ISSN 2039-7941
Anno 2014_Numero 274

rapporti tecnici INGV

Quality control system of the Mediterranean Forecasting System products

274

Istituto Nazionale di
Geofisica e Vulcanologia

ISSN 2039-7941
Anno 2014_Numero 274

rapporti tecnici INGV

QUALITY CONTROL SYSTEM OF THE MEDITERRANEAN FORECASTING SYSTEM PRODUCTS

Marilisa Trani[1], Marina Tonani[1], Alessandro Grandi[1], Jenny Pistoia[2], Paolo Oddo[1],
Giacomo Girardi Feruzza[1] and Massimiliano Drudi[1]

[1]**INGV** (Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna)
[2]**Centro euro Mediterraneo sui Cambiamenti Climatici, Bologna**

274

Index

Introduction

The INGV Mediterranean Forecasting System (INGV-MFS) is currently one of the Monitoring and Forecasting Production Unit of the European MyOcean2 project. The system was implemented in 2000 by the INGV National Group of Operational oceanography (GNOO) and has been developed in years thanks to a number of European projects (see INGV Technical Report n. 163). The centre at INGV operates a near-real time validation system (Cal/Val and MFS-MyOcean System Evaluation) - which provides comparisons between model and observational data of the main ocean variables (SST, 3D temperature and salinity, sea level and currents) on dedicated websites and it releases a Quality Information Document (QuID) on MyOcean website, that illustrates validation diagnostics and quality of its products. Furthermore, it is a partner of the MyOcean work package dedicated to product quality and to improve verification procedures to assess product quality.

This work deals with the verification process of analysis/forecast performances, briefly reviewing methods and measures developed in the context of weather forecasting and illustrating those currently used in ocean forecasting systems. We expose the history and development of verification methods and the common practices and protocols adopted by international organisms that coordinate various national ocean forecasting systems. Then, we give an example of verification of sea level height operated by the INGV-MFS, reporting the on-going improvements and illustrating an application of metrics to validate assimilation of different sets of observations.

1. Quality control systems of forecasting

Assessing the quality of a forecast is essential to improve it and to inform users about the confidence that can be placed in its products. However, planning an efficient and informative evaluation method and defining what could be a ‘good’ forecast is not straightforward.

Murphy (1993) addressed the concept of ‘goodness’ of forecasts identifying three types: 1) consistency - degree of correspondence between forecast and forecasters’ best judgement derived from their knowledge base; 2) quality - degree of correspondence between forecast and some “truth” or reference; and 3) value - incremental benefits deriving from the employment of the forecast by users for decision-making. Furthermore, he described the relationships between the three types of ‘goodness’, highlighting that a high quality not necessary implies a high value of forecasts, and that gaining information about users preferences and being informative on forecast products benefits are crucial to make them valuable.

Quality and value of a forecast are the results of verification and evaluation processes designed to answer to users’ needs. Furthermore, quality assessment is of interest for forecasters to improve the system itself: a process called validation. It consists in comparing the hindcasts of different version of the forecasting system to test developments in one or more components; once the validation has proved improvement of the forecast, the new version can be implemented operationally [Schiller and Brassington, 2011].

The results of these processes provide clues for administrative, scientific (or diagnostic) and economic purposes, as classified by Brier and Allen (1951): determining the state of the art and trends in quality of operational forecast performances over time serves to justify and gain financial aid for advancing training and equipment; assessing the strength and weakness of the forecast allows to better understand underlying physical processes for improving forecasting models; tailoring the types of products released and verification schemes on users’ needs helps to effectively communicate forecast outcomes for enhancing its value.

1.1. Historical development of verification methods

Techniques of forecast verification have been developed since the end of the nineteenth century, in the context of weather forecasting, thanks to a *curious* event known as the “Finley affair” [Murphy, 1996]. Although weather forecasting services were operative since the 1850s in the USA and Western Europe, the attention to the question of forecast verification arose only after 1884, triggered by a paper about an experimental tornado forecasting program published in the *American Meteorological Journal* by John Finley (a Sergeant of the U.S. Army). In the paper, he assessed the quality of his forecast by a yes/no table, validating his outcome as the percentage of correct tornado/no tornado forecasts and coming with a 96.6% of accuracy [Finley, 1884].

One of the early responses to Finley's work noted that the use of yes/no table would have led to an higher accuracy if he had forecasted only no tornadoes (98.2%) [Gilbert, 1884], thus implying that such criterion did not led to meaningful information about the actual skill of his forecast.

The following decade, 1884-1893, saw a fervent and proactive discussion about effective measures and methods to verify forecasts that involved researchers from many different fields. Some of the main focuses of debate were on the appropriate measures to use to account for right forecasts obtained by chance, errors in observations and what to do with forecasts of non-events. Most of the measures proposed at that time are currently in use or were reassessed in years, and some concepts and methods issued are still object of discussion [Murphy, 1996].

The baselines of verification methods were theorized and developed in a landmark paper by Murphy and Winkler (1987) and in a series of subsequent papers by Murphy (1988, 1991, 1993, 1996).

The verification process can be addressed using two main approaches: the measures-oriented approach and the distributions-oriented approach [Brooks and Doswell, 1996; Murphy, 1997]. The former (the traditional approach) consists in judging forecasting performances by looking at numerical scores that quantify some verification attributes; while the latter (more recent) considers the statistical characteristics of the joint probability distributions of forecast and observations as a whole.

Verification attributes address different aspects of forecast quality. The most common used (according to the measure-oriented approach) are bias, association, accuracy and skill, that express respectively: the correspondence between forecast and observation means, the strength of the linear relationship between pairs of forecast and observation, the degree to which a forecast corresponds to observations and the degree to which a forecast corresponds to a standard reference (i.e. climatology, persistence or old model forecast). Other attributes are, for example, reliability, resolution and sharpness (related to the distribution-oriented approach), that concern with the properties of the joint probability distribution: the first two are respectively related to the conditional bias and the standard deviation or variance of the conditional distribution of observations; while the latter considers only the unconditional distribution of the forecast and refers to its ability in predicting extreme events.

Attributes are quantified through different scores, depending on the type of forecast and predictands (deterministic, probabilistic, dichotomous, continuous, categorical) and the purposes of verification. For example, dichotomous forecasts, such as the Finley's tornado forecast, may be verified by using contingency tables that reports the conditional and marginal distributions of forecast and observations; while deterministic forecasts that involve continuous variables (though operationally discrete) may be verified by measuring accuracy, skill and association over forecast-observations pair through statistics, as *mean squared error* (MSE), *root mean squared error* (RMSE), *mean average error* (MAE) or correlation.

Graphics are also used in addition to attribute measures, for example using scatter plots of forecast values against observed values (Figure 1, a); box plots, that shows the median and data between the 25th and 75th percentiles over the full range of data (Figure 1, b); and Taylor diagram (Figure 1, c), based on the correlation coefficient and the standard deviation, that combines measures and graphic allowing to compare simultaneously the skills of various forecasts.

In verifying a forecast, uncertainty in observational data have also to be taken into account, since they are affected by instrument and sampling biases, so, in addition to improving observational systems, confidence intervals should be also associated to attribute scores.

A verification system may be schematized into three main steps: identify a group of attributes that fulfil the desired goals; select measures and graphics that best represent the attributes; choose the appropriate reference of comparison to set a base level of skill. In addition, references and forecasts have to be matched together, for example through interpolation.

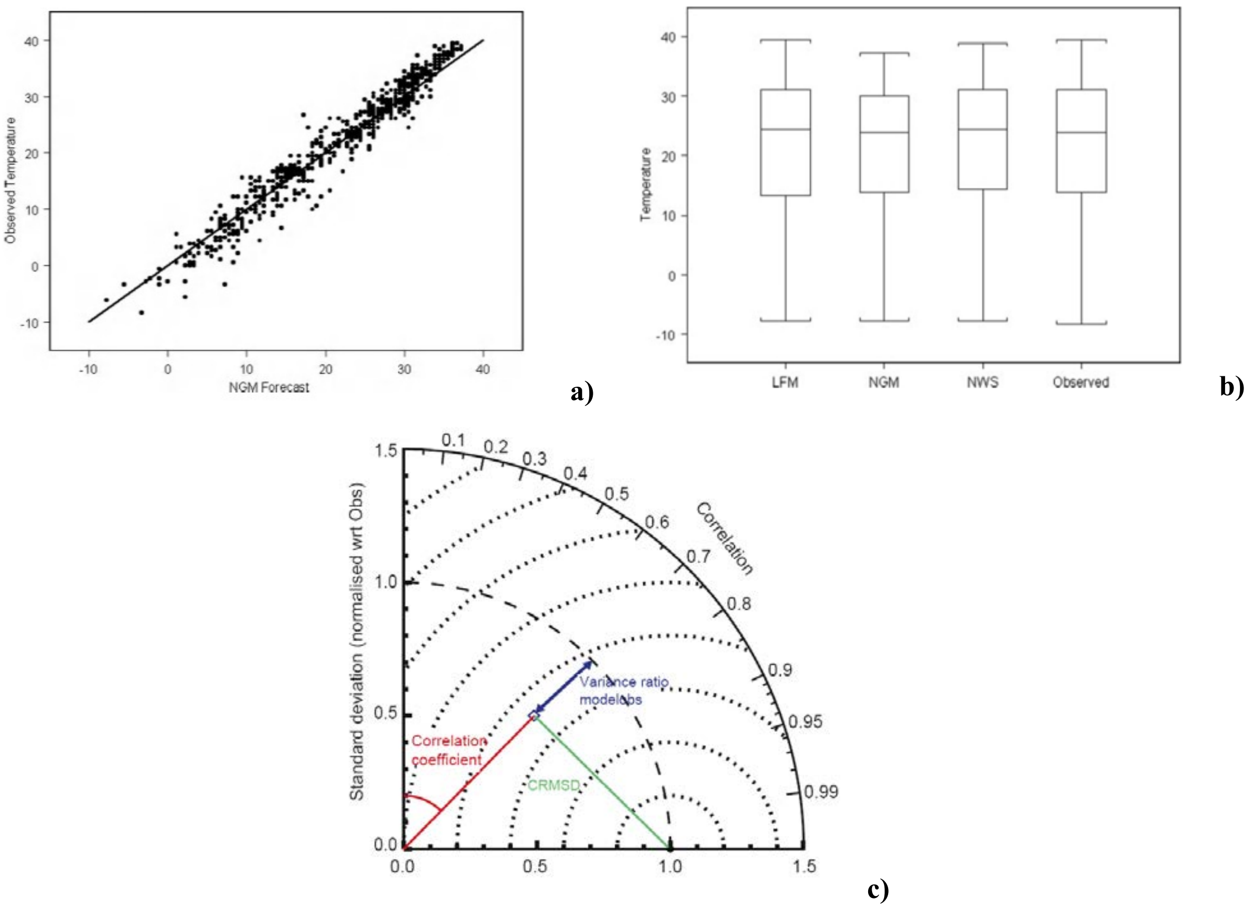


Figure 1. a) scatter plot, forecast (x-axis) against observed values (y-axis); b) box plot comparing three forecasts and observations (the boxes show the range of data between the 25th and 75th percentiles, horizontal lines inside the box the median value and the whiskers the complete range of the data); c) Taylor diagram. (The first two graphics are from Joliffe and Stephenson, 2003, and the third is from Schiller and Brassington, 2011).

1.2. Verification of ocean forecasting systems

Operational oceanography systems and related international programs involve several efforts and dedicated working groups to enhance, implement and standardize procedures to monitor forecast performances. Their aim is either to facilitate the comparison among the various systems and to provide users with the most informative - and simple to understand - measures of the products' quality and value.

Ocean forecast and analysis products can be used for a number of applications - like marine environment monitoring, scientific research, weather forecasting and climate predictions, marine safety, ship routing, oil spill and pollution forecasting and fisheries management – each requiring a different level of quality: for example, what could be a satisfactory accuracy for open ocean circulation studies may not be sufficient for marine safety activities.

The state variables of interest for most of the users are: sea surface height, three-dimensional temperature, salinity and currents, and sea-ice concentration, velocity and thickness. Other products that can be needed for particular applications are some diagnostic quantities, like mixed-layer depth and transport. Since forecasting models produce a large amount of information - that are only selective needed by users - some post-processing is performed to synthesise information, for example providing daily means.

Many aspects related to the type of model and to the data assimilation scheme affect forecasts' quality: horizontal and vertical resolution of the model, surface forcing and lateral boundaries and the type, number and quality control of assimilated observations can have an impact on forecasting performances. However, addressing each source of error is not a simple task, therefore specific designed procedures and measures are needed to focus on one or more sources: for example, to test the validity of the assimilation scheme.

Observations used for assimilation and/or verification are satellite SST and SLA, which undergoes various levels of processing, salinity and temperature profiles from ARGO floats, CTD and XBT, and current velocity from drifting buoys. *In situ* measurements can also be provided by moored buoys. Observational data are quality controlled, so they are provided to the operational system with an associated error, however an error of representativity has also to be taken into account, as observations are point measures whereas model values are area-averages spanning the grid size of the model.

