Below: annual records distribution.

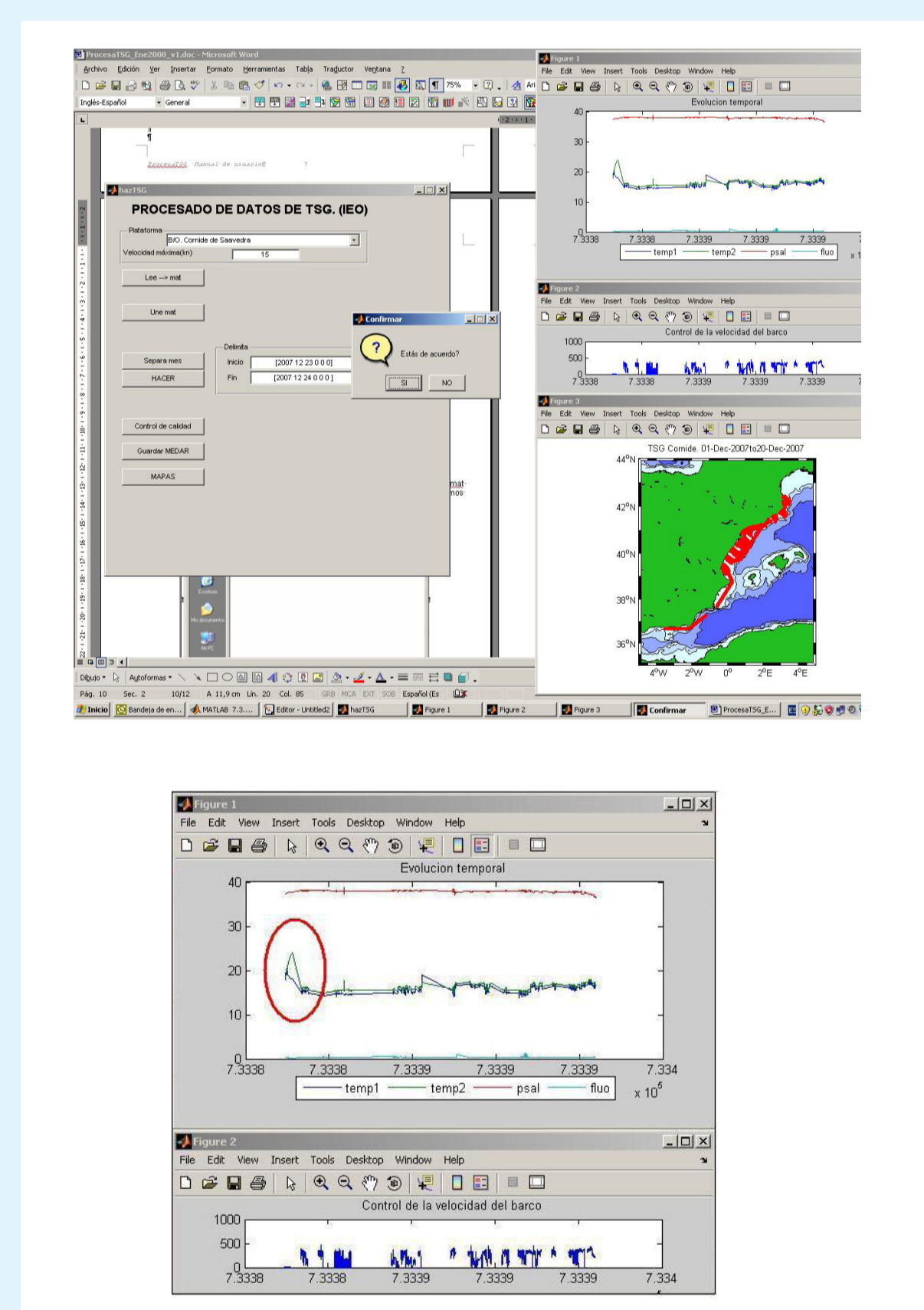


Fig. 4. Sensor stabilization and vessel velocity, output from procesaTSG. In some cases anomalies are found. The temperature spike marked red is probably due to the sensor stabilization or the overheating of the seawater at the pipe intake.

