Ocean Sci., 8, 211–226, 2012
www.ocean-sci.net/8/211/2012/
doi:10.5194/os-8-211-2012

Ocean Science

ENSURF: multi-model sea level forecast – implementation and validation results for the IBIROOS and Western Mediterranean regions

B. Pérez[1], R. Brouwer[2], J. Beckers[2], D. Paradis[3], C. Balseiro[4], K. Lyons[5], M. Cure[5], M. G. Sotillo[1], B. Hackett[6], M. Verlaan[2], and E. A. Fanjul[1]

[1]Organismo Público Puertos del Estado (OPPE), Madrid, Spain
[2]Deltares, Delft, The Netherlands
[3]Météo-France, Toulouse Cedex, France
[4]MeteoGalicia, Santiago de Compostela, Spain
[5]Irish Marine Institute, Co. Galway, Ireland
[6]Norwegian Meteorological Institute, MET-NO, Norway

Correspondence to: B. Pérez (bego@puertos.es)

Received: 1 March 2011 – Published in Ocean Sci. Discuss.: 8 April 2011
Revised: 18 January 2012 – Accepted: 1 February 2012 – Published: 30 March 2012

Abstract. ENSURF (Ensemble SURge Forecast) is a multi-model application for sea level forecast that makes use of several storm surge or circulation models and near-real time tide gauge data in the region, with the following main goals:

1. providing easy access to existing forecasts, as well as to its performance and model validation, by means of an adequate visualization tool;
2. generation of better forecasts of sea level, including confidence intervals, by means of the Bayesian Model Average technique (BMA).

The Bayesian Model Average technique generates an overall forecast probability density function (PDF) by making a weighted average of the individual forecasts PDF's; the weights represent the Bayesian likelihood that a model will give the correct forecast and are continuously updated based on the performance of the models during a recent training period. This implies the technique needs the availability of sea level data from tide gauges in near-real time. The system was implemented for the European Atlantic facade (IBIROOS region) and Western Mediterranean coast based on the MATROOS visualization tool developed by Deltares. Results of validation of the different models and BMA implementation for the main harbours are presented for these regions where this kind of activity is performed for the first time. The system is currently operational at Puertos del Estado and has proved to be useful in the detection of calibration problems in some of the circulation models, in the identification of the systematic differences between baroclinic and barotropic models for sea level forecasts and to demonstrate the feasibility of providing an overall probabilistic forecast, based on the BMA method.

1 Introduction

The increase in computing and networking facilities over the last decades has made advances possible in operational sea level forecasting, using numerical models that account for astronomical tide and meteorological forcing. These systems have become critical for some countries where the magnitude of storm surges can reach over 3 m in occasions and cause considerable inundation and damage along the coast. Countries surrounding the North Sea, for example, where the land is both low-lying and densely populated, and historic storm surges have caused thousands of deaths, have dedicated storm surge warning services that rely on these models. More recently, also regions that are less prone to these dramatic events have begun to make use of these forecasts, such as the Mediterranean coast where the meteorological component is of the same order of magnitude as the tide and the forecasts are used for large vessel manoeuvring inside harbours or for dredging operations.

For many regions, a collection of models, statistical methods and post-processing techniques is available for describing the response of the sea level to an imposed weather field and astronomical tide. Barotropic 2-D models take into consideration irregular boundaries and variable water depth which affect surge propagation and magnitude and have proved to be adequate for this application during the last 30 yr (Flather, 1981, 1987; Alvarez-Fanjul et al., 1997, 2001) and have been the basis of the existing operational sea level forecasts up to now. On the other hand, more recent improvements in computer skills have allowed the development of 3-D baroclinic models for ocean circulation forecasts; their operational implementation has led to the availability of alternative sea level forecasts in some regions. For these general circulation models, a validation of sea level output is critical for a correct characterization of the sea surface elevation and consequently for an adequate description of the circulation patterns. However, it is well known that these 3-D circulation models do not generally perform better for storm surge simulations, although they include a more complete description of the physical processes that produce sea level variations, something we confirm within the ENSURF application for the IBIROOS region in this paper. Nevertheless, they do provide a sea level forecast that could be considered as an additional source of information.

Despite careful calibration, these numerical models often present a bias with respect to observations. This may be corrected for by making use of data-assimilation or post-processing techniques, which include information from real time tide gauge or altimetry data into the forecast. Thus, an optimal operational sea level forecasting system can be based on a combination of numerical models and observations.

Storm surge and ocean circulation forecasts are generated and distributed by several operational centres throughout Europe, each using their own forecasting system. Usually these systems provide deterministic and independent forecasts of sea level for their specific regions, sometimes geographically overlapping in part. Their mutual comparison and, if possible, integration, in order to improve their skills at the common domains or points, pose a new challenge. Recent studies have demonstrated the advantages of the multi-model and the ensemble approach for validation and improvement of predictive capabilities. This provided the rationale for the creation of the ENSURF system (Ensembles SUrge Forecast), within the ECOOP European project (European Coastal-shelf sea Operational observing and forecasting system), Contract No. 3655, whose overall goal is to consolidate, integrate and further develop existing European coastal and regional seas operational systems. ENSURF constitutes one of the main products of this project (http://www.ecoop.eu/summary.php), as it represents a perfect example of this integration, not only because it involves different forecasting systems, but also because it makes use of observations and new statistical techniques that may improve the independent forecasts. This is something that could be valuable for other operational systems. In the particular region studied in this paper, an improvement of the forecasts at the harbours is found often with this integration.

2 ENSURF system: objectives and general description

ENSURF is a multi-model application for sea level forecast that makes use of some existing storm surge/circulation models currently operational in Europe, as well as near-real time tide gauge data in the region. The application was first implemented for the NOOS region, which is running operationally at Deltares (http://noos.deltares.nl). It involves an integration of existing operational sea level forecasts, with potential for relocation in new coastal areas and the following main objectives:

1. providing easy access to existing forecasts as well as to the performance and validation of the different models through a common visualization tool;
2. generation of overall probabilistic forecasts of sea level, including confidence intervals, by means of statistical post-processing techniques such as the Bayesian Model Average (BMA);
3. becoming a joint European service in the framework of the ECOOP project.

In this paper, we describe the implementation of ENSURF for the IBIROOS and Western Mediterranean regions by Puertos del Estado (Fig. 1). The reason for the two separate implementations at Deltares and Puertos del Estado was the different status and experience on sea level data exchange policy, both from models and observations, in the two regions. Initially, it was not possible to develop a component for the MOON region, due to an insufficient number of operational models with sea level output in the Mediterranean Sea. Nevertheless, in this work we have included the Western Mediterranean, where Spanish and French forecasts and data were available. The system has shown its usefulness as a user-friendly operational validation tool, and its ability to provide a probabilistic forecast by means of the Bayesian Model Average Technique. It is the first time such a kind of tool has been implemented for sea level forecasting in the South of Europe.

For the ENSURF implementation for the IBIROOS region, it was necessary to select the locations of the storm surge forecasts for available tide gauges and to establish the data exchanges (real time measurements and forecasts) between partners. The system is currently running operationally at Puertos del Estado (http://ensurfibi.puertos.es) and ready to incorporate more stations and sources in the future (Fig. 2). We present the first validation results of the different models

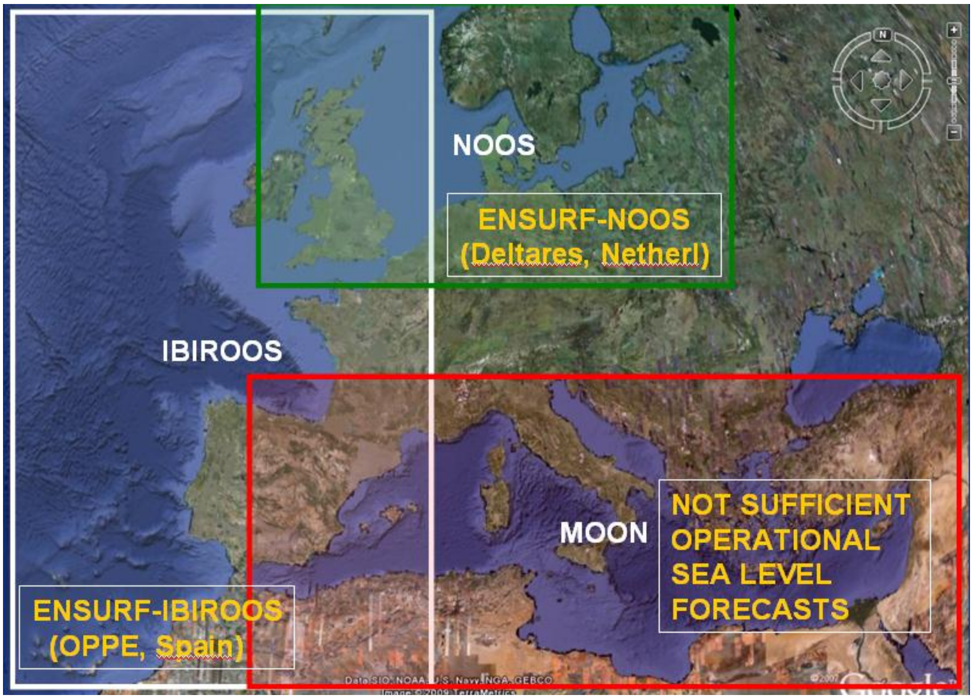

NOOS
ENSURF-NOOS (Deltares, Netherl)
IBIROOS
MOON
NOT SUFFICIENT OPERATIONAL SEA LEVEL FORECASTS
ENSURF-IBIROOS (OPPE, Spain)

Fig. 1. ENSURF implementation for the main operational oceanographic regions in Europe.

and the performance of the Bayesian Model Average Technique for a specific period in this region.