The selection of a number of ocean forecast scores started with the definition of standardized diagnostics and a set of metrics - mathematical tools that computes scalar measures from system outputs and compare them to "reference" [Crosnier et al., 2006] - in the context of the European MERSEA project and the international GODAE program (see Chapter 1.2.1), that leaded to a series of inter-comparison exercises. In 2003-2004, the MERSEA Strand 1 inter-comparison project provided the first attempt to compare different ocean forecasting systems, defining principles of validation and four classes of metrics, together with common quantities, grids and file format to share results [Crosnier et al., 2006, 2007]. Three principles was identified: consistency –correspondence of forecast outcomes with the current knowledge of the ocean circulation and climatologies; quality - degree of the difference between forecast best-estimate and an independent set of observations; and performance – short term forecast capacity of each system, i.e. against persistence or climatology. A fourth principle considered is benefit – assess quality standards to reach before a product is useful for an application (survey on end-users' needs).

The inter-comparison project required each system to compute a set of metrics for the same geographical area and following the same procedure, storing results in a netCDF file on a local server, to make them readily available and usable by other systems.

The four classes of metrics used during MERSEA and GODAE programs are:

- Class 1 metrics, that provide a general overview of ocean dynamics by comparing some forecast daily mean fields with climatologies or hindcasts (consistency assessment) on predefined 2D and 3D grids. These fields can be used to study time series or spatial variability, derive some diagnostic quantities as eddy kinetic energy, or performing spectral analysis.
- Class 2 metrics, that compare forecast outcomes at selected locations (e.g. on virtual moorings or along section tracks of the model domain) that match areas where measurements are most available (e.g. sea level from tide gauges, temperature, salinity and currents from moorings, WOCE and CLIVAR repeated sections and the roots covered by ships of opportunity). Such metrics are complementary to Class 1 for they allow to assess outcomes on higher temporal and spatial resolution, but storing a reduced amount of data form respect to Class 1.
- Class 3 metrics, that are based on computing integrated physical quantities form model variables during the model run, on the native grid at each time step, as volume and heat transport through selected sections or overturning streamfunctions. These metrics allow to check the model from the physical point of view.
- Class 4 metrics, that aim to measure data assimilation performance and forecast skill assessing all fields (hindcast, nowcast, analysis, climatology and persistence) for the same day with the same criteria. Statistics, like RMSE, are computed on differences between observational data and model values on the "observational space".

The European project MERSEA conducted a real time inter-comparison of five operational forecast systems for the North Atlantic and Mediterranean Basins and was followed by the MERSEA Integrated Project (2004-2008).

In 2008, a GODAE inter-comparison project involved seven ocean forecasting systems, of which four global and three regional, adopting the MERSEA framework and defining more metrics adapted to the Global Ocean and Arctic and Antarctic seas, for example adding specific metrics for the ice [Hernandez, 2009].

These projects proved that large scale features of ocean circulation was well represented by the various systems despite the different model configurations, resolutions and assimilation techniques; discrepancies was observed in regional areas. Furthermore GODAE promoted observational projects like Argo and the Global High-Resolution Sea Surface Temperature (GHRSST), which provided data in real time, and created centres for data storage and distribution as US GODAE and Coriolis. The methodologies designed and tested also constitutes a basic framework for developing dedicated metrics for coastal areas and biogeochemical models.

1.2.1. International GODAE Inter-comparison and Validation Task Team

The Global Ocean Data Assimilation Experiment started in 1997 aiming to create a net of operational systems capable of estimate and predict the state of the global ocean for social and economical applications. On the wake of the First Global Atmospheric Research Program Global Experiment (FGGE) [Bengtsson, 1981] - which proved the large benefits of an efficient observation system on the feasibility and quality of global weather predictions – GODAE allowed the development of global ocean predictions taking advantages from the increasing availability of near-real time observational data provided by satellite measurements and from the rapid improvement in numerical modelling due to advances in computing power [Smith and Lefebvre, 1997]. Furthermore, the program made accessible observations, forecasts and analysis products to users.

It ended in 2008 as a demonstration program, establishing a new group, the GODAE OceanView Science Team (GOVST), that coordinates and promotes international collaboration and research for improving operational oceanography systems. This group is complemented by the Expert Team on Operational Oceanographic Forecasting Systems (ET-OOFS), established by the Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) for (1) providing guides and standards (i.e. type of products, validation procedures, metrics, nomenclatures, symbology, etc.) to the operational oceanography systems; and (2) evaluating the suitability for operational adoption of new scientific and technological findings and leading inter-comparison activities, to deliver consistent and valuable ocean forecasting products to the users.

The GODAE OceanView Science Team (GOVST) consists of five task teams dedicated to enhance scientific knowledge and technical capabilities on a particular topic, i.e. the observing system evaluation, coastal areas, marine ecosystem and short to medium term predictions and product evaluation. In particular, product evaluation is addressed by the Inter-comparison and Validation Task Team (IV-TT), that provide an inter-comparison and validation framework shared by the operational oceanography systems, coordinating and promoting the development of methodologies and metrics.

1.2.2. European MyOcean2 quality control system

MyOcean2 is a pre-operational project granted by the European Commission within the GMES Program (Global Monitoring for Environment and Security), scheduled to run from April 2012 to September 2014. Continuing the first MyOcean project (2009-2012), which had extended MERSEA program and unified some previous projects (Polar View, ECOOP, Glob Colour, MarCoast), it is testing and developing a European service for ocean monitoring and forecasting. The service, which is planned to become operative as a component of the GMES Marine Service, integrates national centres that collect data and produce forecasts, avoiding overlapping of resources and standardizing operational processes and quality. As a result, users of all marine applications can access free to high quality observational, model-based data and up-to-date information on the global ocean and European seas. The products provided are physical (temperature, salinity, currents, sea ice, sea level and wind) and biogeochemical parameters (chlorophyll, oxygen, nitrate, phosphate and some optics properties).

All data and information about the products - including file format, downloading and interpretation, observational data processing, model's structure, product quality and a detailed description of the main application fields - are available on MyOcean website (www.myocean.eu).

MyOcean2 is structured (Figure 2) in nineteen work packages and eleven production units, of which four Thematic Assembly Centres (TAC) and seven Monitoring and Forecasting Centres (MFC). Thematic Assembly Centres deal with the collection, treating and delivering of observational data (sea level, ocean colour, sea ice-wind-SST, and *in situ* data); while Monitoring and Forecasting Centres operate 3D models of different regions in the past (25 years), in real-time and in the future (1-2 weeks) (Global Ocean, the Arctic area, the Baltic Sea, the Atlantic North-West shelves area, the Atlantic Iberian-Biscay-Ireland area, the Mediterranean Sea and the Black Sea).

Each production centre constitutes a work package, while other work packages are dedicated to specific tasks, one of which is responsible of the product quality (WP17).

The Product Quality Work Package, led by the UK Met-Office, provides guidelines and conventions to production centres for performing routine validation, sharing information on quality control between MFCs and TACS and enhancing validation capabilities. Guidelines are released once per year and are drawn on the basis of previous results and change proposals from the WP17 partners.

Each production centre releases a Quality Information Document (QuID) on MyOcean website, providing information about the quality of their products through EANs (Estimated Accuracy Numbers, mean and root mean square of the difference between the operational product and a reference, computed for particular depths and sub-regions if present).

Guidelines set standards for the production and presentation of metrics (EAN) that summarize products' accuracy, covering aspects as type of metrics, units, precision, regions, vertical layers and format of presentation.

A systematic procedure of validation is at an implementation phase, whose goal is to provide web-based reports of product accuracy to users on a quarterly basis. The report will include tabulated summary statistics and an explanatory text and will be drawn by the UK Met-Office centre that collects statistics from all the production units.

The minimal set of statistics required are: the number of data values used to compute statistics, the mean of model-based products and of reference and the mean squared error. Optional statistics are the variances of products and reference and the correlation or anomaly correlation, that can allow the presentation on a Taylor diagram. Production centres send their statistics in NetCDF file format accompanied by a description of the verification process, specifying: how statistics are calculated, if some temporal processing has been performed on products or reference and if reference data are 'not independent' (already assimilated) or 'independent' (not assimilated).

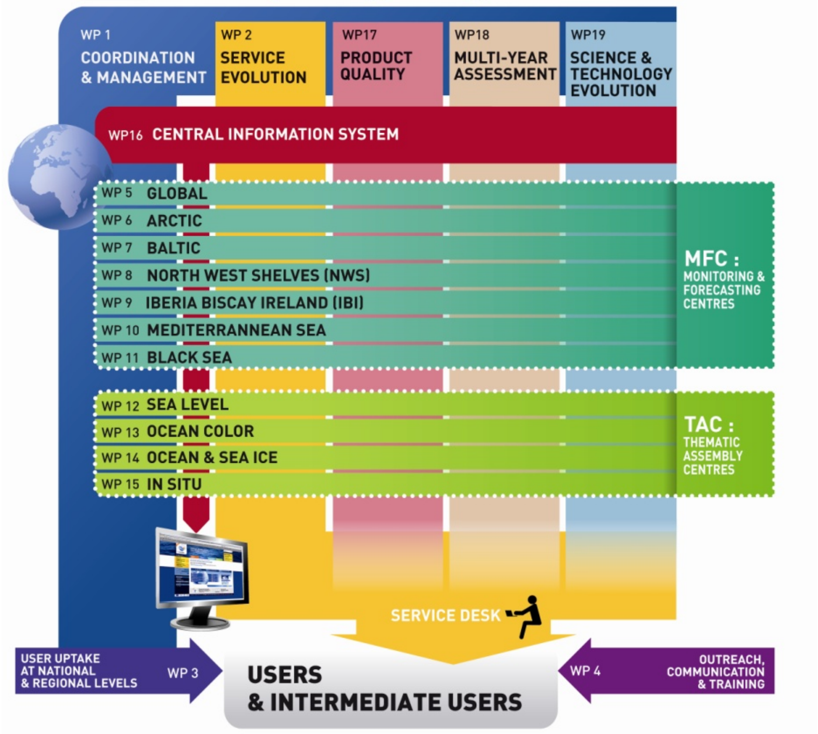


Figure 2. MyOcean 2 work packages organization (from MyOcean website, www.myocean.eu).

1.2.3.INGV MFS quality control system

Information about the quality of the products is currently provided by INGV MFS through a document on MyOcean website (Quality Information Document, www.myocean.eu) and two electronic reports (gnoo.bo.ingv.it/mfs/myocean/calval/ and gnoo.bo.ingv.it/mfs/myocean/evaluation.html).

The Quality Information Document (QuID) assesses differences between model values (forecast and analysis) and observations through Estimated Accuracy Numbers (EAN) and describes the validation system.

Med-MFC provides tabulated EANs for temperature (T) and salinity (S) at selected depth intervals and for sea surface temperature (SST) computed from respect to the first and third day of forecast and the analysis. Observations used as reference are vertical profiles of T and S from ARGO floats and moored buoys and satellite SST. EANs presented in the current QuID (MYO2-MED-QUID-006-001-V3.1, March 2013) are computed using analysis and forecasts produced by the Med-MFC V3 system during October-November-December 2012.

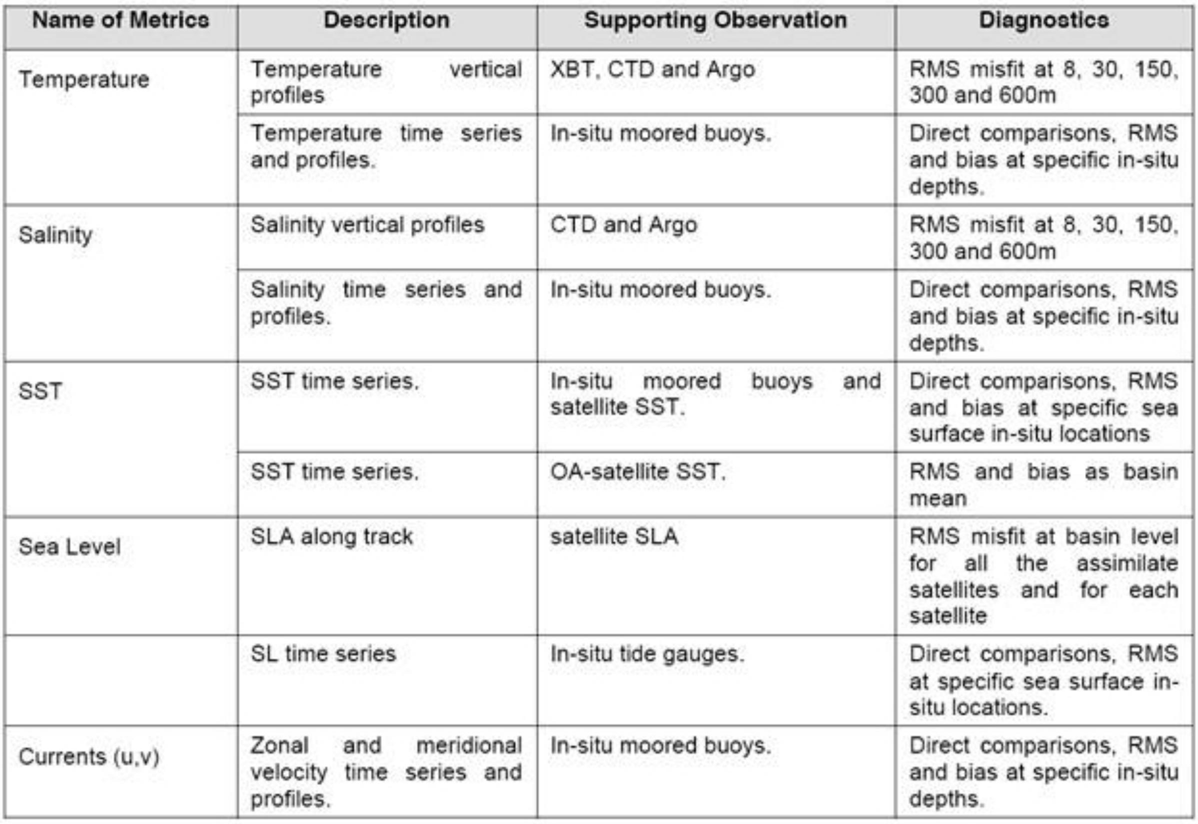
The validation system, focused on model-observation pairs comparison, provides RMS misfits and bias that are available online in near real time. Temperature, salinity, currents and sea level are compared to satellite SST, vertical profiles of T and S from ARGO floats, moored-buoys, CTD and XBT, current velocity from moored buoys and sea level measures from tide gauges (Table 1). All assimilated observations are considered semi-independent (Tonani et al. 2009) due to the kind of assimilation system implemented (see Chapter 2.2 for further details).