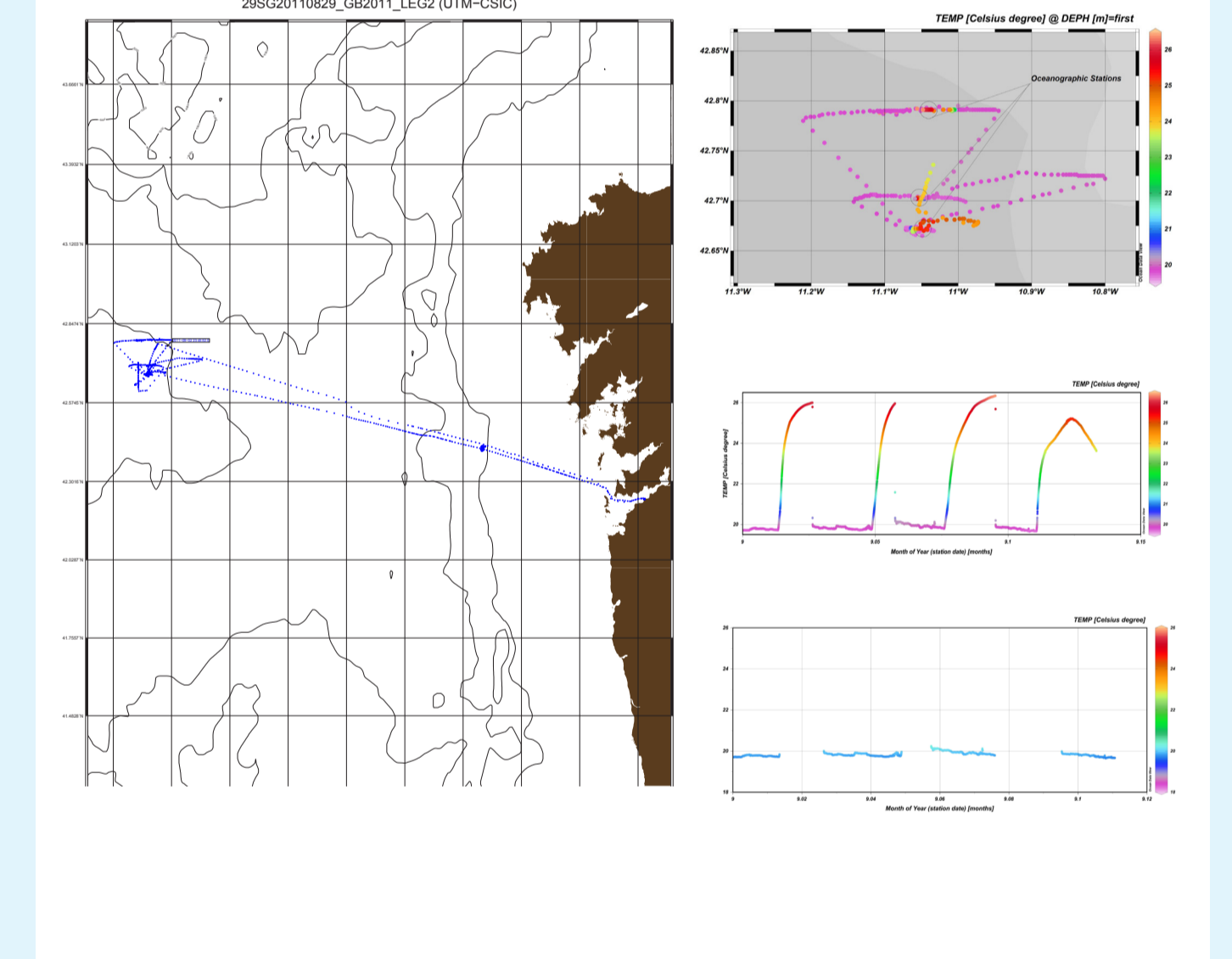


Fig.5. ODV data visualizations show bad values of temperature measurements (SDN QC=4) each time the vessel is performing oceanographic stations. It is possible to filter the data and visualize just the corrected data which show the temperature variation just along of the track.

### CONCLUSIONS

Good quality data coming from continuous measurements give valuable information to complete the world ocean datasets.

The two procedures showed along this work achieve similar results in systematic errors detection and validation of the final datasets, although the output formats are different: ODV-SDN and MEDATLAS.

IEO procesaTSG MATLAB tool:

- was developed to complement the IEO QC/DAMAR quality control system (EU MEDATLAS/MEDAR) which didn't include trajectories.
- allows a high accuracy control system and realize automated quality data control. Also it is useful as a feedback input for the equipment operators and the quality control data team in order to realize even more appropriated operating procedures. Next step of our data quality control procedures, will take into account the comparison with the TSG-QC quality control software of the Institut de recherche pour le développement (IRD).

ODV software:

- allows good and flexible visualizations, due to its QC managing system.
- The simultaneous utilization of both ODV and MATLAB tool procesaTSG, allowed the anomalies detection in such large datasets.

The control procedure highlights the convenience of a careful planning for the sampling procedures. Furthermore, in some cases, it should be recommendable to turn off the instrument during a fixed sampling station, especially in estuarine areas. Enclosed maps with the sample stations points coordinates and event annotations could be useful to better understand data anomalies.

The assembly of the two dataset after this validation improves the spatial coverage and allows to elaborate spatial distribution maps for the regions with enough data. An example is given in the following figures. Such maps are obtained after data quality control procedures and ODV visual control validation.