ENSURF is based on the MATROOS (Multifunctional Access Tool for Operational Ocean Data Services) visualization tool developed by Deltares. It is installed on a server where automatic scripts handle the acquisition of data from both models and tide gauges via ftp sites maintained by the partners. So the first step is establishing the adequate data exchange and formats for an operational integration into our system. Through this scheme, both time series of data (forecasts and observations) and forecasted fields can be included in an internal database, allowing easy access and visualization by remote users (Fig. 3).

The models output can be simply the surge component (when they are forced just with meteorological forecasts) or the total sea level (including the tide). In the first case the tide needs to be added later in order to provide a total sea level forecast. Some of the models are run in barotropic mode which is normally sufficient for storm surge applications, while other forecasts are generated from general circulation or baroclinic models which, in principle, include all the different sea level signals (e.g. density changes). For the first time, all these different applications can be validated in near-real time thanks to the ENSURF system.

2.1 ENSURF-IBIROOS sources and data

The sources currently contributing operational sea level forecasts to the IBIROOS component of ENSURF are shown in Table 1. As already mentioned, the characteristics of the models differ, some being barotropic and others baroclinic, with different resolutions and bathymetry, and with normally different model forcings. They also lead to different outputs of sea level, depending on just having meteorological forcing or including the tide.

Of course, implemented by different institutions in different countries, the domains of the models are also diverse (Fig. 4), although sharing part of the coastline in some cases; these will be the coastlines and harbours where the advantage of multi-model approach to improve the forecasts will be explored. A brief description of each source, without entering into too many details, is given below.

2.1.1 Nivmar system

In operation since 1998 at Puertos del Estado, it is based on the HAMSOM circulation model and the use of near-real time tide gauge data from the REDMAR network (Alvarez Fanjul et al., 2001). The model is run vertically integrated in barotropic mode, forced only with meteorological data (atmospheric pressure and wind) from the HIRLAM meteorological model (Undén et al., 2002). The forecast is run twice a day (00:00 and 12:00 UTC cycles), with a 72 h forecast horizon, and the domain covers the Spanish Atlantic coast and Canary Islands as well as the whole Mediterranean Sea.

HAMSOM (Backhaus, 1983; Rodríguez et al., 1991; Alvarez et al., 1997) uses a finite difference semi-implicit scheme on a variable size grid, being the resolution of the

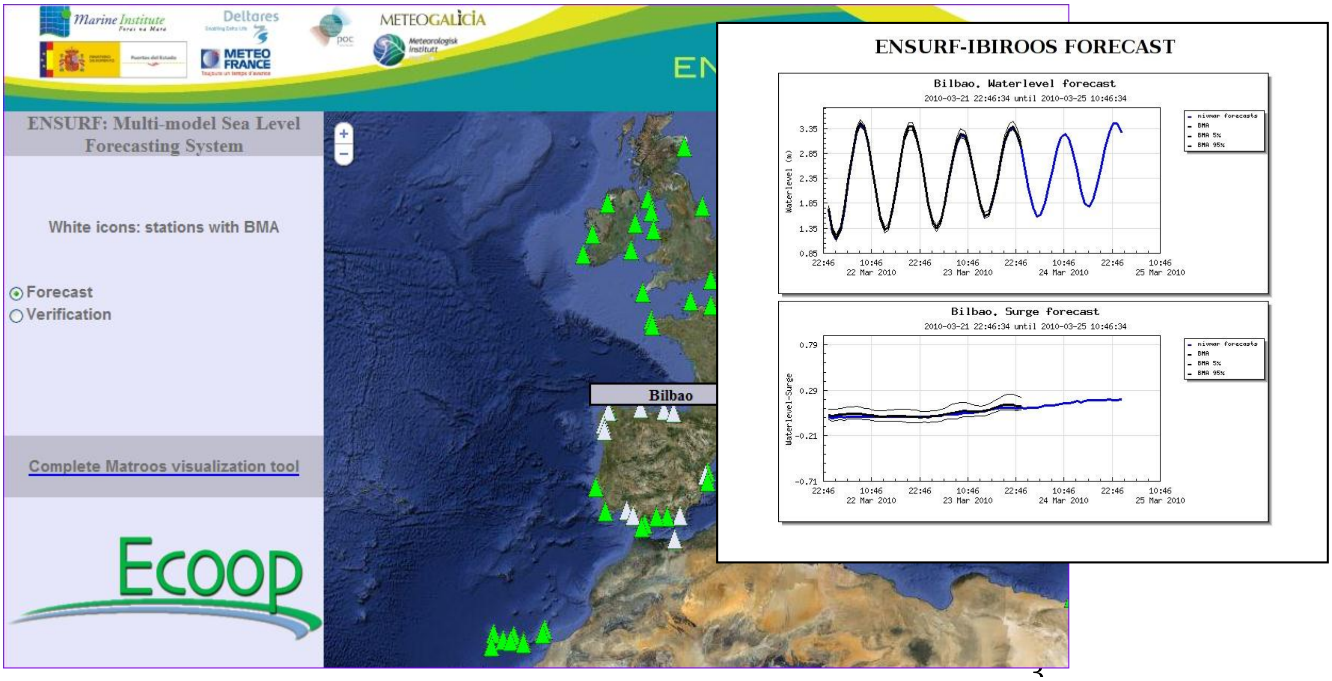


Fig. 2. ENSURF IBIROOS portal (http://ensurfibi.puertos.es) allows the display of the forecast including the BMA confidence interval or the verification with tide gauge data. The figure shows an example of forecast for Bilbao station, both for total and meteorological sea level.

Table 1. Sources or models contributing to ENSURF-IBIROOS. Name of the sources are the ones used within MATROOS visualization tool.

Institution/Country	Source/Model	Model Resolution	Met. Forcing Resolution	Output
OPPE/Spain	Nivmar/HAMSOM (barotropic) Eseoat/POLCOMS (baroclinic)	$10' \times 15'$ $3'$	0.16° 0.16°	Surge and total sea level Total sea level
Météo-France/France	Metfr_arpege/MF model (barotropic) Metfr_aladin/MF model (barotropic) Metfr_ecmwf/MF model (barotropic)	$5'$ $5'$ $5'$	0.25° 0.10° 0.5°	Surge Surge Surge
Marine Institute/Ireland	Imi/ROMS (baroclinic)	$0.6'$–$1.4'$	0.5°	Total sea level
MeteoGalicia/Spain	Metga_sm/MOHID (barotropic)	$3.6'$	0.3°	Surge

central area of the domain $10' \times 15'$ for latitude and longitude respectively. For the bottom friction it makes use of a quadratic function in terms of the current velocity, and for the wind stress it uses the Charnok parameterization (Charnok, 1955), which consists of the use of a constant non-dimensional surface roughness or Charnok coefficient ($\alpha = z_0 g W^{-2}$, where z_0 is the roughness length, W the friction velocity and g the gravitational acceleration). The open boundary conditions consist of the inverted barometer effect. The HIRLAM meteorological model is a limited area model with 0.6° and 6 h spatial and temporal resolution, being run twice daily by the AEMET (Spanish Meteorological Agency).

The bathymetry employed is the DTM5 data set (GETECH, 1995). Output data are hourly values of meteorological residual at all the points of the domain (no tide) and total seal level at special points (harbours) where a tide gauge is available, which allows the addition of the tidal component derived from observations to the model result.

Nivmar includes a simple data assimilation scheme for the forecast at the harbours, improving the results of the predictions by correcting the mean value of the simulated residuals. The correction is done by adding a constant value which is the difference of the means of the predicted and the observed time series during a recent time window. This is in fact the same technique used by the BMA to deal with the bias problem that will be explained later.