On Cal/Val website (gnoo.bo.ingv.it/mfs/myocean/calval/) model values and *in situ* observations from moored-buoys can be directly compared as time series or vertical profiles by selecting the desired product, a particular buoy, time and depth.

On the evaluation website (gnoo.bo.ingv.it/mfs/myocean/evaluation.html), statistics are available as temporal series of RMS of the first seven days of analysis' and observation' misfits, updated each Wednesday, of sea level anomaly (SLA), SST and temperature and salinity at various depths. SLA and SST are compared to satellite data. The RMS of satellite-model misfits of SLA are provided from respect to data of all satellites and of each satellite separately. For SST, bias is also shown as well as a map comparison of satellite surface temperature fields (measured during night time) and the forecast daily mean fields of the first four days of forecast produced on Friday.

Temperature and salinity are compared with XBT (only temperature), CTD and ARGO data and skill scores are also shown, comparing the relative accuracy of the forecast from respect to the persistence and the analysis and of the analysis from respect to the persistence.

Table 1. Details on metrics and diagnostics provided by INGV MFS (table from QuID on MyOcean website, www.myocean.eu).



Name of Metrics	Description	Supporting Observation	Diagnostics
Temperature	Temperature vertical profiles	XBT, CTD and Argo	RMS misfit at 8, 30, 150, 300 and 600m
	Temperature time series and profiles.	In-situ moored buoys.	Direct comparisons, RMS and bias at specific in-situ depths.
Salinity	Salinity vertical profiles	CTD and Argo	RMS misfit at 8, 30, 150, 300 and 600m
	Salinity time series and profiles.	In-situ moored buoys.	Direct comparisons, RMS and bias at specific in-situ depths.
SST	SST time series.	In-situ moored buoys and satellite SST.	Direct comparisons, RMS and bias at specific sea surface in-situ locations
	SST time series.	OA-satellite SST.	RMS and bias as basin mean
Sea Level	SLA along track	satellite SLA	RMS misfit at basin level for all the assimilate satellites and for each satellite
	SL time series	In-situ tide gauges.	Direct comparisons, RMS at specific sea surface in-situ locations.
Currents (u,v)	Zonal and meridional velocity time series and profiles.	In-situ moored buoys.	Direct comparisons, RMS and bias at specific in-situ depths.

2. Metrics implementation for sea level quality assessment

Sea level is an important property of the ocean used for a number of applications. It is the integral of the ocean interior and its variations give information about many physical processes which contribute to shaping the ocean surface at various spatial and temporal scales. It is directly related to ocean currents through geostrophic balance, allowing for the reconstruction of ocean circulation over large scales and for studying mesoscale variability; on seasonal, inter-annual or longer temporal scales, information about sea level changes are useful, for example, in coastal erosion surveys and climate studies. For these reasons, it is also one of the core variables produced by ocean models.

Sea level is measured locally by tide gauges and on a global basis by satellite altimeters. Tide gauges provide long temporal series at the same point, but are placed at a few locations near coasts, while satellites provide near real time observations of sea surface height (SSH) at high space and time resolution. Indeed, the development of operational oceanography is strongly linked with that of satellite altimetry, since sea level from altimetry is assimilated by ocean models to constrain circulation and used as a reference for validation of products.

2.1. Principle of satellite altimetry

An altimeter mission combines radar technique and a positioning technique, computing SSH as the difference between the satellite orbit altitude and the distance between the satellite and ocean surface, and recording along-track measurements on repetitive ground tracks.

The altimeter measures the round-trip travel time of a microwave pulse emitted downward by the radar and reflected back by the ocean surface, while a tracking system determines the three-dimensional position of the satellite relative to a given earth ellipsoid. The accuracy in the computation is affected by various factors: measurement errors due to the interaction of light with the ionosphere and water vapour, which depends on the frequency used and can be corrected using dedicated instrumentation onboard (bi-frequency altimeter radar and a microwave radiometer); and errors due to the positioning system.

SSH derived by altimetry comprises the geoid (the equipotential of the earth gravity field to which a motionless ocean would exactly conform) and the ocean surface topography (the departures of sea surface from the resting level due to atmospheric pressure distribution, tides and currents).

Tides and atmospheric pressure contributions to the total altimeter measure have to be removed, leaving only the contribution to SSH due to ocean currents, i.e. dynamic topography, which is of most relevance for oceanographic studies.

Tidal contribution is estimated by analysing the tidal signal on a long time series of altimeter records, while atmospheric one is approximated by the inverse barometric effect (1 mbar increased pressure lowers sea level by 1 cm). The most difficult contribution to be removed is that due to the Earth's gravity, since it is 100 times larger than that due to the other factors, hence geoid should be known with the precision of few centimetres to obtain a reliable ocean dynamic topography.

At the present time, since geoid is not yet well known, only the variable part, i.e. the Sea Level Anomaly (SLA), can be obtained with confidence by the altimeter measure. It is extracted from the total SSH by removing a mean profile computed over many orbit cycles, when a long enough time series of records on the same orbit is available.

The orbit cycle of a satellite can vary from 10 to 35 days, thus the orbit track must be repeated for several years to compute a reliable time-mean topography. For this reason, when planning a new altimeter mission that have to replace an old one, is essential to fly the new altimeter in the same orbit and possibly with an overlapping time of six months. If a new mission is not in a repetitive orbit phase, an alternative Mean Sea Surface (MSS) is used, that is obtained by combining the along track data from different altimeter missions referencing them to a particular mission (see AVISO website, www.aviso.oceanobs.com, and MyOcean documentation, www.myocean.eu).

Table 2 illustrates the main characteristics of the altimeter missions in operation during 2013 or part of the year (AVISO and MyOcean SL TAC documents). Jason 1 and Jason 2 continue the Topex/Poseidon mission ended in 2005; they have the shortest repeat cycle but with a wide spacing between tracks. Saral is on the same ground track as the previous ERS and Envisat missions and is complementary to Jason 2; it has a longer repeat cycle and a narrower spacing between tracks. Cryosat-2 is polar orbiting and designed to measure ice cover; it has a repeat cycle similar to the Saral's. Panels a-d of Figure 3 show the coverage of the Mediterranean Sea during a cycle of each satellite.

Time-mean profile used for Jason 2 is computed with 10 years of Topex/Poseidon and 6 years of Jason1 data; for Jason 1 and Cryosat-2 a MSS computed using multi-mission data with the 10 years T/P as reference is used, because there is no repetitive track; for AltiKa (Saral's altimeter) is computed with cycles ERS2 and cycles Envisat data.

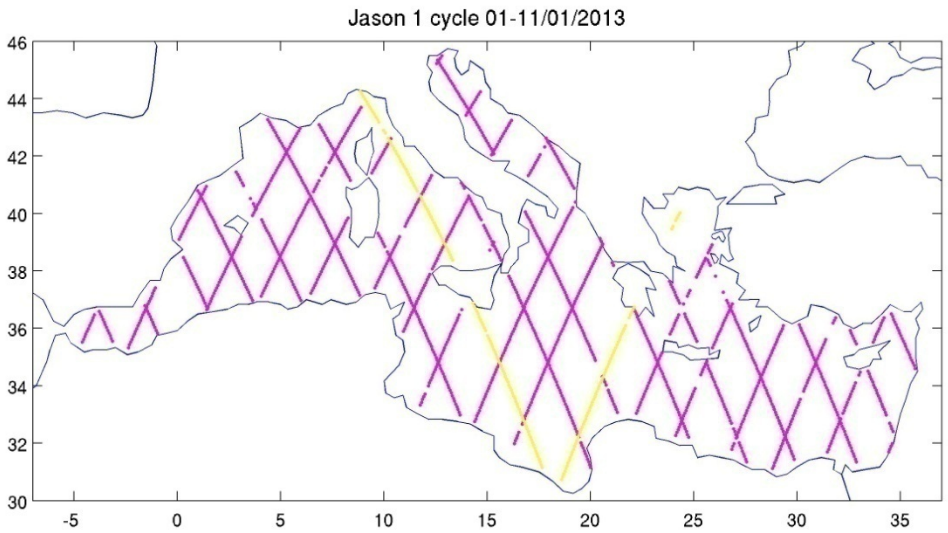
Altimeter products for operational oceanography or other applications are provided worldwide by the SSALTO/DUACS system. The centre collects and processes data in order to make them directly usable.

Main processing steps are: error corrections, cross-calibration and merging of measurements from different altimeter missions and production of high quality along track and maps of sea level anomalies.

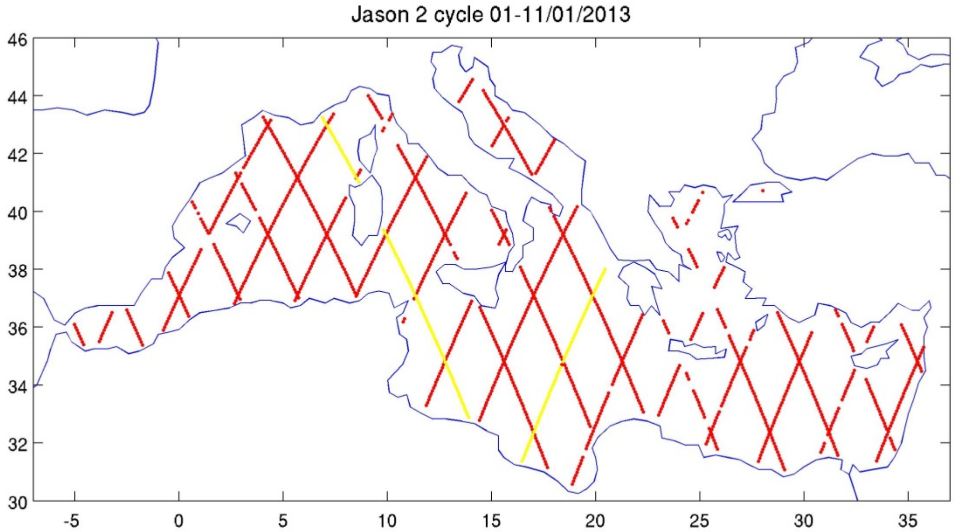
Products are freely available both in near real time and delayed mode from MyOcean catalogue.

Table 2. SLA satellites's characteristics.