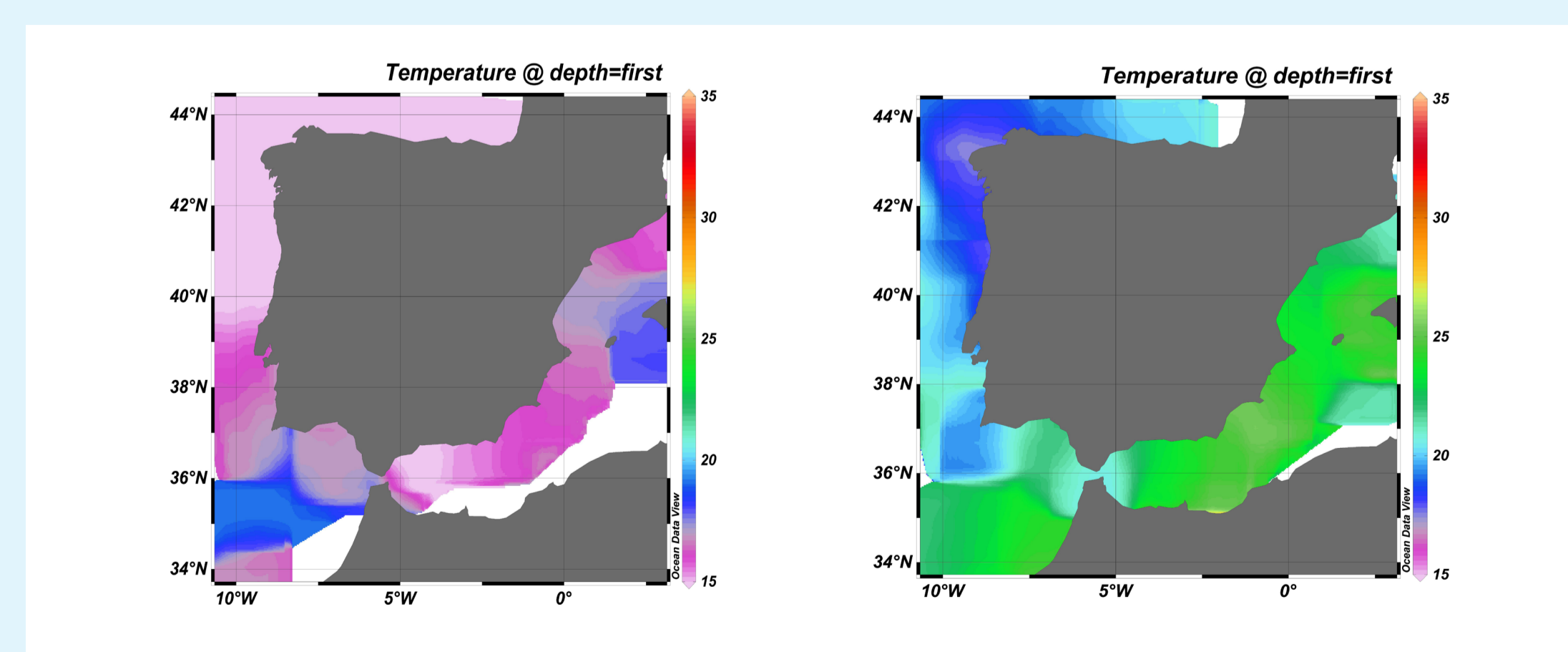


Fig.6. Seasonal situation of Spanish coastal sea temperature distribution. Left: winter. Right: summer.

References:  
 1. Marcelli, M., Pannocchi, A., Piermattei, V., Mainardi, U. 2012. New technological developments for oceanographic observations. *Oceanography*, InTech Publ. ISBN 979-953-307-541-8.  
 2. Guidelines for Surface Underway Data. ICES WGMMD (2001).  
 3. IOC Manuals and Guides. No. 22, Revised Edition. IOC/2010/MG/22Rev.  
 4. Tel, E., Arranz, L. 2008. ProcesasTSG: Análisis de Datos de Termosalinógrafo. Aplicación Informática en Matlab. Manual de Usuario. IEO, UNESCO-IOC 2010. QTSPP Real-Time Quality Control Manual, First Revised Edition.

5. Vitoria, A., Tel, E., González-Pola, C., Marino, A., Reguera, I., Rodríguez-Puente, C., Lavín, A. 2012. 10 years of TSG records on the Cantabrian Coast. *Proceedings 7th Spanish-Portuguese Assembly for Geodesy and Geophysics*.  
 6. Fichaut, M., Bregent, S. 2012. NEMO 1.4.4. User's Manual, IFREMER.  
 7. Schlitzer, R., Ocean Data View, <http://odv.awi.de>, 2011.