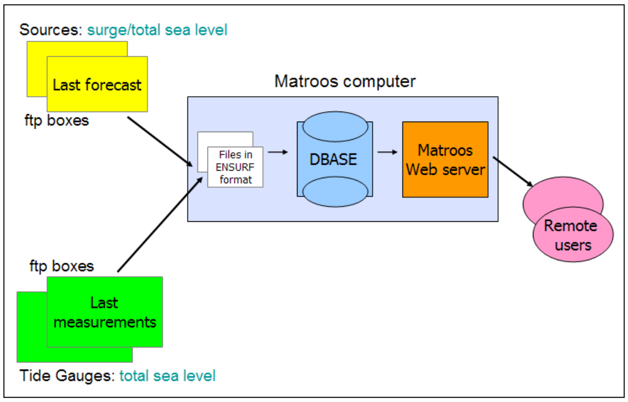


Fig. 3. ENSURF system architecture, showing the data flow and MATROOS structure.

2.1.2 ESEOAT system

ESEOAT is an ocean forecasting system operational at Puertos del Estado since 2006 (Sotillo et al., 2007, 2008). It is based on the 3-D baroclinic model POLCOMS (Proudman Oceanographic Laboratory Coastal Ocean Modelling System), which uses a finite differences scheme, and covers the Iberian Atlantic waters with an 1/20° horizontal resolution and 34 vertical S-levels. A flux/radiation boundary condition scheme is used for elevations and water column mean velocities; relaxation of temperature and salinity and inverse barometer conditions are also included at the open boundaries. Near bed velocities are computed by means of friction coefficients. The system is forced with the same meteorological fields as Nivmar (HIRLAM system). The wind stress makes use of the Charnok parameterization. Bathymetry used is derived from GTOPO30 data base and tidal forcing, based on 15 harmonic constituents imposed at the open boundaries, is also included. Hourly outputs of total sea level (including tides) and surface fields are provided, as well as daily averaged 3-D fields (temperature, salinity and currents). ESEOAT does not include tide gauge data assimilation.

2.1.3 Météo-France system

Météo-France provides three different forecasts to the system which make use of the same circulation model (the Météo-France Storm Surge model), but with different meteorological forcing. The storm surge model is a 2-D barotropic model which uses finite differences on an uniform grid, with a resolution of 5′ (around 9 km), the Chézy bottom roughness condition (Chézy, 1776) which implies the dependence of the bottom friction coefficient on a constant Chézy value for the whole domain (i.e. no depth dependency) and the Wu formulation for the wind stress (Wu, 1982), which considers it varies linearly with U_{10} wind velocity measured at 10 m above the mean sea surface. At the open boundary, an inverted barometer effect is imposed to the sea level elevation and a radiation condition is used for the current (gravity waves). Tide is included with 9 harmonic constituents, given by 17 border tide gauges (for the Atlantic only). The bathymetry is based on the GEBCO 1′ × 1′ plus local and regional fixes. The three forecasts correspond to the following meteorological forcings:

- Metfr_ecmwf: IFS: ECMWF global model with 4DVar, 25 km, 0.5° every 6 h.
- Metfr_arpege: Arpege: Météo-France global model with 4DVar, 23 km, 0.25°, every 3 h.
- Metfr_aladin: Aladin: Météo-France, LAM+3-DVar coupled by Arpege, 9 km, 0.1°, every 3 h.

The output consists of 10 min surges or meteorological sea levels at tide gauge locations and special points (harbours, vulnerable places…). No data assimilation from tide gauge data is performed.

2.1.4 IMI system

The circulation model used by the Irish Marine Institute is the Regional Ocean Modeling System (ROMS) which is a free-surface, hydrostatic, primitive equation ocean model described in Shchepetkin and McWilliams (2005). ROMS uses orthogonal curvilinear coordinates on an Arakawa-C grid in the horizontal while utilizing a terrain-following (σ) coordinate in the vertical. The model domain (NE_Atlantic) covers a significant portion of the North-West European continental shelf at a variable horizontal resolution between 1.2 and 2.5 km and with 40 σ levels. The model bathymetry utilizes data from a number of sources (e.g. Irish National Seabed Survey multibeam dataset) to produce the best possible bathymetry for the area. Surface forcing (at three-hourly intervals) is taken from the half-degree Global Forecasting System (GFS) forecast while tide forcing is prescribed at the model boundaries by applying elevations and barotropic velocities for ten major tide constituents which are taken from the TPXO7.2 global inverse barotropic tide model (Egbert and Erofeeva, 2002). The NE Atlantic model is nested within the high resolution (1/12°) Mercator Ocean PSY2V4R2 operational model of the North Atlantic whereby daily values for potential, temperature, sea surface height and velocity are linearly interpolated from the parent model onto the NE Atlantic model grid at the boundaries. Bottom stress is applied using the logarithmic "law of the wall" with a roughness coefficient of 0.01 m. Surface stress is calculated using the COARE algorithm (Fairall et al., 1996). The output consists of 10 min total sea level at tide gauge locations. No data assimilation of tide gauges is performed.

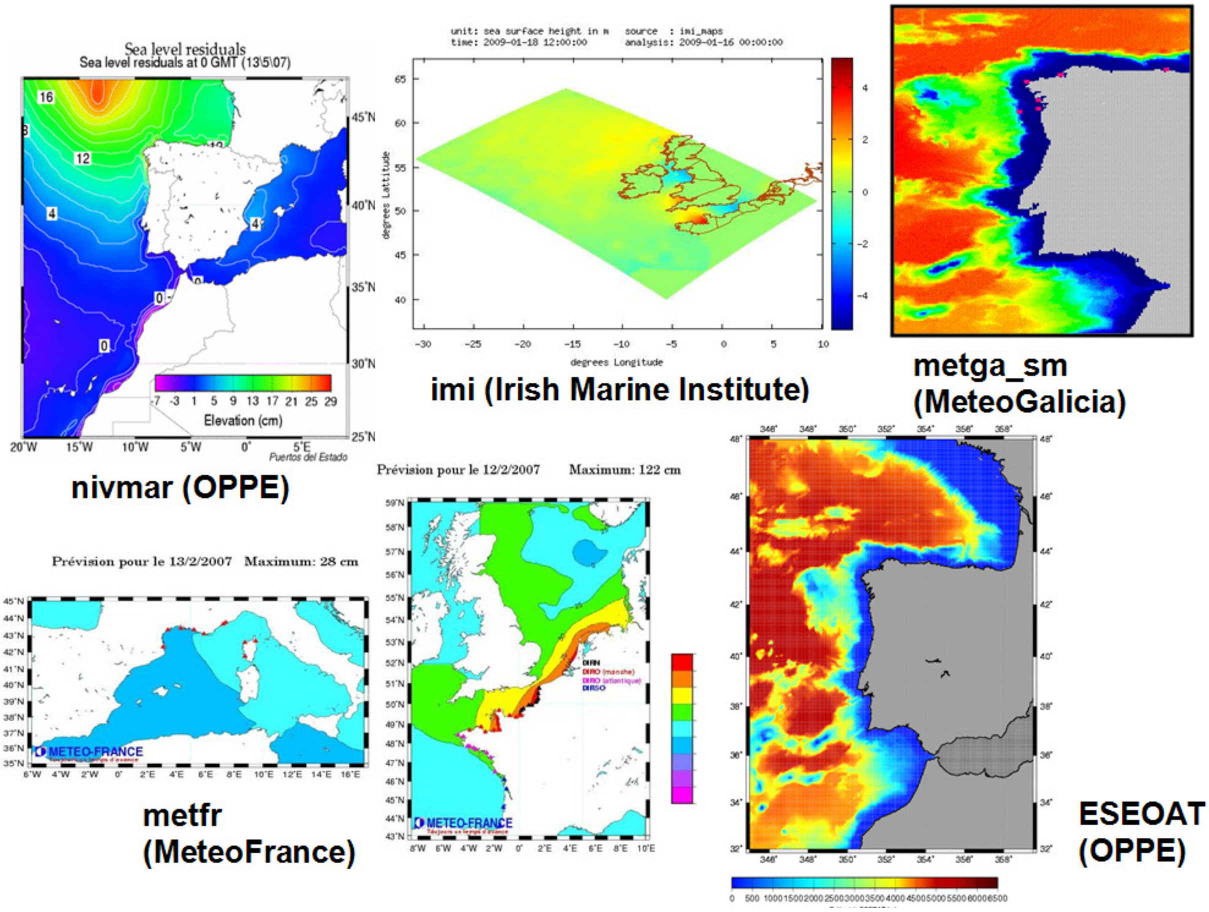

Sea level residuals
Sea level residuals at 0 GMT (13\5\07)
-7 -3 1 5 9 13 17 21 25 29
Elevation (cm)
Puertos del Estado
nivmar (OPPE)
unit: sea surface height in m source : imi_maps
time: 2009-01-18 12:00:00 analysis: 2009-01-16 00:00:00
degrees Latitude
degrees Longitude
imi (Irish Marine Institute)
metga_sm
(MeteoGalicia)
Prévision pour le 12/2/2007 Maximum: 122 cm
Prévision pour le 13/2/2007 Maximum: 28 cm
METEO-FRANCE
metfr
(MeteoFrance)
ESEOAT
(OPPE)

Fig. 4. Domains of the sources available for the ENSURF-IBIROOS component. Nivmar covers the whole Mediterranean Sea, but only results at Western Mediterranean are presented.