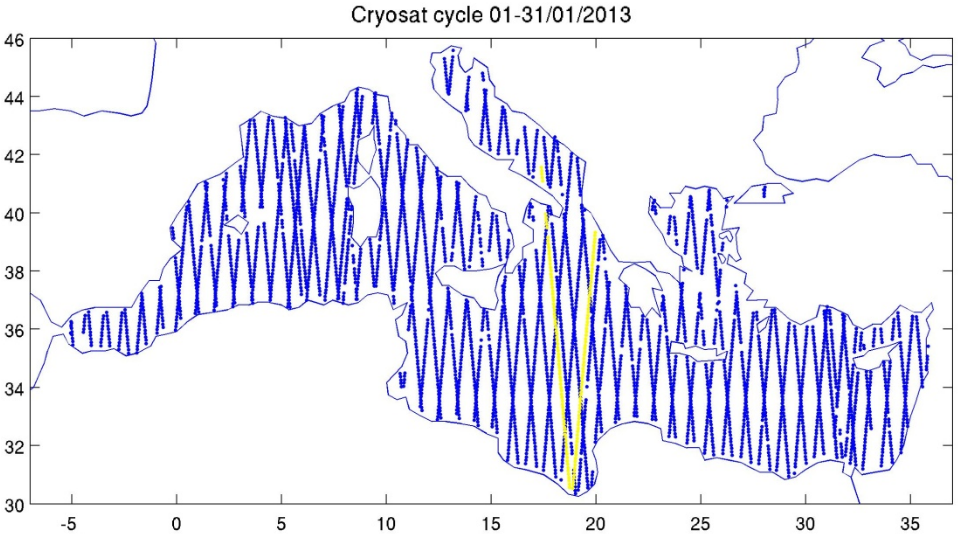
Satellite	**Jason 1**	**Jason 2**	**Cryosat**	**Saral**
Agency	Cnes/Nasa	Cnes/Nasa/	Esa	Isro/Cnes
Launch on	07/12/2001	20/06/2008	08/04/2010	25/02/2013
End Date	01/07/2013	On going	On going	On going
Satellite characteristics				
Altitude	1336 km	1336 km	717 km	~800 km
Inclination	66 °, non sun-syncronous	66 °, non sun-synchronous	92°, non sun-synchronous	98.55°, sun-synchronous
Repetitivity	~10 days	~10 days	30 days	35 days
Ground track spacing	315 km at the equator	315 km at the equator	60 km at the equator	75 km at the equator
Instrument	Poseidon-2, Ku-band	Poseidon-3, Ku-band 13,6 GHz (altimetry) and 5.3 GHz (ionospheric correction);	Siral, Ku-band 13.575 GHz	AltiKa, Ka-band 34GHz
Product characteristics				
Spatial along track resolution	14 km	14 km	14 km	14 km
Estimated Accuracy Numbers (RMS cm)	< 5	< 4	< 6	< 5


Jason 1 cycle 01-11/01/2013
46
44
42
40
38
36
34
32
30
-5
0
5
10
15
20
25
30
35

a)


Jason 2 cycle 01-11/01/2013
46
44
42
40
38
36
34
32
30
-5
0
5
10
15
20
25
30
35

b)


Cryosat cycle 01-31/01/2013
46
44
42
40
38
36
34
32
30
-5
0
5
10
15
20
25
30
35

c)

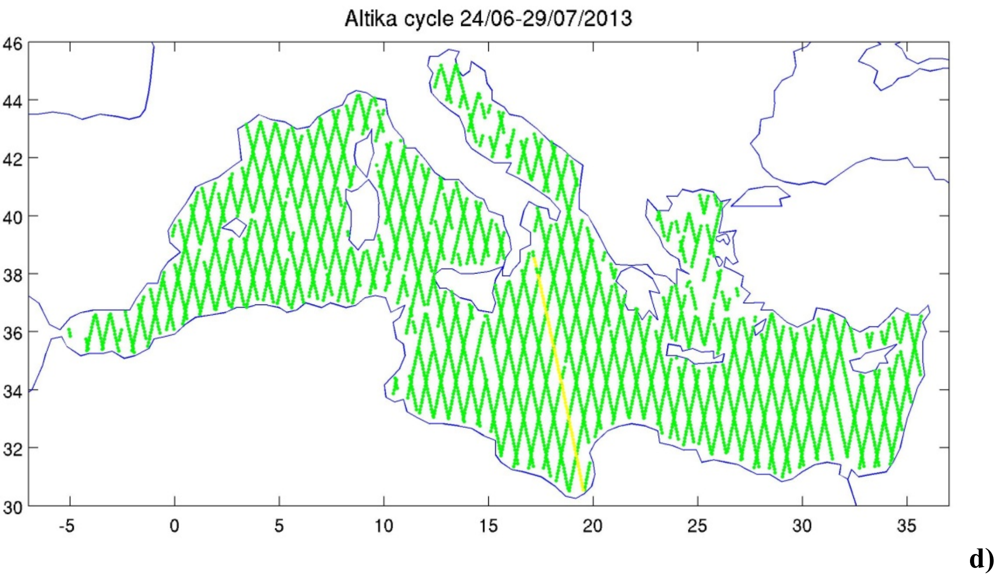


Figure 3. Satellites's cycles: a) Jason1, b) Jason2, c) Cryosat-2 and d) AltoKa. The yellow lines highlight the repeated tracks.

2.2. Mediterranean Monitoring and Forecasting Centre (Med-MFC) operational chain

The model products are provided as daily means centred at midnight (Figure 4, a) on a regular horizontal grid of 1/16° and at 72 unevenly spaced vertical levels. The model implemented (Oddo et al. 2009) is based on the NEMO 3.4 code (Nucleus for European Modelling of the Ocean, version 3.4) [Madec et al. 2008] coupled with a wave model based on the WaveWatch-III code [Tolman, 2009]. The assimilation system is a 3DVAR scheme (OCEANVAR) developed by Dobricic and Pinardi (2008), which makes use of vertical profiles of T and S from ARGO, CTD, and XBT and satellite along track SLA.

The production cycle (Figure 4, b) consists of a 10-day forecast produced every day initialized by a 24h hindcast. The analysis is produced each Tuesday for the previous 14 days: solutions every 24h are corrected by daily assimilation and used to initialise the subsequent 24hr hincast model run (INGV report n.163 and MyOcean MED-MFC Product User Manual).

Validation of model SSH is performed using altimeter along track SLA.

Near real time SLA observations are daily downloaded from MyOcean website, where they are provided in NetCDF format based on CF convention. Each file contains daily passes of a single satellite over the Mediterranean, stored as along track SLA on one time dimension (see Table 3 for full file content).

Figure 5 shows Mediterranean Sea coverage of the satellite tracks available from the 15th of July 2013 to the 29th of July 2013, i.e. during one assimilation cycle. Of these SLA tracks only those of Jason 2 are assimilated. Jason 1 data were also assimilated by the system until the end of June 2013, when the mission had to be terminated due to the failure in transmission of the satellite which could not be restored.

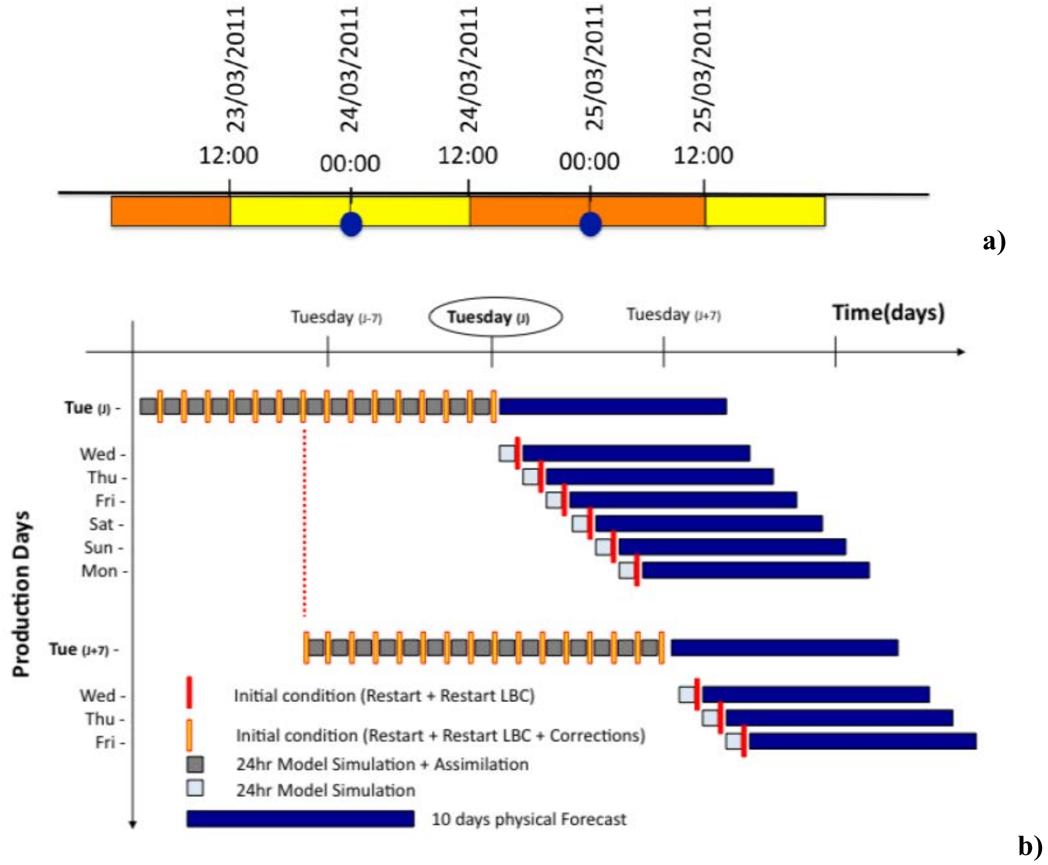


Figure 4. MyOceam Med-MFC product characteristics. The analysis and forecast products are 24hr means centered at midnight, panel a). The production chain scheme is represented in panel b).

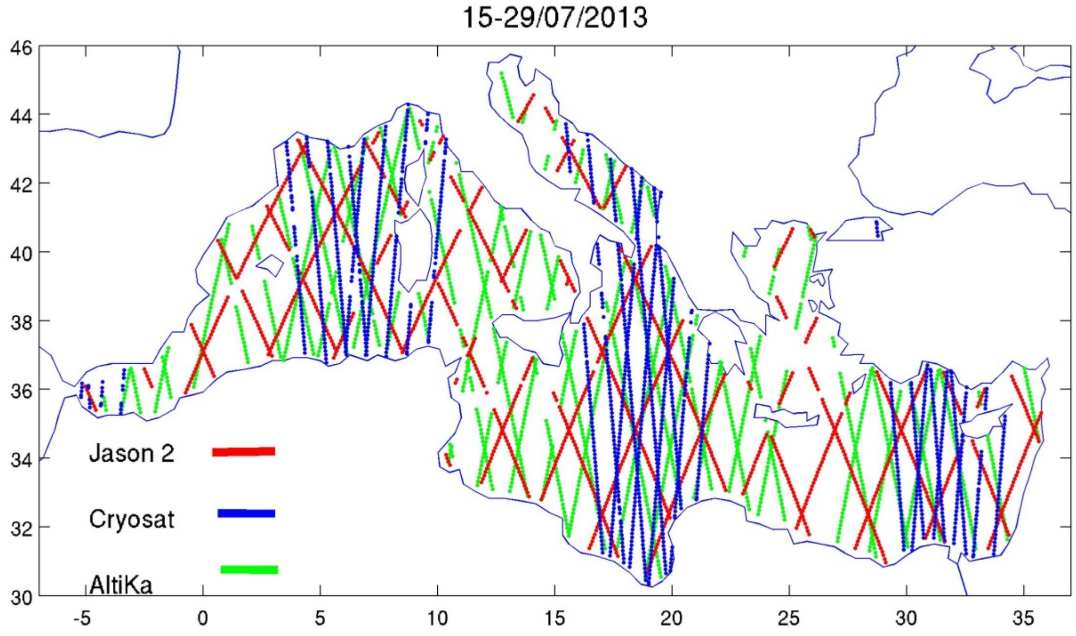


Figure 5. Satellite track coverage during an assimilation periodo f 15 days.

Table 3. Description of the Along Track SLA file content.

	Name	**Description**
Dimension	time	identify the variables depending on time
Variables	time	time in Julian days of each point along a track
	longitude	longitude of each point along a track
	latitude	latitude of each point along a track
	cycle	cycle number for each measurement
	track	track number for each measurement
	flag	'0' if the processed data comes from OGDR; '1' if data comes from IGDR
	SLA	Sea Level Anomaly values for each time given
Global attributes	Comment	name of the satellite mission
	Title	'*along track Sea Level Anomaly*'
	Conventions	'*CF-1.5*'
	Source	*'satellite altimetry'*
	References	URL of the reference document
	Institution	name of the institution that provide the data
	History	date and time of creation of the file, name of creator and number of version