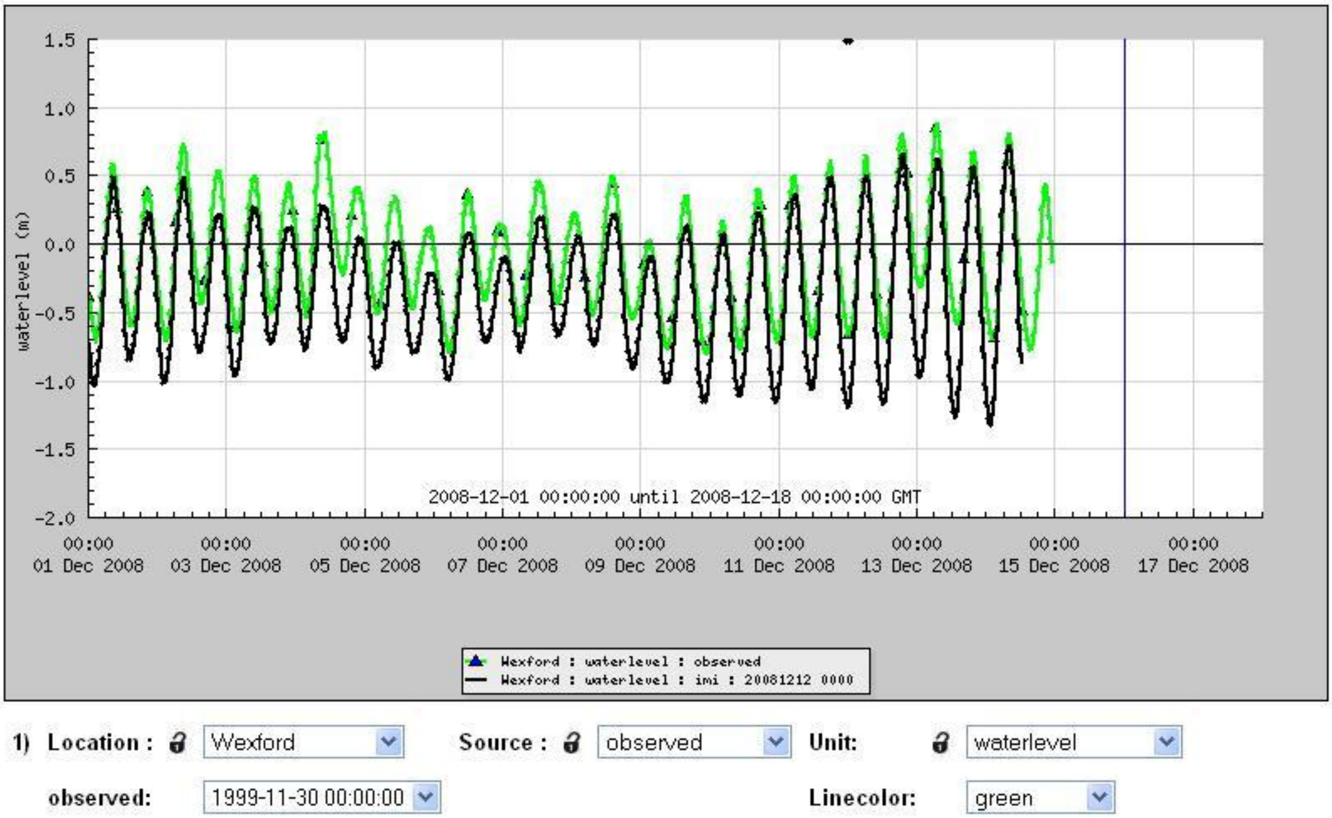
2.1.5 Metga system

The MeteoGalicia operational storm surge forecast is based on a 2-D-barotropic version of the MOHID circulation model that uses a finite volumes numerical scheme and the Large and Pond (1981) parameterization of the wind stress. Although several spatial scales have been defined with the aim of defining the storm surge processes in Galicia Coast and inside the Rias, for ENSURF just the coarse resolution grid (0.06°) covering the Iberian Peninsula is used. The bathymetries were obtained without any type of filtering based on the GEBCO arc-second dataset and data from local nautical charts to correct near coast zones. Meteorological forcing is provided by the local atmospheric model, WRF, with boundary conditions provided by the GFS global model. The WRF model is running daily in 3 nested grids with 36, 12 and 4 km resolution forcing the different MOHID scales with 1h temporal resolution. Also an inverted barometer effect is imposed at the open boundary. The system produces daily three-day forecasts with hourly values of meteorological sea level and current velocity fields, as well as surface elevation maps.

2.1.6 Tide gauge data

A common set of tide gauge stations was selected for reception of sea level data in near-real time. All the models must provide output from these special points if they fall within the model domain; the purpose of this is not just the validation of the different models with observations at the harbours, but also the implementation of the Bayesian Model Average Technique (BMA) for statistical forecast at these specific locations as will be explained later. The important role of tide gauge data for improving sea level forecasts at the coast has been recognized in the implementation of the Nivmar system (Alvarez-Fanjul et al., 2001), for example, and it is also mentioned by Mourre et al. (2006), who found how the use of tide gauges led to better global statistical performance of high-frequency barotropic models.

Data sampling can vary from 10 to 60 min (multiples of 10), and latency required can be of several hours. Automatic quality control of data in near-real time was implemented for this ENSURF-IBIROOS component, to avoid wrong values entering MATROOS and affecting model calibration and BMA results. Time needs to be Universal Time. As will be



1) Location : Wexford Source : observed Unit: waterlevel

observed: 1999-11-30 00:00:00 Linecolor: green

Fig. 5. Example of bias between Irish Marine Institute forecast (*imi*source, black) and tide gauge observations (green) at Wexford harbour.

explained in the next section, for each individual tide gauge entering ENSURF at least one year of data is required for previous computation of the tide. This will be needed to compute the total sea level provided by the system at a particular harbour and also for the implementation of near-real time quality control of the observations. Sea level data from tide gauges have been kindly provided by the following institutions: SHOM (France), POL (UK), DMI (Denmark), Marine Institute (Ireland), Geographic Institute (Portugal) and Puertos del Estado (Spain).

2.2 Tide, bias and datum correction

Several facts complicate the immediate comparison between different sea level forecasts and observations, which necessitate the requirement for the making of some decisions and pre-processing before sea level data enter the MATROOS tool:

- some models provide total sea level (tide + meteorological + density effects) and others just the surge component (meteorological variations);
- reference or datum of sea levels differ between models and data: models refer their output to “mean sea level”, which in this case it is a spatial average that depends on the model domain and the boundary conditions. “Mean sea level” from a tide gauge station is a temporal and local (one point) average, so it depends on the period of data and the station position. To further complicate things, observations of sea level from tide gauges are normally referred to the “harbour” or “chart” datum, i.e. close to the Lowest Astronomical Tide, not to mean sea level;
- when the model includes the tide, this differs from the one obtained from observations as models use just a few set of harmonics, and not all the models use the same set. The most precise tide at a particular harbour comes from harmonic analysis of tide gauge observations.

One of the consequences of this is that all the models present significant bias with respect to sea level observations, both in surge and total sea level (Fig. 5), as well as differences in the tide and reference. In order to minimize the bias problem during the period of the ECOOP project, and facilitate the visualization and comparison within the MATROOS tool, this bias was computed for all the sources and stations based on two months of data previous to the Target Operational Period (TOP) of the ECOOP project which started on January 2009, and then it was applied operationally to the sources before integration into the system. This was obviously not needed for Nivmar as the bias is already corrected in this case by the use of tide gauge data in near-real time.

Differences in the tide, on the other hand, have been solved within MATROOS in the following way:

- for models providing just total sea level: harmonic analysis is performed for one year of model output (for all the grid points (Fig. 6) (and at the tide gauge points; Fig. 7)). From the obtained harmonic constants, the tide (Model Tide) can be computed and the surge component (total – tide) of the forecast extracted from:

Forecasted Surge = Forecasted Total Sea Level – Model Tide.

- A harmonic analysis is also performed for one year of tide gauge observations with the same software (to avoid any differences due to the number and set of constituents used), in order to compute, in the same way:

Observed Surge = Observed Total Sea Level – Observations Tide.

- Finally, the total sea level forecasted by the ENSURF system for a particular source will be the result of the Tide obtained from the observations and the Forecasted surge:

Total Sea Level ENSURF = Forecasted Surge + Observations Tide.

One of the advantages of this need for pre-computing and extracting the tide from the models that provide total sea level is that it has allowed the detection of problems in some of the sources, which after harmonic analysis and tide extraction showed large oscillations on the residuals. Sometimes these problems are related to the wrong introduction of the tide. On the contrary, a normal appearance of the model surge component may be an indicative of the correct performance of the model sea level output (Fig. 7). For this task we have used the Foreman harmonic analysis and prediction software.

3 Bayesian Model Average (BMA) technique

One of the advantages of multi-model systems is that they provide the opportunity to apply multi-model ensemble techniques, such as the Bayesian Model Average (BMA). This method was first employed in social and health sciences (Leamer, 1978), and later applied to dynamical weather forecasting models by Raftery and co-workers (Raftery et al., 2005; Kass and Raftery, 1995; Hoeting et al., 1999). In 2008, the technique was implemented for forecasting sea level at stations along the Dutch coastline, making use of six different forecasts from the NOOS region (Beckers et al., 2008).

As one of the main objectives of ENSURF, we will present later the validation results of several implementations of the BMA for the IBIROOS and Western Mediterranean regions, as compared with the validation of the existing independent sources. The BMA will also provide a probabilistic forecast including confidence intervals.

3.1 Description of the technique

When selecting a particular model for prediction there is always a source of uncertainty that is normally ignored and then underestimated. The BMA method solves this problem by conditioning, not on a single "best" model, but on an ensemble of competing models, becoming a standard method for combining predictive distributions from different sources. Our uncertainty about the best of these sources is quantified by the BMA.

It is important to stress that the dominant approach to probabilistic weather forecasting has been the use of ensembles in which a model is run several times with different initial conditions or model physics (Leith, 1974; Toth and Kalnay, 1993; Molteni et al., 1996; Hamill et al., 2000). In our case, the approach is slightly different as we make use of existing operational systems based on different models and even physics, and of course more limited in the number of members.

The basic idea is to generate an overall forecast probability density function (PDF) by means of a weighted average of PDF's centered on the individual bias-corrected forecasts; the mean of this total PDF is expected to have a smaller root mean square (RMS) error than those of the different models, i.e. there should be an improvement of the performance with respect to those of the individual forecasts (Fig 8). The weights used on this average represent the probability that a particular model will give the correct forecast PDF, and this is determined and updated operationally based on the performance of the models during a recent training period. The technique thus relies on the availability of sea level data from tide gauges in near-real time, as has been mentioned before. Moreover, the overall PDF, being reasonably well-calibrated, can provide a forecast confidence interval which is important for many practical applications. The BMA weights can also be used to assess the skill of ensemble members and for their pre-selection.

The variance of the total PDF is the result of two components: the first one associated with the spread of the ensemble members, the second one with the variance of the individual model forecast PDF's. This latter component should also be determined over a training period, which can be different from the training period mentioned earlier, although in ENSURF the same training period is used to determine the BMA weight and the variance of the individual models.

The computation of the optimal BMA forecast PDF is done by means of the EM algorithm, an iterative algorithm that alternates between two steps, the *E* (or expectation) step and the *M* (or maximization) step:

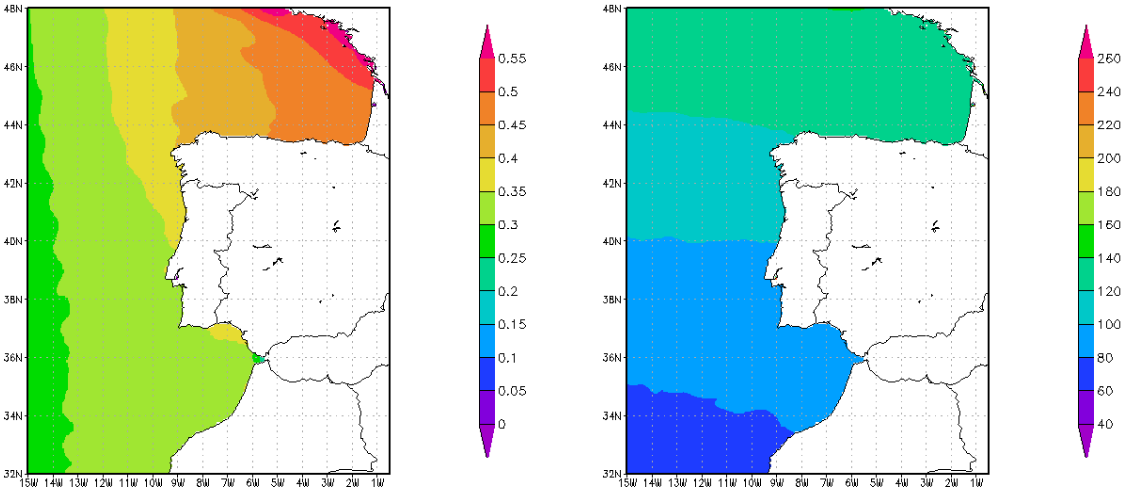


Fig. 6. S_2 harmonic constituent (amplitude, left, and phase, right), result of the harmonic analysis of one year of data at all the grid points of *eseoat* source (ENSURF-IBIROOS).

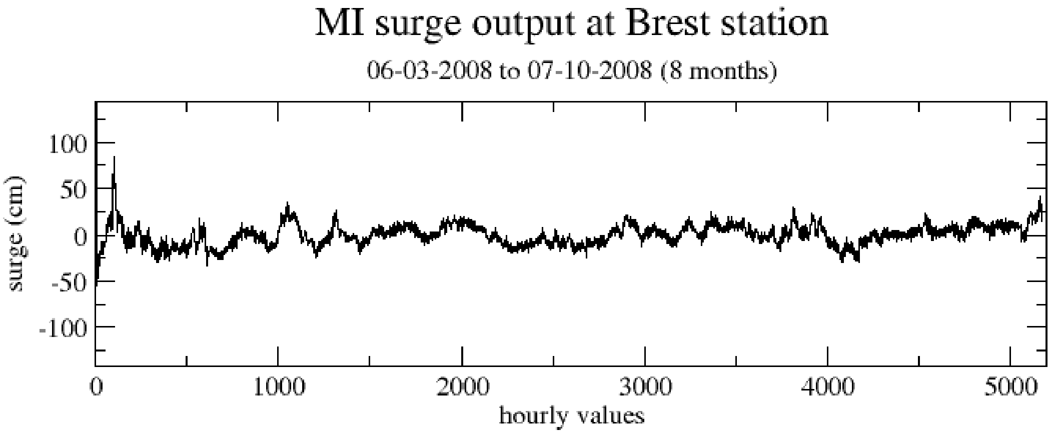


Fig. 7. Surge component of *imi* source (Irish Marine Institute) at Brest tide gauge, after tide extraction. The normal appearance of the surge from the model (no spikes or reference changes or tidal oscillations) usually confirms the correct introduction of the tide in the model. This was not always the case during ENSURF development.

1. the E step starts from an initial guess for the weights $w(k)$ of each individual model and estimation of the matrix **z***(k,s,t)*, which represents the probability that model k gives the best forecast for station s at time t:

$$\mathbf{z}^j(k,s,t) = \frac{w(k)g(k,s,t)^{j-1}}{\sum_i w(i)g(i,s,t)} \quad (1)$$

where j refers to the *jth* iteration of the algorithm, and g represents the probability that the observed value obs*(s,t)* was predicted correctly by model k, i.e. the forecast PDF of each model which is assumed to be a normal distribution with variance $\sigma(k)$:

$$g(k,s,t) = \frac{\sigma(k)}{\sqrt{\pi}} \exp\left(\frac{(\mathrm{obs}(s,t) - \mathrm{forecast}(k,s,t))}{2\sigma(k)^2}\right) \quad (2)$$

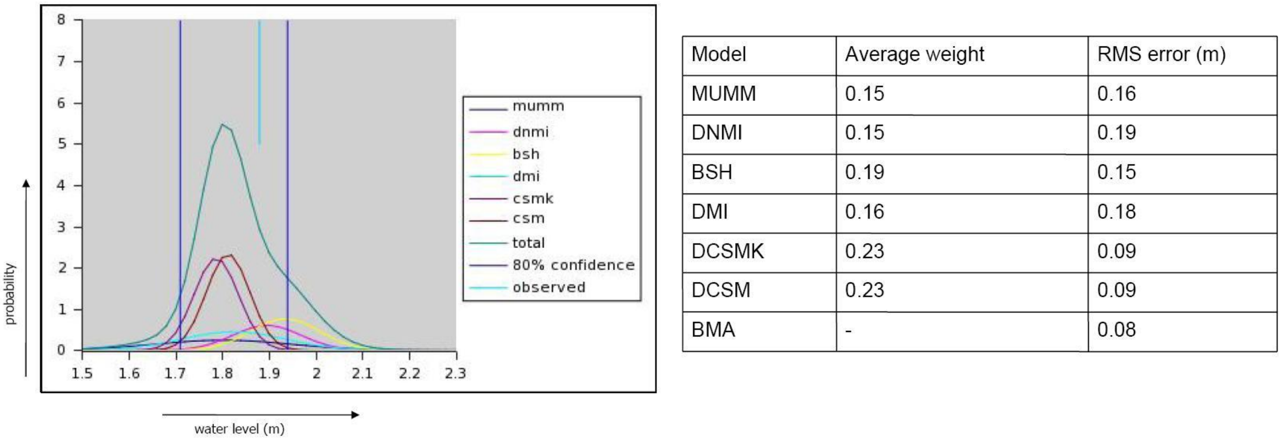
2. the M step consists then on the determination of weights $w(k)$ and variances $\sigma(k)$ of each of the models (k), based on the values of *z(k,s,t):*

$$w^j(k) = \frac{1}{n}\sum_{s,t} z^j(k,s,t) \quad (3)$$

$$\sigma^j(k)^2 = \frac{1}{n}\sum_{s,t}\sum_{k} z^j(k,s,t)(\mathrm{obs}(s,t) - \mathrm{forecast}(k,s,t))^2 \quad (4)$$

where n is the number of observations in the training period.