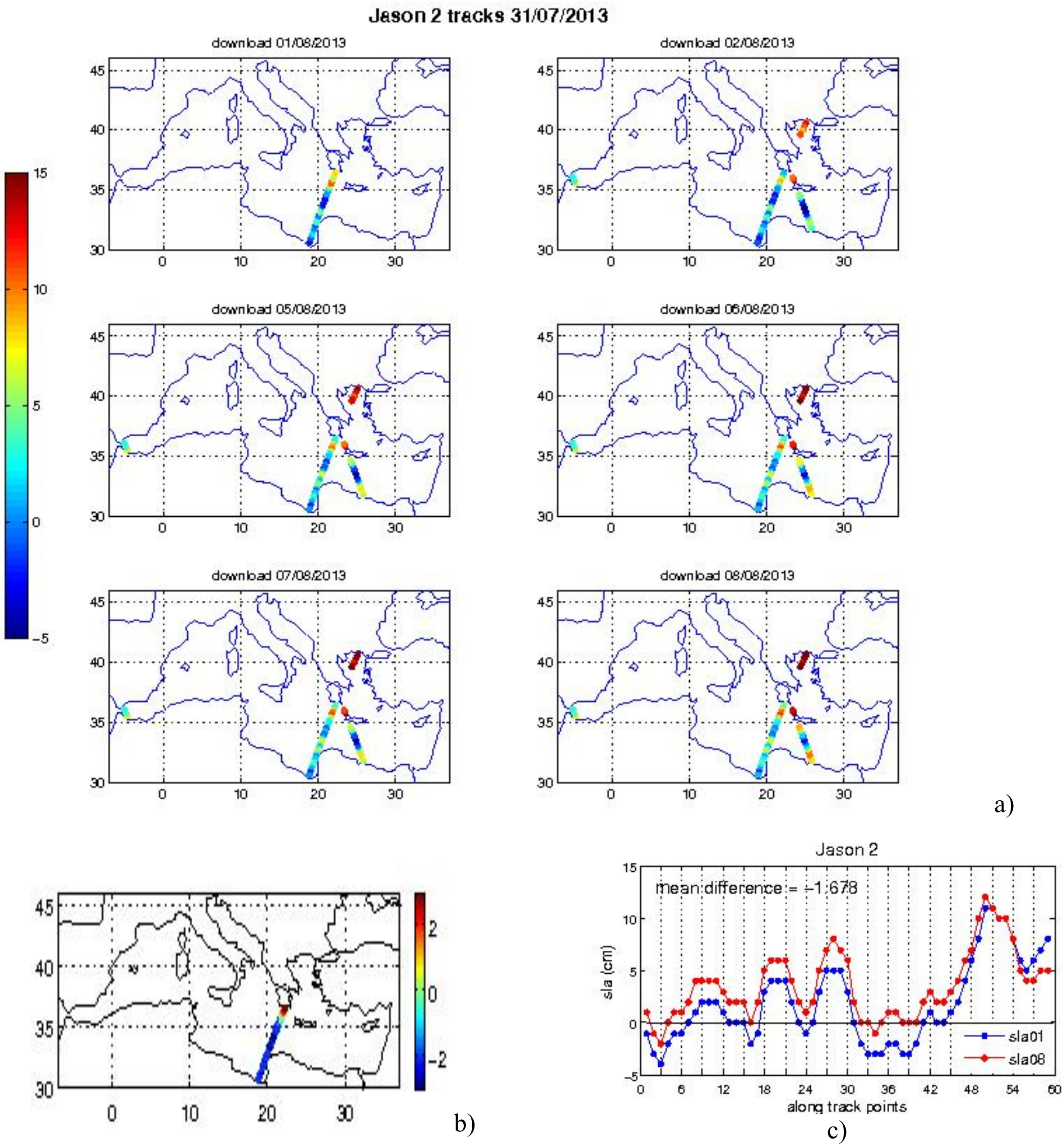
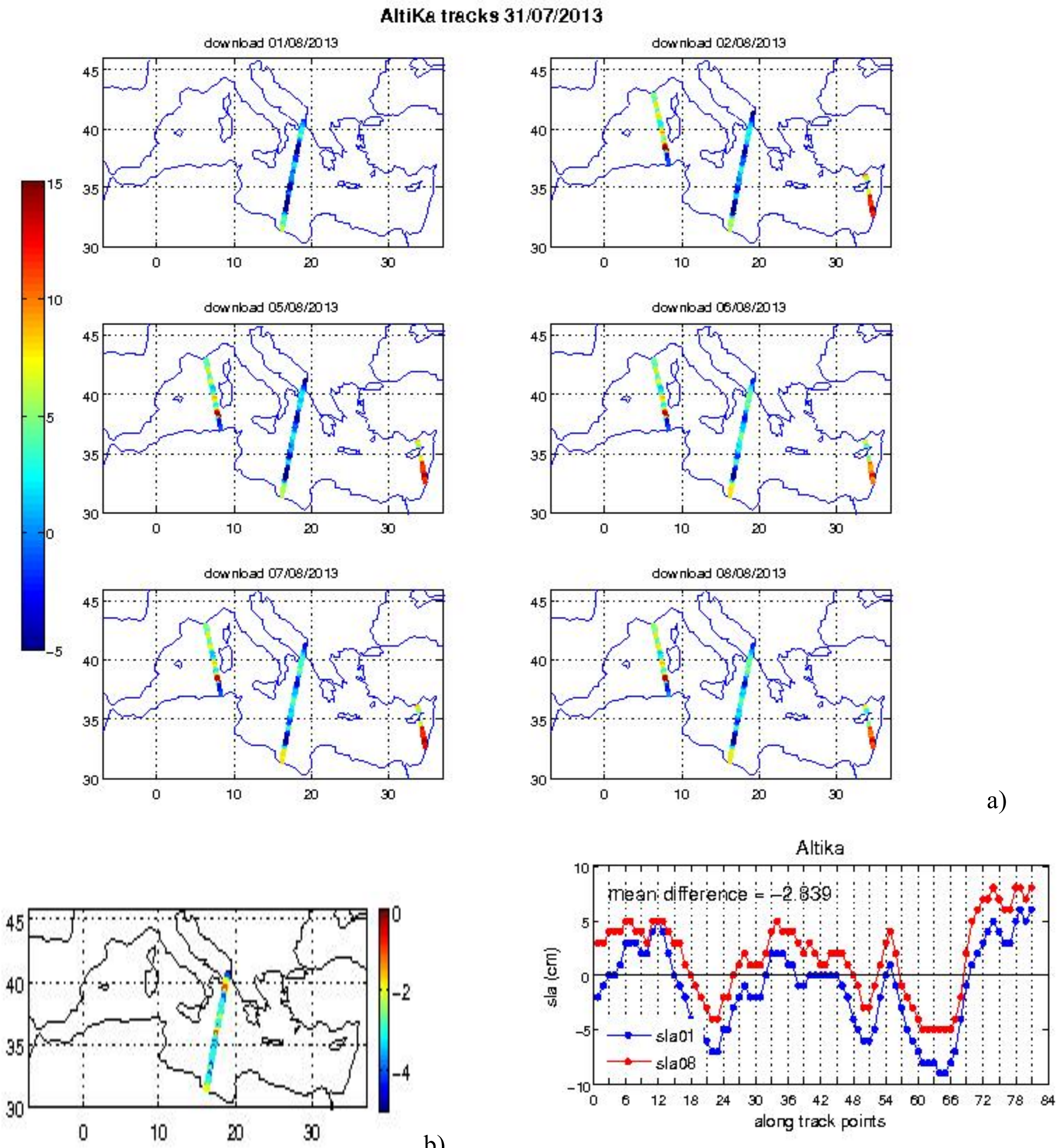


Figure 6. Jason 2 tracks referred to the same day (31/07/2013) downloaded in NRT and DT. The number of observations increases in the DT product and also the quality of the measurement. The different number of observations are shown in panel a). The difference of the observation value at the same location in the different releases of the product are shown in panel b). The comparison is computed only where the observations are available in NRT and DT (5 days later) product.

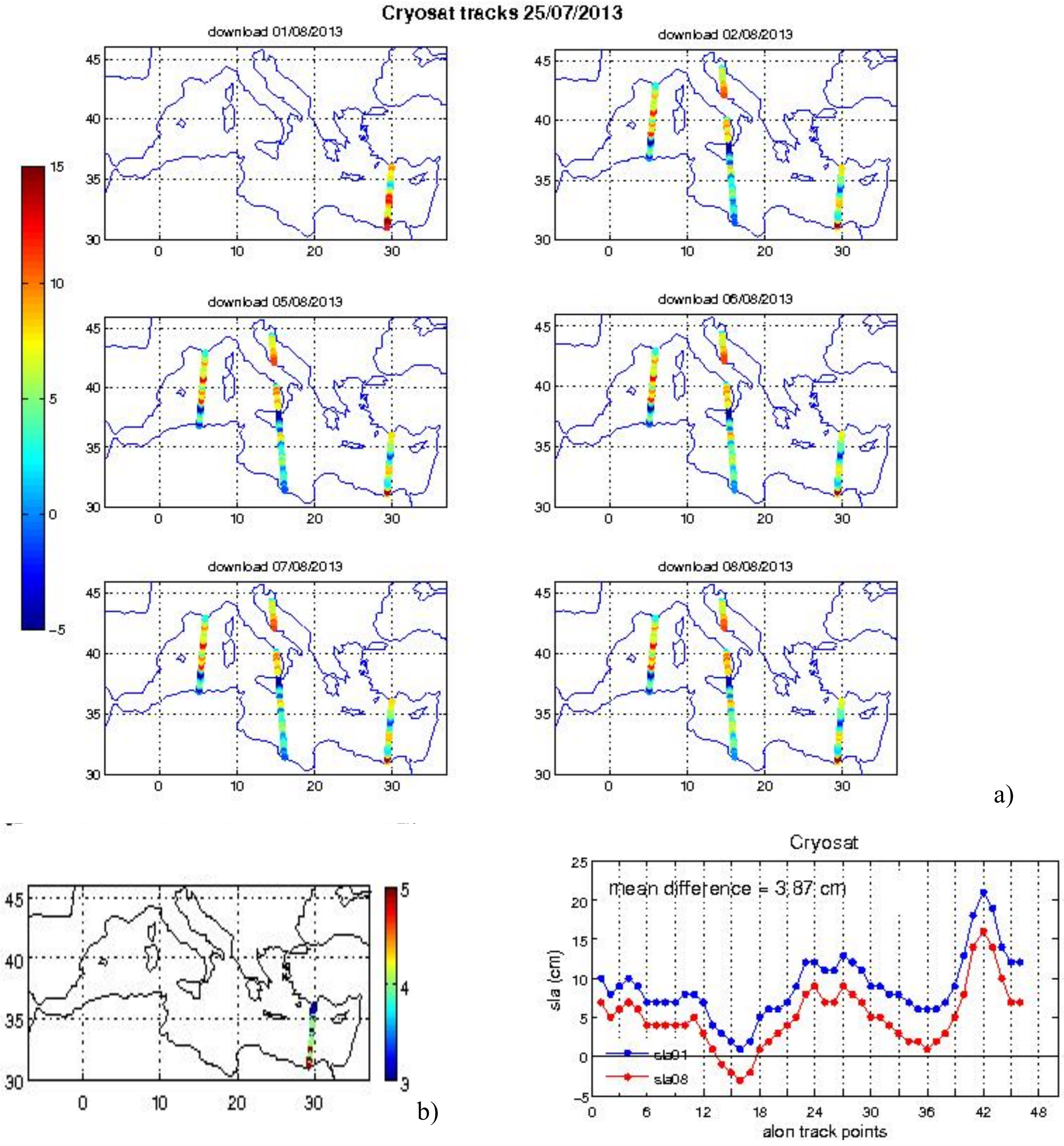


a)

b)

c)

Figure 7. AltiKa tracks referred to the same day (31/07/2013) downloaded in NRT and DT. The number of observations increases in the DT product and also the quality of the measurement. The different number of observations are shown in panel a). The difference of the observation value at the same location in the different releases of the product are shown in panel b). The comparison is computed only where the observations are available in NRT and DT (5 days later) product.



a)

b)

c)

Figure 8. Cryosat tracks referred to the same day (25/07/2013) downloaded in NRT and DT. The number of observations increases in the DT product and also the quality of the measurement. The different number of observations are shown in panel a). The difference of the observation value at the same location in the different releases of the product are shown in panel b). The comparison is computed only where the observations are available in NRT and DT (5 days later) product.

The production centre of altimeter data (MyOcean SC-TAC) reprocesses daily products after the first release on MyOcean catalogue, so files of a given day are updated in days after the first time of release. As a consequence, the same file downloaded in different days can contain added tracks and slightly different values of SLA. In Fig. 2.4 SLA data of the 31st of July 2013 of Jason 2 and AltiKa and of the 25th of July 2013 of Cryosat are displayed, that have been released for the first time the 1st of August 2013 (NRT product). Displays of data downloaded in days after (DT products) that show that tracks have been added and values of SLA changed a bit. For Cryosat, it can be observed how data from the same file downloaded the 1st (upper-left panel of Figure 6, Figure 7 and Figure 8) and the 2nd (upper-right panel of Figure 6, Figure 7 and Figure 8) of August exhibit a notable difference between SLA values along the same track.

2.3. Validation of the Med-MFC SSH product

2.3.1. Method used to match model values and observations

Sea Surface Height (SSH) model values are provided as daily means on a regular grid of 1/16°. Mean Dynamic Topography (MDT), calculated from Dobricic (2005), is subtracted to obtain anomaly values; then, they are interpolated in the along track points to be compared with altimeter SLA.

The model is incompressible and the Bussinesq approximation is adopted, therefore model SSH accounts only for the non-steric component [Mellor and Ezer 1995], while altimeter SLA includes both steric and non-steric contribution. Steric height contribution must be removed from the observations subtracting the mean of misfits (observations minus model values) along the satellite track [Pujol et al., 2010], averaging values over portions of track where the distances among points are lower than 100 km.

A Fortran code was written to perform the above calculations, interpolating linearly model SLA at the observations locations, removing points where data are not available and taking into account that the model daily means are centred at midnight. Therefore, model values for a given day are compared with tracks at times before 12:00 of the same day and after 12:00 of the day before.

The Fortran code also includes an estimate of the steric height contribution to SSH computed by using daily means of temperature and salinity produced by the model.