These two steps are repeated until convergence by using a convergence criterium or by fixing the number of iteration



Model	Average weight	RMS error (m)
MUMM	0.15	0.16
DNMI	0.15	0.19
BSH	0.19	0.15
DMI	0.16	0.18
DCSMK	0.23	0.09
DCSM	0.23	0.09
BMA	-	0.08

Fig. 8. Left: individual and overall pdf's for a single 24 h forecast (27 October 2006, 13:00, Delfzjil), based on 6 models for the North Sea component of ENSURF. The 80 % confidence interval is marked on dark blue, the actual observed values was 1.87 m (blue), within the confidence interval. Right: results for the BMA and the individual forecasts for the period 2003–2006 (6 stations) (extracted from Beckers et al., 2008).

cycles that should guarantee convergence. Beckers (2008) found that 10 iterations are normally sufficient. In the ENSURF implementation, being an operational application, weights and variances from the previous time step are used as a starting point for the new iteration. Once the convergence is reached, the overall forecast mean for each of the stations can be computed from:

$$\text{forecast}(\text{overall}, s, t\_fc) = \sum_k w(k)\text{forecast}(k, s, t\_fc) \quad (5)$$

and the overall forecast confidence intervals can then be obtained by integrating the weighted sum of the individual forecast PDFs.

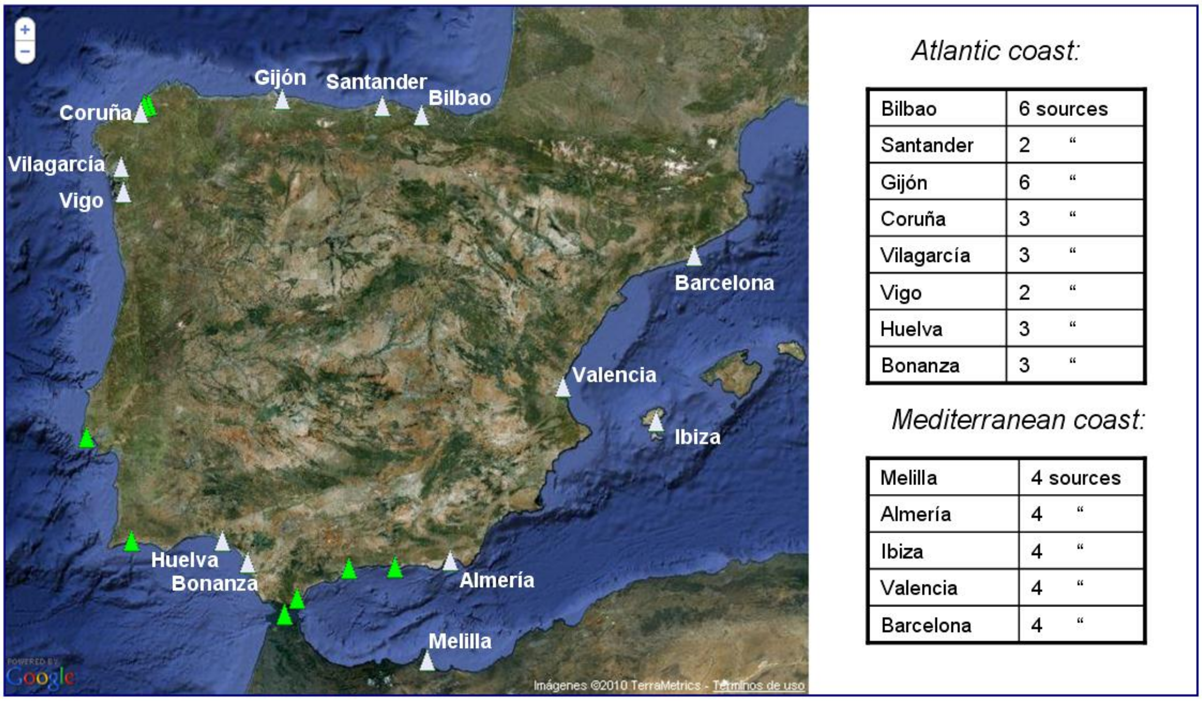
3.2 BMA experiments for ENSURF-IBIROOS

We have implemented several BMA trial versions making use of the flexibility of the MATROOS visualization tool. The final BMA version is applied to the surge component of sea level forecast. This was done because this component can be approximated by a normal distribution to a reasonable degree of accuracy, which is not the case for total sea level including tides, especially for strong semidiurnal tidal regimes. All the validation results at the end of this chapter refer to the surge or meteorological component. The total sea level forecast is nevertheless available operationally in ENSURF as we add the tide computed from tide gauge data to the different forecasts, including the BMA, as explained in the previous section (Fig. 2).

As has been mentioned, the implementation of ENSURF for IBIROOS and the Western Mediterranean represents the first activity of an operational multi-model forecast in the region, so several experiments were performed during the development of the system. In many cases, some of the sources available had not been well validated with respect to sea level, as their initial objective was general ocean forecast including parameters such as currents, temperature and salinity. In these cases, ENSURF has allowed the detection of problems related to the tidal modeling, to the boundary conditions or to the re-initialization scheme. All these problems propagate into the forecasted sea level time series. The forecasts that had a poor Correlation Index with observations were not included in the BMA implementations of ENSURF. Several institutions are still working on the improvement of some aspects of the models that will hopefully provide better forecasts of sea level in the near future. This is the case for MeteoGalicia (Spain) and the Marine Institute (Ireland).

The following initial BMA forecasts were implemented in the region, taking into account the reliable sources available and their common domains (we will distinguish between Atlantic and Western Mediterranean coast):

- Atlantic: available sources: *nivmar, eseoat, imi, metfr* (3 sources) and *metga*. In this case, we have output from two baroclinic sources, *eseoat* and *imi*. Four BMA versions were implemented (TP being the Training Period), avoiding *metga* and *imi* sources, due to the low Correlation Index that will be shown later:
  - BMA0: *eseoat* and *nivmar*, TP = 15 days,
  - BMA_ibi1: *eseoat, nivmar, metfr* (3), TP = 7 days,
  - BMA_ibi2: *eseoat, nivmar, metfr* (3), TP = 4 days,
  - BMA_ibi3: *eseoat, nivmar, metfr* (3), TP = 15 days.
- Mediterranean: available sources: *nivmar* and *metfr* (3). In the Mediterranean the four sources are barotropic. We used all the sources available, 4 in total, for the BMA implementation, changing also the TP, as in the Atlantic coast:



Bilbao	6 sources
Santander	2 “
Gijón	6 “
Coruña	3 “
Vilagarcía	3 “
Vigo	2 “
Huelva	3 “
Bonanza	3 “

Melilla	4 sources
Almería	4 “
Ibiza	4 “
Valencia	4 “
Barcelona	4 “

Fig. 9. Stations for which the BMA was implemented (white icons) for the first ENSURF-IBIROOS implementation, based on the availability of quality control of tide gauge data in near-real time, and more than two sources of forecast.

- BMA_med1: *nivmar, metfr* (3), TP = 7 days,
- BMA_med2: *nivmar, metfr* (3), TP = 4 days,
- BMA_med3: *nivmar, metfr* (3), TP = 15 days.

The BMA was implemented at particular stations or harbours that were selected based on the availability of a sufficient number of sea level forecasts or sources, and automatic near-real time quality control of tide gauge data (Fig. 9). At the present stage of ENSURF implementation, there are still several harbours where only one forecast exists. At the beginning of this project only Puertos del Estado tide gauges and the REDMAR network (Pérez et al., 2008) were using a common quality control procedure. This software is currently being extended to all IBIROOS stations within MyOcean project, and will be applied by Puertos del Estado also for other Mediterranean stations included in ENSURF in the future. All the BMA versions were in operation during the ECOOP Training Operational Period (TOP). Results of the validation of all the sources and the BMA will be shown in the next section.

It should be kept in mind that both observations and forecast data may not be complete, so the BMA must deal with missing data, and a weight $w(k)$ and a $\sigma(k)$ will be determined as long as there is at least one forecast-observation combination in the training period. It may seem that the BMA method has little to offer if there is only one forecast available. However, taking into account the problem of the bias between models and observations previously mentioned, the BMA implementation in ENSURF includes a bias correction, which in many situations still improves on the original forecast.

4 Validation results

Basic statistic parameters (Root Mean Square Error: RMSE, Correlation Index: CI, Maximum Error: RMAX and Mean difference: Bias) were computed from the comparison between the different models and the tide gauge observations, for the period September 2008 to December 2009. The different BMA versions were treated as additional model forecasts. In order to synthesize all the data, we have averaged and plotted the CI and RMSE parameters of all the stations and sources on the Iberian Atlantic coast and on the Mediterranean coast (Figs. 10 and 11, respectively).

4.1 Barotropic vs. baroclinic sources

One of the first objectives of the validation was comparing the output of baroclinic general circulation models operational in the region to the standard storm surge applications based on barotropic, vertically-integrated models. At the time of writing this paper this was only possible for stations on the Atlantic coast.

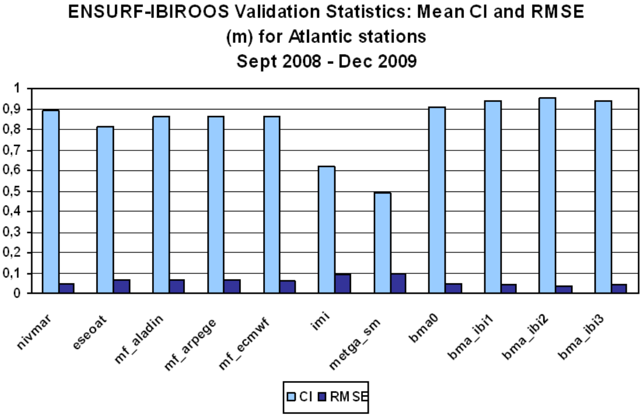


Fig. 10. Mean CI (Correlation Index) and RMSE (Root Mean Square Error) for the sources and stations of the Iberian Atlantic coast (September 2008 to December 2009).

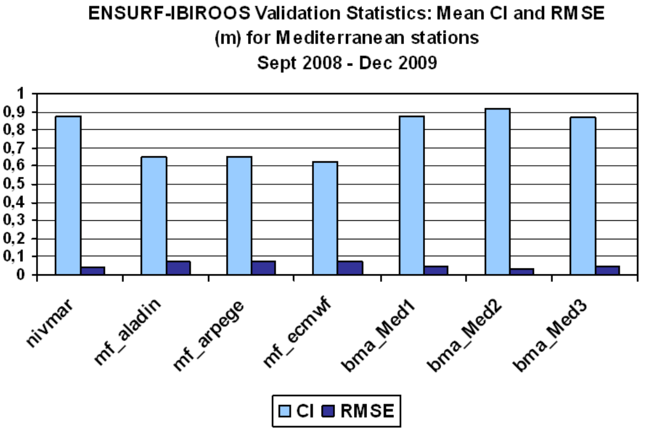


Fig. 11. Mean CI (Correlation Index) and RMSE (Root Mean Square Error) for the sources and stations of the Mediterranean coast (September 2008 to December 2009).

Figure 12 shows the RMSE and CI of several sources for Bilbao station (North of Spain), with respect to tide gauge observations. Best performance is found for *nivmar*, which was to be expected, since it is the only source that automatically and dynamically corrects the bias based on the observations. Interestingly, the Météo-France forecasts (*mf-aladin*, *mf-arpege*, and *mf-ecmwf*), which are also barotropic but without tide gauge data assimilation, give better statistical results than the baroclinic forecasts from *eseoat* and *imi*. This is an important although not new conclusion about the capability of baroclinic models to correctly reproduce sea level variations. Averaged statistical parameters for all the Atlantic stations (Fig. 10) confirm this point. Figure 10 also reveals some problems with sources *imi* and, especially, *metga_sm* which show poorer statistics. This is the reason why they were not used for the BMA forecast. Institutions responsible for these systems are investigating the causes. On the other hand, *eseoat* shows a relatively good performance, taking into account that it does not make use of tide gauge data, although not enough to improve on the results of the barotropic sources.

4.2 General performance of the BMA's

Next, the performance of the BMA implementations was compared to that of the individual forecasts. From Fig. 10 it can be seen that the BMA performance is, in general, outperforming the individual models for the Atlantic coast, with higher CI and lower RMSE. This is true for practically all of the BMA's, but more clearly for *bma_ibi2*, having a 4 day training period.

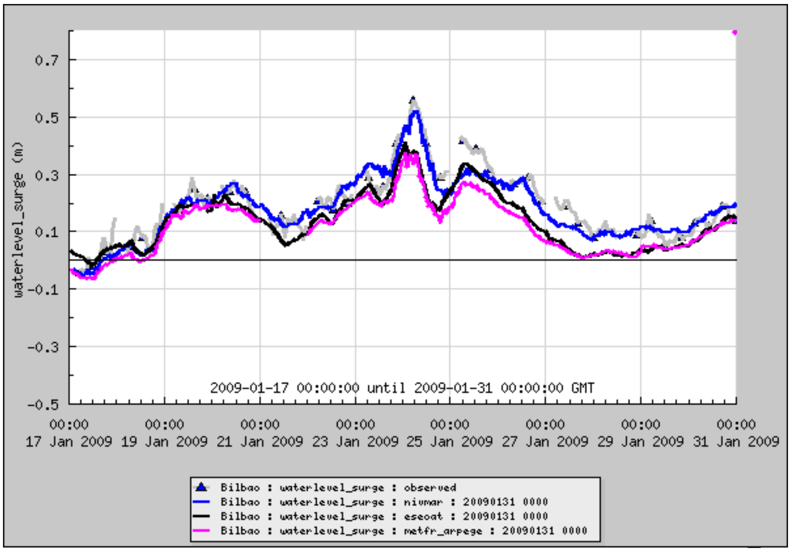
Results for the Mediterranean are presented in Fig. 11. In this case, as already mentioned, there were no baroclinic sources available in ENSURF at the time of writing this paper and the BMA versions do not improve the results of *nivmar* so clearly, with *bma_med2* using 4 days of training period the only one showing a slight improvement in terms of CI and RMSE. The forecasts from Météo-France are poorer than in the Atlantic, possibly due to the presence of a boundary of the Météo-France model domain around Sardinia and Corsica, which disturbs the results at these stations. Taking into account the experience with the *nivmar* system, we recommend the Mediterranean Sea to be completely covered by the model domain.

In conclusion, all the BMA versions produce good results in the Atlantic, improving the performance of the best of the individual sources, but this is not always the case in the Mediterranean. It is important to notice that the performance indicators that were used were mean statistic parameters and that results can differ slightly depending on the station and the period of data.

4.3 Influence of the data period on the validation results

In order to determine the influence of the data period used for the validation, and taking into account the existence of months with very low storm activity in the initial period September 2008 to December 2009, we repeated the performance assessment for other periods. We selected for the Atlantic stations the stormy season of January to February 2009, where most of the largest surge events since the implementation of ENSURF were present (see example of results for the two periods at Gijón station, Table 2).

The first result is that all the sources show an improvement in terms of their performance indicators, especially for those performing poor during the previous period, such as *imi* or *metga_sm*. One possible explanation for this could be the fact that this test period is close to the period that was used for bias correction of the models; in fact most of the sources present a drift over a period of months with respect to observations. Some institutions are investigating the reason for this drift; an interesting point is that, although easy



Model	RMSE	CI
Nivmar	0.041	0.96
Eseoat	0.062	0.91
Mf-aladin	0.049	0.95
Mf-arpege	0.049	0.95
Mf-ecmwf	0.050	0.95
Imi	0.054	0.95

Fig. 12. Comparison of baroclinic models (*eseoat* and *imi*) and barotropic ones (*nivmar* and *metfr* models) for the stormy period January–February 2009 at Bilbao station. Blue colour in the table (right) used for the sources with better statistical parameters. (Data in meters, RMSE: root mean square error, C.I.: Correlation Index).

Table 2. Statistical parameters of the validation for the different sources at Gijón station (North Spain), for the whole period and for just the stormy months of January and February 2009. RMSE: Root Mean Square Errors (in meters), CI: Correlation Index.