Steric height is computed as:

$$\zeta_{st} = -\frac{1}{\rho_0} \int_{-H}^{0} \rho dz$$

where ρ_0 is the reference density at T=0°C and S=35, ρ is the density from *in situ* temperature, salinity and pressure in each cell of the model. The model provide potential temperature and salinity, so the code compute pressure from the depth of the cell, i.e. a vertical level of the model, and *in situ* temperature from model potential temperature. Then, at each grid point, values are integrated vertically and mediated over the entire Mediterranean surface to obtain a daily basin mean.

Figure 9 shows, as an example, the model SSH analysis field of the 1st of January 2013 from which model SLA was computed, while in Figure 10 a,b (upper panels) are displayed for comparison SLA values of the first and third day of forecast and analysis (respectively green, light blue and blue lines) and along track SLA from Jason1 and Jason2 altimeters (red lines) for the same day at times before 12:00 (encircled in Figure 10 a,b bottom panels). As expected, analysis approaches a little better altimeter values, although differences between analysis and forecasts are very small.

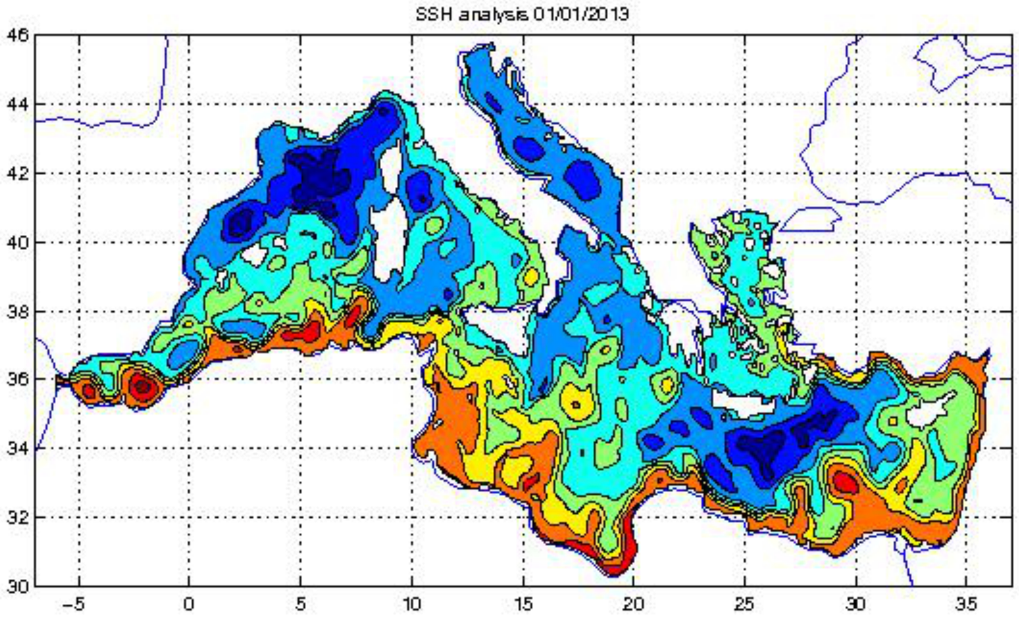
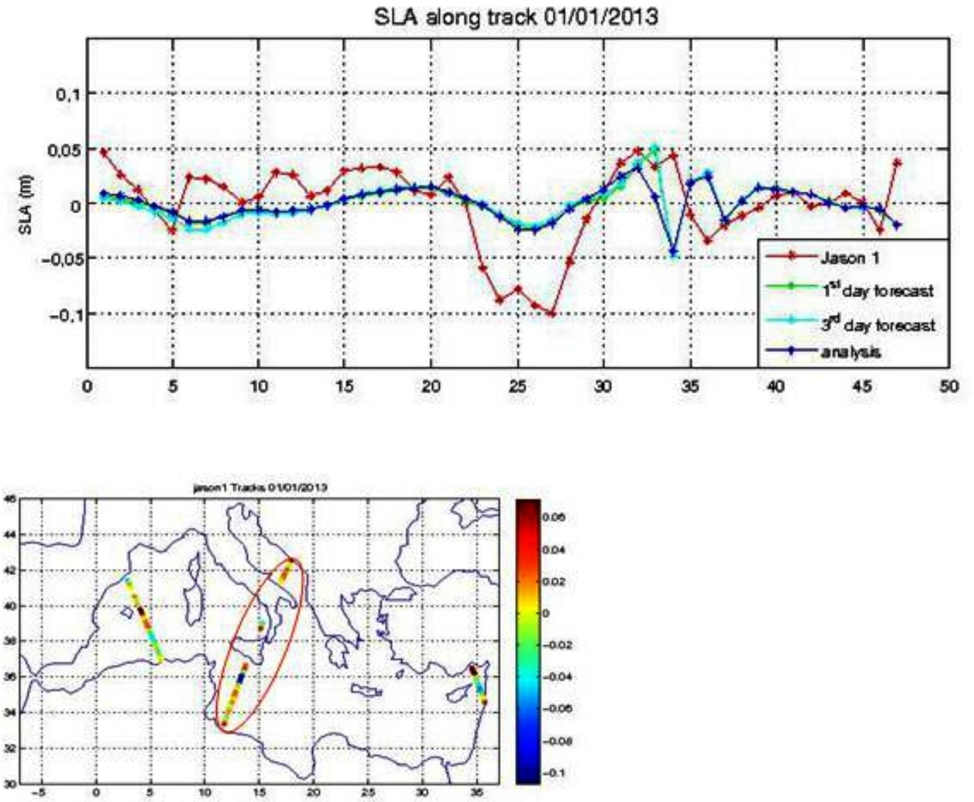
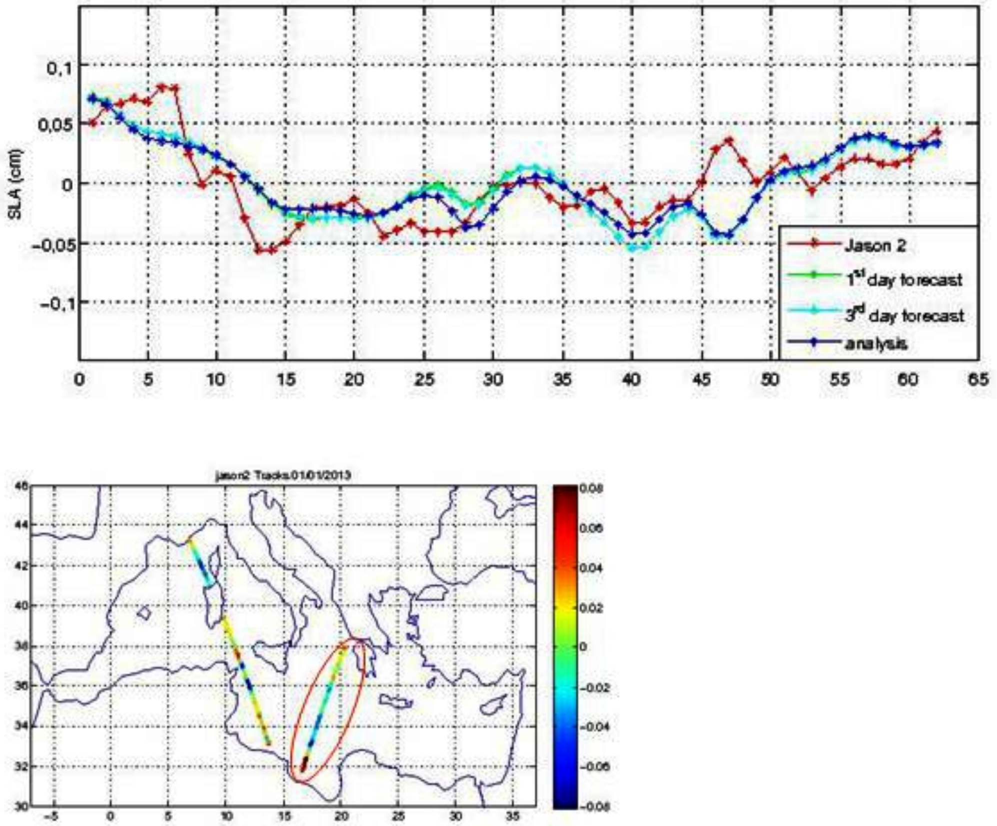


Figure 9. January 1st 2013, Analysis SSH field.



a)



b)

Figure 10. The upper panels of figure (a) and (b) show the comparison among the along track model SLA values of first (green lines) and third (light blu lines) and along track SLA (red lines) of Jason1 (a) and Jason 2 (b). The bottom panels of figure (a) and (b) show along track SLA of the two satellites available the 1st of January 2013. The tracks encircled are those displayed as red lined in the upper panels.

2.3.2. Statistics for MyOcean quarterly validation report

MyOcean 2 project aims to provide users with information about the accuracy of all products, therefore guidelines and conventions are developed for a routine validation procedure and the style of presentation of results. On MyOcean website each production centre releases a Quality Information Document (QuID) for each product, where information about the product quality are given as Estimated Accuracy Numbers (EAN) and commented. EANs are mean and root mean square of the difference between the operational product and a reference, and they are computed for particular depths and sub-regions if present. Med-MFC provides tabulated EANs for temperature and salinity at selected depth intervals and for SST.

Validation is also required as quarterly reports, where a number of statistics have to be presented to give a more informative and consistent presentation of the accuracy of products.

Statistics required for MyOcean quarterly report are:

- Number of data values used to compute statistics
- Mean of model values (product)
- Mean of observations (reference)
- Mean square error computed as

$$MeanSquareError(x,y) = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2$$

where x_i is the product (analysis or forecast) value and y_i is the reference (observation) value.

These statistics are computed daily to compare the first and third day of forecast and the analysis with SLA observations from two satellites, i.e. Jason 1 and Jason 2, covering the period January-February-March 2013. Results are shown in Fig. 2.7-2.8.

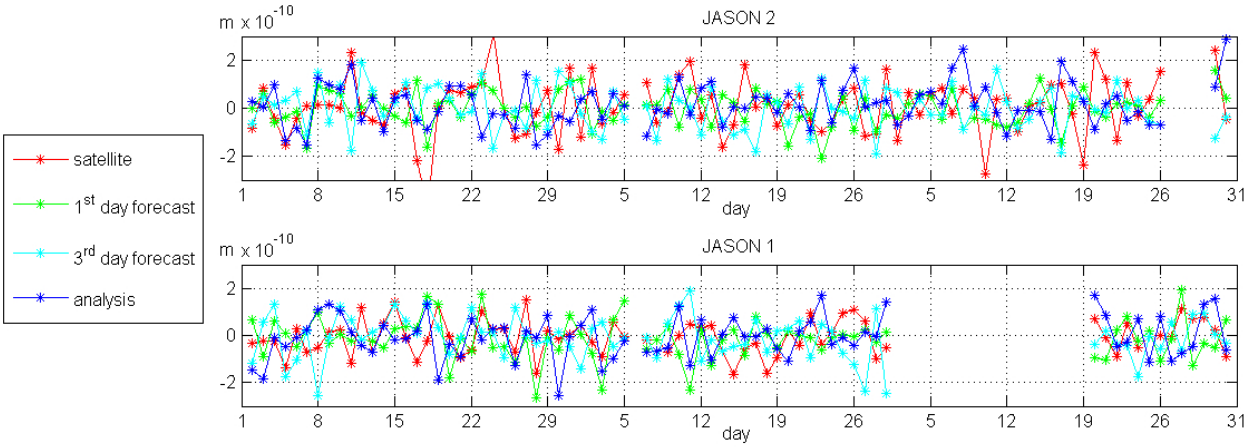


Figure 11. D Daily means of along Jason 2 tracks (a) and Jason 1g (b) SLA of observations (red line), first day if forecast (green line), third day of forecast (light blu line) and analysis (blu line).

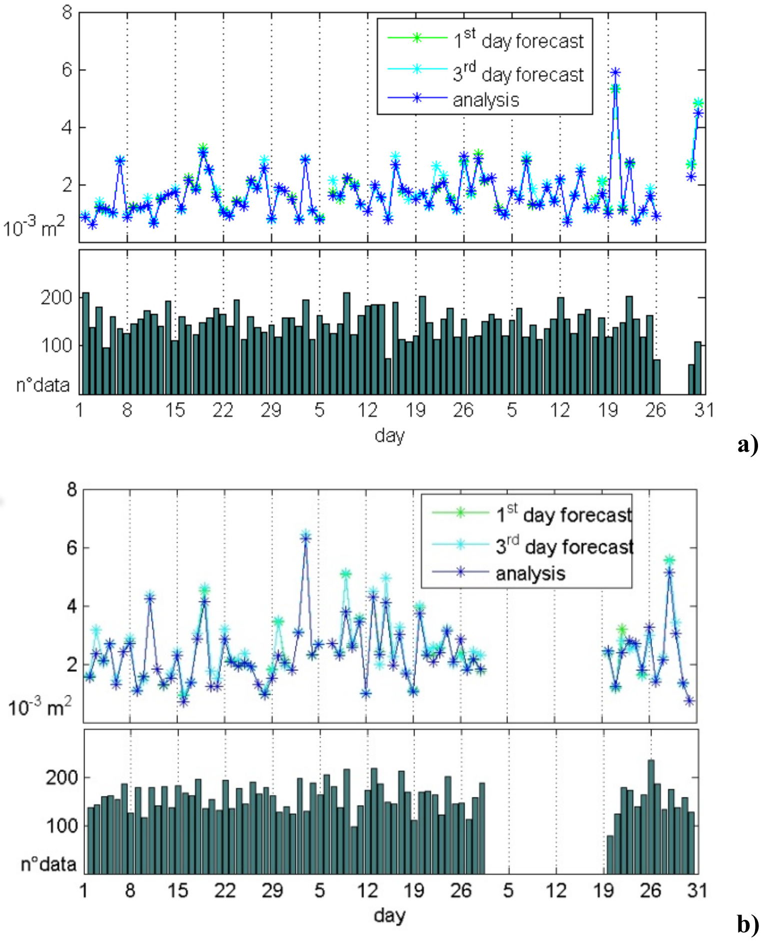


Figure 12. Daily MSE SLA differences along Jason2 (a) and Jason1 (b) tracks computed for the differences between satellite and the three forecast products.

2.4. Case study of the impact of multi-altimeter sea level assimilation

Estimates of RMS of misfit (observations - model values) are used to test the impact of assimilation of one to three satellites on the analysis quality. Four experiments(A, B, C and D) were carried out in the period from the 26th of June to the 5th of October 2013, assimilating along track SLA from a different number of satellites. The satellite assimilated in each experiment are:

A. Jason 2;
B. Jason 2 and AltiKa;
C. Jason 2 and Cryosat;
D. Jason 2, Cryosat and AltiKa.

Misfits are obtained as described in chapter *2.3.1*, interpolating SLA analysis values along satellite tracks, and RMS is computed for each experiment over all misfits in a day for each satellite independently. The along track bias is subtracted to each track-leg. If a track-leg crosses a island or a peninsula of significant dimension (the check is done over a 100km distance), the track-leg is divided in sub-track-leg and the bias computed and subtracted for each sub- track-leg. The tracks in Figure 10, for example, have respectively 4 sub-track-leg in panel a) and 3 sub-track-leg in panel b).

The following equations described in detail the computation of the along track RMSE.

$$DIFF_i^{leg1} = SLA_i - \left(SSH_i^{obs} - MDT_i^{obs}\right)$$

$$BIAS^{leg1} = \sum_{i=1}^{i=nobs} DIFF_i \Big/ nobs$$

$$RES^{leg1} = \frac{DIFF_{i=1}^{leg1}}{BIAS^{leg1}}, \frac{DIFF_{i=2}^{leg1}}{BIAS^{leg1}}, \dots\dots, \frac{DIFF_{i=nobs}^{leg1}}{BIAS^{leg1}}$$

$$RES^{leg2} = \frac{DIFF_{i=1}^{leg2}}{BIAS^{leg2}}, \frac{DIFF_{i=2}^{leg2}}{BIAS^{leg2}}, \dots\dots.., \frac{DIFF_{i=nobs}^{leg2}}{BIAS^{leg2}}$$

$$RES^{legn} = \frac{DIFF_{i=1}^{legn}}{BIAS^{legn}}, \frac{DIFF_{i=2}^{legn}}{BIAS^{legn}}, \dots\dots, \frac{DIFF_{i=nobs}^{legn}}{BIAS^{legn}}$$

$$RES = RES^{leg1}, RES^{leg2}, \dots\dots., RES^{legn}$$

$$RMSE = \sqrt{\frac{\sum RES^2}{nn}}$$

where $nobs$ is the number of observations along a single leg track; SLA_i is the value of along track observation at point i, SSH_i^{obs} is the model SSH interpolated at the observation position, MDT_i^{obs} is the Mean Dynamic Topography, which is on the model grid, interpolated at the observation position and nn is the number of observation along all the legs available for the 24hr corresponding to the model product time range.

Observations from all the three considered satellite were available during the period of the experiment (see Figure 14) except for 5 days in September during which Jason2 was not transmitting (see Figure 14 and upper right panel of Figure 13). The weekly sum of the number of observation is around 900 for Jason2 and Altika and more than 1000 for Cryosat.

The RMSE time series for the four experiments (A, B, C and D) is shown in Figure 13. The RMSE computed for each day of analysis respect each satellite is shown in the first three panels. The last (bottom-right) panel shows the RMSE computed respect all the three satellite tracks. The mean RMSE over the whole experiment period are shown in Table 4. The analysis performs better increasing the number of satellite. When only two satellite are assimilated (exp B and exp C), the experiment B performs slightly better than experiment C probably due to the higher measurement error associated to Cryosat. The analysis decrease its error of ca. 0.4 introducing the observation of Altika and Cryosat in the assimilation cycle.

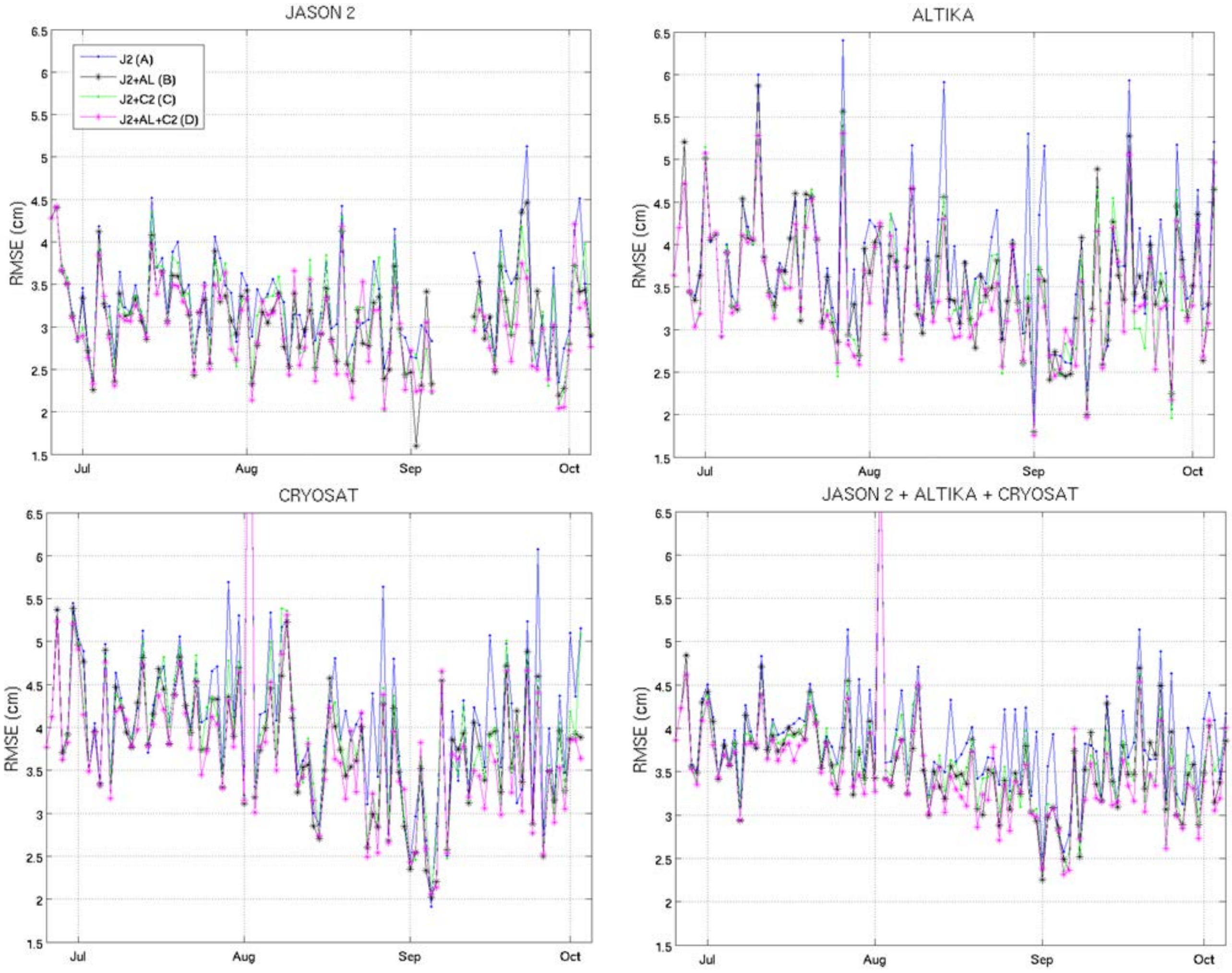


Figure 13. Along track RMSE for the four different experiments (A=assimilation only of Jason2, B=assimilation of Jason 2and Altika, C=assimilation of Jason2 and Cryosat, D=assimilation of Jason2, Altika, Cryosat). The different panels show tha along track RMSE for each satellite and for all the three satellites.

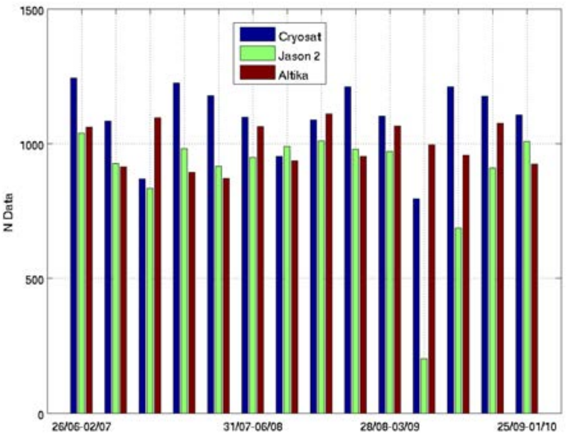


Figure 14. Weekly sum of the number of observations assimilated in Experiment A, B, C and.

Table 4. RMSE (cm) along the track of each satellite (Jason2, Altika and Cryosat) and along all the tracks for the 4 differnt experiments.

	JASON2 (J2)	ALTIKA (AL)	CRYOSAT (CR)	J2+AL+CR
Exp(A)	3.3cm	3.8cm	4.1cm	3.8cm
Exp(B)	3.1cm	3.6cm	3.9cm	3.6cm
Exp(C)	3.1cm	3.5cm	3.9cm	3.6cm
Exp(D)	3.0cm	3.4cm	3.8cm	3.5cm

This result is confirmed also from the Taylor Diagram (Figure 14) where experiment D has always the smaller RMSE and the higher correlation.

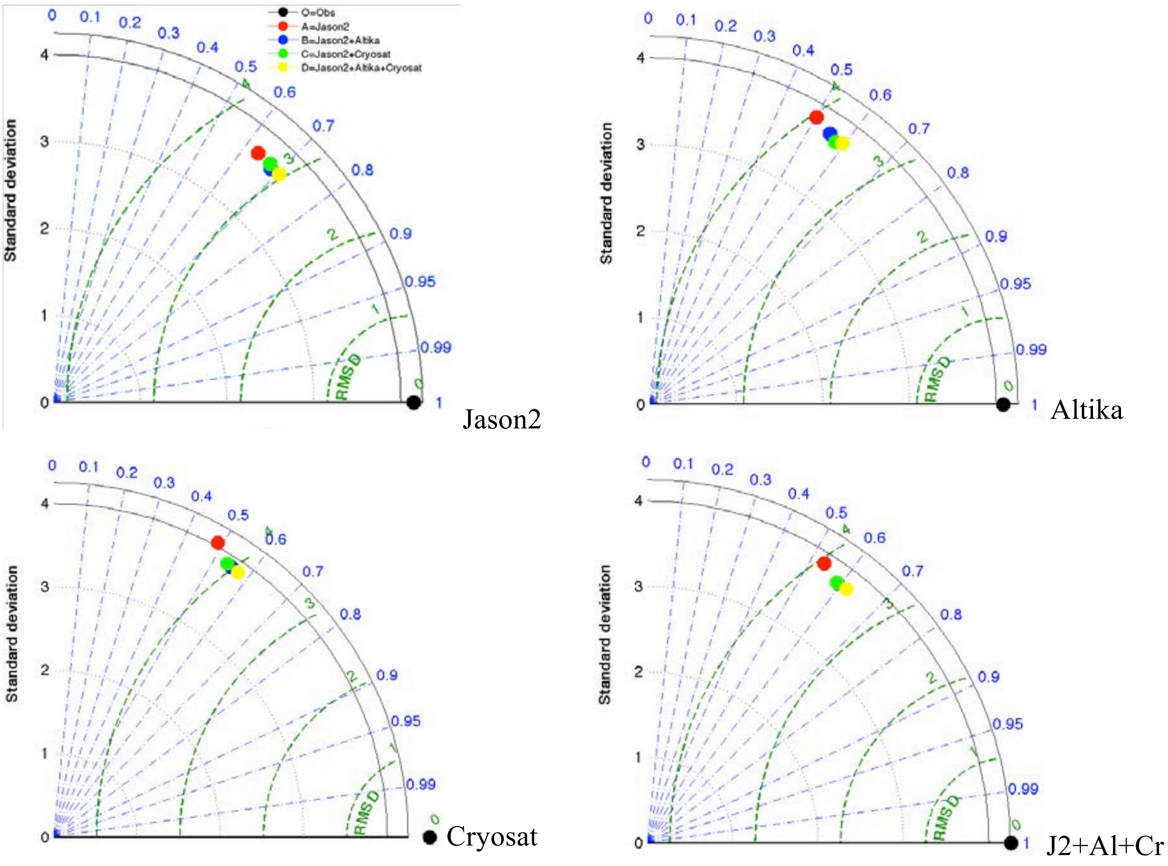


Figure 15. Taylor diagram for the four experiments (A, B, C and D).

3. Output files format

Results of verification procedures have to be provided from all forecasting systems as netCDF files using the same conventions about dimensions, variables and attributes names.

MyOcean quarterly report contains time series of the statistics, while the file required for GODAE inter-comparison contains observations used as reference and forecast products interpolated on observation positions.

Below the headers of the two netCDF files containing statistics of SLA product and along track SLA values of the first and third day of forecast and analysis and observations from Jason 2 and Jason 1 satellites are shown.

Header of the netCDF file containing statistics for MyOcean quarterly validation report

```
netcdf product_quality_stats_SLA_01012013_31032013 {
dimensions:
string_length8 = 18 ;
areas = 1 ;
metrics = 4 ;
forecasts = 3 ;
surface = 1 ;
time = UNLIMITED ; // (0 currently)
variables:
char area_names(string_length8, areas) ;
area_names:long_name = "Mediterraenan Sea" ;
area_names:descriptions = "region over which statistics are aggregated" ;
char metric_names(time, string_length8) ;
metric_names:long_name = "metric names" ;
float forecasts(forecasts) ;
forecasts:long_name = "forecast lead time" ;
forecasts:units = "days" ;
float time(time) ;
time:long_name = "validity time" ;
time:units = "days since 1950-01-01 00:00:00 UTC" ;
float stats_sla_jason2(time, forecasts, surface, metrics, areas) ;
stats_sla_jason2:_FillValue = -999. ;
stats_sla_jason2:parameter = "SLA" ;
stats_sla_jason2:reference = "Jason2 along track SLA" ;
stats_sla_jason2:units = "m" ;
float stats_sla_jason1g(time, forecasts, surface, metrics, areas) ;
stats_sla_jason1g:_FillValue = -999. ;
stats_sla_jason1g:parameter = "SLA" ;
stats_sla_jason1g:reference = "Jason1g along track SLA" ;
stats_sla_jason1g:units = "m" ;

// global attributes:
:contact = "alessandro.grandi@bo.ingv.it" ;
:product = "MEDSEA_ANALYSIS_FORECAST_PHYS_006_001_a" ;
:creation_date = "Created Wed Nov 27 15:29:10 2013" ;
:start_date = "20130101" ;
:end_date = "20130331" ;

}
```

Header of the netCDF file containing observations and forecast values for GODAE class 4 metrics inter-comparison

```
netcdf class4_20130101_MFS_SYS4b_2.0_SLA {
dimensions:
numobs = 150 ;
numvars = 1 ;
numdeps = 1 ;
string_length8 = 8 ;
string_length28 = 28 ;
nfumfcsts = 3 ;
variables:
float best_estimate(numdeps, numvars, numobs) ;
best_estimate:_FillValue = -999. ;
best_estimate:long_name = "Model forecast day offset" ;
best_estimate:comment = "Days between forecast production and validity time" ;
best_estimate:Units = "Days" ;
float forecast(numdeps, nfumfcsts, numvars, numobs) ;
forecast:_FillValue = -999. ;
forecast:long_name = "Model forecast counterpart of obs. value" ;
forecast:comment = "Model daily mean valid at noon used for calculation" ;
char id(string_length8, numobs) ;
id:long_name = "Observation id" ;
double modeljuld(numobs) ;
modeljuld:long_name = "Model field date in Julian days" ;
modeljuld:units = "Days since 1950-01-01 00:00:00 UTC" ;
float latitude(numobs) ;
latitude:long_name = "Latitudes" ;
latitude:units = "Degrees" ;
float longitude(numobs) ;
longitude:long_name = "Longitudes" ;
longitude:units = "Degrees" ;
float observation(numdeps, numvars, numobs) ;
observation:_FillValue = -999. ;
observation:long_name = "Observation value" ;
double juld(numobs) ;
juld:long_name = "Observation time in Julian days" ;
juld:units = "Days since 1950-01-01 00:00:00 UTC" ;
char type(string_length28, numobs) ;
type:long_name = "Observation type" ;
char unitname(string_length8, numvars) ;
unitname:long_name = "Unit name" ;
char varname(string_length8, numvars) ;
varname:long_name = "Variable name" ;
float leadtime(nfumfcsts) ;

// global attributes:
:title = "forecast class 4 file" ;
:version = "1" ;
:creation\ date = "2013/11/26" ;
:contact = "alessandro.grandi@bo.ingv.it" ;
:obs_tipe = "Jason 2 along track Sea Level Anomaly" ;
:system = "MFS" ;
:configuration = "SYS4b2" ;
:institution = "Istituto Nazionale di Geofica e Vulcanologia - Bologna, Italy" ;
:validity_time = "2013-01-01 00:00:00 utc" ;
:best_estimate_description = "analysis produced from 8 to 15 days behind real time" ;
:time_interp = "daily average fields" ;
:NCO = "4.1.0" ;
}
```

APPENDIX
Computer programs used/implemented for this work

Calculations for this work have been conducted on the Ekman HPC Cluster, that is part of the computing infrastructure of the Operational Oceanography Project Unit of INGV Bologna section.
Two different procedures and sets of programs have been used/implemented:

1. To compute statistics for MyOcean quarterly report, model forecast and analysis daily mean fields are put in different directories named by the day to which the field refers. These directories include the files that contain satellite observations of the same day and of the day before. The path is */home/trani/statistiche/yymmdd*.
   The directory */home/train/statistiche/OUTPUT* contains: the codes and the relative executable files of the two programs (slasysqbdd_4MDTstat.f90 and sla_qrep_4ms.f90) – the former that matches model values with observations calculating misfits and the latter that computes statistics - a shell script (runconfsla_4int.sh) that gets input files, runs the two programs and put output files in the */home/trani/statistiche/yymmdd* directories; the netCDF file of the MDT field and a template, i.e. a file containing all zero values that is get by the program to replace a missing satellite file (for example, if satellite data are not available for a particular day).
   A second version of the program that computes statistics, named sla_qrep_4ms_wr.f90, contains also some lines of code which write the header of the netCDF file requested by MyOcean for the quarterly report. This program has to be completed by adding data to the variables.

2. To compute RMS for the multialtimeter assimilation experiments, analysis daily mean field outputs for each experiment and files of satellite observations of the same day and of the day before are put in different subdirectories, named by the day to which the field refers, contained in a directory with the name of the relative experiment (example for experiment A: */home/trani/esperimenti/output/ Aout/yymmdd*).
   The directory */home/trani/esperimenti* contains: the MDT field, the code and the relative executable file that matches model values with observations calculating misfits (slasysqbdd_8MDTexp.f90); and a shell script (runconfsla_mt8.sh) that gets input files, runs the program and put output files in the same directories of the input files. Also, there is a shell script (slaexp_merge8.sh) that copies each output file from the different directories named by the date in the directory named by the experiment.
   Two matlab programs (biasrmse1_all.m and biasrmse7_all.m, in */home/trani*) read data in the output files – those copied in the directory named by the experiment – and compute respectively the daily and weekly RMS of misfits along the tracks of a single satellite or of all satellites (shown in Fig. 2.10a,b,c,d).

References

Bengtsson L., (1981). *The impact on FGGE on global medium range forecasts. International Conference on early Results of FGGE and Large-scale Aspects of its Monsoon Experiments.* Tallahassee, Jan 1981, ICSU/WMO, Geneva, 1-26.

Brooks H.E. and Doswell C.A. (1996). *A Comparison of Measures-Oriented and Distributions-Oriented Approaches to Forecast Verification*. Weather Forecasting, 11, 288–303.

Brier G.W. and Allen R.A., (1951). *Verification of weather forecasts*. Compendium of Meteorology, T.F. Malone, American Meteorological Society, 841-848.

Crosnier L. and Le Provost C., (2007). *Inter-comparing five forecast operational systems in the North Atlantic and Mediterranean basins: The MERSEA-strand1 methodology*. Journal of Marine Systems, 65, 354-375.

Crosnier L., Le Provost C. and MERSEA Strand 1 Team, (2006). *Internal metrics definition for operational forecast systems inter-comparison: Examples in the north Atlantic and Mediterranean Sea*. In: Ocean weather forecasting: an integrated view of oceanography (Chassignet and Verron eds.), pp. 455-465. Springer, France.

Dobricic S., (2005). *New mean dynamic topography of the Mediterranean calculated from assimilation system diagnostics*. Geophysical Research Letters, 32, L11606.

Dobricic S. and Pinardi N., (2008). *An oceanographic three-dimensional variational data assimilation scheme*. Ocean Modelling, 22, 89-105.

Finley J.P., (1884). *Tornado predictions*. American Meteorology Journal,1,85-88.

Gilbert G.K., (1884). *Finley's tornado predictions*. American Meteorology Journal,1,166-172.

Hernandez F., Bertino L., Brassington G., Chassignet E., Cummings J., Davidson F., Drevillon M., Garric G., Kamachi M., Lellouche J-M., Mahdon R., Martin M. J., Ratsimandresy A. and Regnier C., (2009). *Validation and inter-comparison studies within GODAE*. Oceanography, vol. 22 n.3, 128-143.

Joliffe I.T. and Stephenson D.B., (2003). *Forecast Verification*. Wiley and Sons Ltd.

Mellor G.L. and Ezer T., (1995). *Sea level variations induced by heating and cooling: an evaluation of the Boussinesq approximation in ocean models*. Journal of Geophysical Research, 100 (C10), 20 565- 20 577.

Madec G. and the NEMO team, (2008). *NEMO ocean engine*. Note du Pole de modelisation, Institut Pierre_Simon Laplace (IPSL), France, 27, ISSN 1288-1619.

Murphy A.H., (1988). *Skill scores based on the mean square error and their relationships to the correlation coefficient*. Monthly Weather Review, 116, 2417-2424.

Murphy A.H., (1991). *Forecast verification: its complexity and dimensionality*. Montly Weather Review, 119, 1590-1601.

Murphy A.H., (1993). *What is a good forecast? An essay on the nature of goodness in weather forecasting*. Weather and Forecasting, 8, 281-293.

Murphy A.H., (1996). *The Finley affair: a signal event in the history of forecast verification*. Weather and Forecasting, 11, 3-20.

Murphy A.H., (1997). *Forecast verification*. In: Economic value of weather and climate forecasts (Katz and Murphy eds.), pp. 19-74. Cambridge University Press.

Murphy A.H. and Winkler R.L., (1987). *A general framework for forecast verification*. Monthly Weather Review, 115, 1330-1338.

Oddo P., M. Adani, N. Pinardi, C. Fratianni, M. Tonani and D. Pettenuzzo, (2009). A Nested Atlantic--Mediterranean Sea general circulation model for operational forecasting, Ocean Sci., 5, 461--473.

Pujol M.-I., Dobricic S., Pinardi N. and Adani M., (2010). *Impact of multialtimeter sea level assimilation in the Mediterranean Forecasting Model*. Journal of Atmospheric and Oceanic Technology, 27, 2065-2082.

Schiller A. and Brassington G.B., (2011). *Operational Oceanography in the 21st century*. Springer.

Smith, N. and M. Lefebvre (1997). *The Global Ocean Data Assimilation Experiment (GODAE)*. In: Int. Symp. Monitoring the oceans in the 2000s : an integrated approach, 15-17October, Biarritz.

Tolman H.L., (2009). *User manual and system documentation of WAVEWATCH III™ version 3.14*. NOAA / NWS / NCEP / MMAB Technical Note, 276.

Tonani M. and Grandi A., (2013). *Quality Information Document for Mediterranean Sea physical analysis and forecasting product*. MYO2-MED-QUID-006-001-V3.1 in MyOcean website.

Tonani M., Grandi A., Fratianni C. and Drudi M., (2010). *European Marine Core Service Mediterranean Monitoring and Forecasting Centre (MFC): manuale d'uso del Mediterranean Forecasting System (Med-MFC-currents)*. INGV Technical Report, 163.

Tonani M., N.Pinardi, C. Fratianni, J. Pistoia, S. Dobricic, S. Pensieri, M. de Alfonso and K. Nittis, (2009). Mediterranean Forecasting System: forecast and analysis assessment through skill scores, Ocean Sci. 5,649--660.

Web Sites

http://www.myocean.eu
http://gnoo.bo.ingv.it/mfs/myocean/calval/
http://gnoo.bo.ingv.it/mfs/myocean/evaluation.html
http://www.aviso.oceanobs.com

Coordinamento editoriale e impaginazione
Centro Editoriale Nazionale | INGV

Progetto grafico e redazionale
Daniela Riposati | Laboratorio Grafica e Immagini | INGV

Via di Vigna Murata, 605
00143 Roma
Tel. +39 06518601 Fax +39 065041181

http://www.ingv.it

Istituto Nazionale di Geofisica e Vulcanologia