Gijón station	Sep 2008–Dec 2009		Jan 2009–Feb 2009	
Source	RMSE (m)	CI	RMSE (m)	CI
nivmar	0.042	0.94	0.042	0.96
eseoat	0.055	0.90	0.047	0.95
mf-aladin	0.072	0.83	0.047	0.96
mf-arpege	0.072	0.83	0.046	0.96
mf-ecmwf	0.071	0.83	0.045	0.96
metga_sm	0.080	0.74	0.045	0.96
imi	0.102	0.58	0.052	0.97
BMA	0.036	0.96	0.038	0.97

to understand that barotropic models do not include all low frequency variations of sea level, especially those related to steric effects, these should be present in baroclinic sources such as *eseoat* and *imi* that should include any kind of forcing for sea level variability. This may indicate that it is necessary to implement an operational bias correction based on observations, as *nivmar* does, for the other sea level forecasts.

Another important point is that the BMA versions performed best in practically all the stations in the Atlantic for the initial period September 2008 to December 2009, but only get this improvement for 50 % of the stations (Gijón, Bonanza, Huelva and Vigo), for the period of January–February 2009. For the rest of stations (Bilbao, Santander, Coruña and Vilagarcía), *nivmar* gives the best results.

Finally, we repeated the validation analysis for the period mid-November 2009 to January 2010, also a stormy season, especially in the Mediterranean, and this time at the end of the initial period and far from the bias correction months. This time we get the opposite situation in the Mediterranean: now the BMA performs better than the individual models, also improving the best source *nivmar* in all the stations, except Melilla. For the Atlantic, all the stations show an improvement by the BMA except for Coruña and Vigo.

In conclusion, the BMA gives better performance improvement when using the whole period of data of one year or more, or the last months of this period, farther away from the period that was used for the individual models' bias correction. This would suggest that the improved performance by the BMA is for a large part due to its integrated bias correction. It is important to take into account that poor performance could reflect other problems such as gaps or anomalies in the sources or the observational data, which depend upon the period. Malfunction of the tide gauges is difficult to avoid in the operational mode, even with near-real time quality control procedures that will be unable to deal with all kinds of errors.

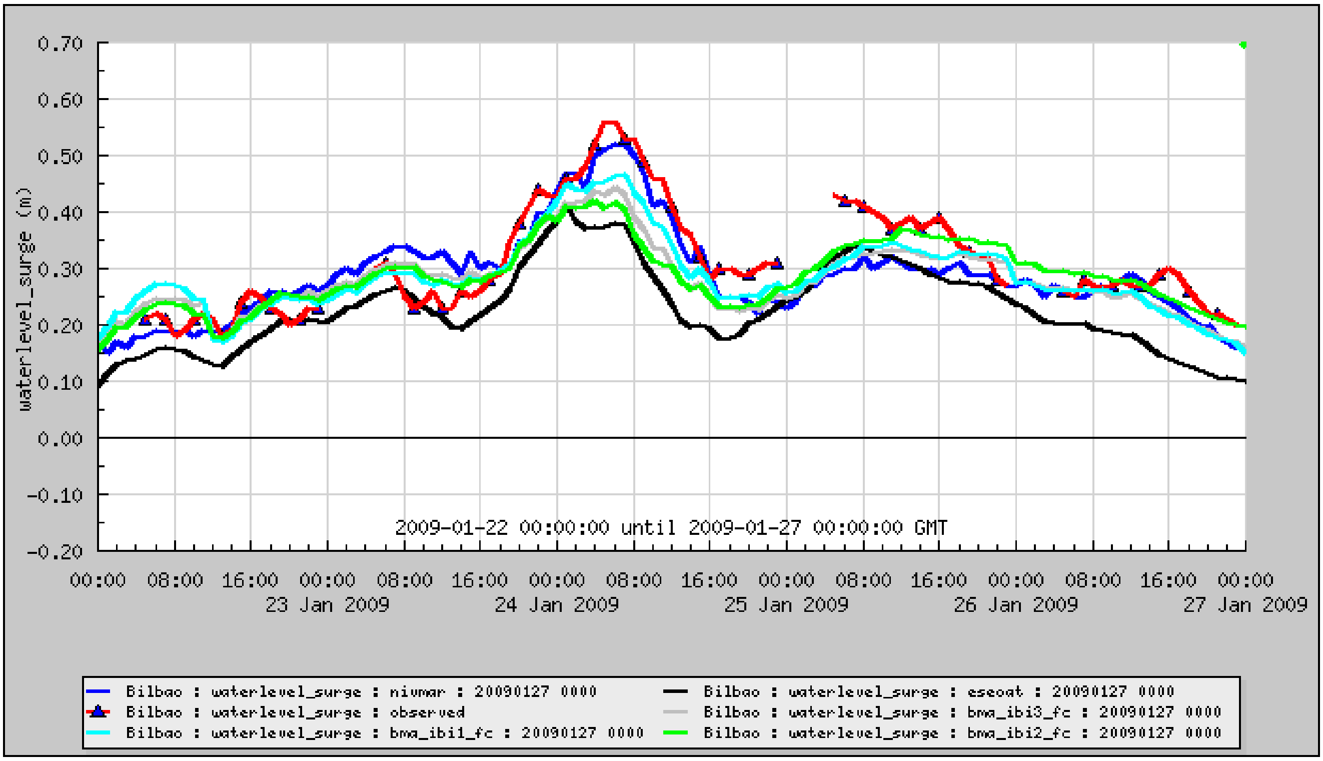


Fig. 13. Comparison of different forecasts during the peak of a storm for Bilbao station, in January 2009. Red line corresponds to observed surge data (tide gauge). The three BMA versions reproduce the peak worse than *nivmar* (blue) and better than *eseoat* (black).

4.4 Performance during the peak of a storm

The performance assessment results presented up to now were based on average statistical parameters, obtained from a relative long period of data. They reflect the general behaviour of the models and the BMA for all meteorological conditions. However, the objective of a good sea level forecast should be an adequate simulation of the peak of a storm. Figure 13 shows again the forecasts for the largest storm of the period of study, at Bilbao station. As we pretend to check the improvement of the BMA, for the sake of simplicity we show now just the output from *nivmar* and *eseoat* sources, as well as the output of the three different BMA's with different training periods.

It can be seen that the peak of the storm is better reproduced by *nivmar* source, and that the BMA do not improve the forecast, although they do better than *eseoat* in any case. This is an important result that should be explored in detail. We also see at this particular harbour that the BMA that reproduces the peak best is *bma_ibi1*, with 15 days of training period. For the whole period for Bilbao, we see however that a training period of 4 days (*bma_ibi2*) performs better. Although this can be different at another harbour, this finding is contradictory to what is expected and to results for the North Sea (Beckers et al., 2008), where peaks were reproduced better with shorter training periods.

The reasonably good forecast of *nivmar* at the peak of the storm is also remarkable. One possible explanation for this is the adequate representation of the continental platform, very narrow here, for the *nivmar* system: it was manually and carefully corrected before final implementation. This is something that was not done in the rest of the sources, which are less focused on sea level, such as *eseoat*.

It seems, in any case, that a better determination of the BMA weights and parameters may be needed for an adequate forecast of extreme events. Beckers et al. (2008) already suggested this idea and propose to determine these weights based on the performance of the models during extreme meteorological conditions instead of during a recent training period.

5 Conclusions and future work

ENSURF has proved its utility as a validation and multi-model forecast tool, and has facilitated the first experience of the exchange of operational forecasts for the IBIROOS and Western Mediterranean regions. It has allowed the detection of problems regarding calibration and bias correction of existing operational models that had not been noticed before. For the first time a probabilistic forecast is feasible, based on existing operational systems and the BMA method. The system is implemented and operational for sea level in the NOOS and IBIROOS regions. Extension to other parameters and regions is, in principle, possible.

First validation results of the surge component, based on the comparison between tide gauge data and the forecasts at the harbours confirm that, at least for the IBIROOS region, baroclinic models do not reach the performance of barotropic models for storm surge applications. There is a general improvement of performance by the BMA forecast compared to the individual models. This improvement is most clear for the Atlantic stations and becomes less evident if we change the data period and concentrate on the stormy season. The BMA has some difficulty in reproducing the peak of the storm, as compared to the *nivmar* source.

Future work will focus on the addition of new sources and extension to the whole Mediterranean, the more relevant and urgent ones being the operational forecasting systems established within the MyOcean project for IBIROOS and MOON. Availability of near-real time data from tide gauges, with automatic quality control to avoid erroneous data entering the system, is a prerequisite for accurate forecasts of sea level at the harbours and for the functioning of the BMA technique. Within the same project, the automatic quality control of tide gauge data will therefore be applied to the rest of sea level stations in Europe, which will allow the completion of the BMA implementation and validation for other countries contributing to ENSURF. Finally, a detailed study of the influence of the training period in the BMA performance or the extension to 2-D fields should be the goal in the near future.

Acknowledgements. This work has been supported through the EU ECOOP project, Contract No. 36355. We would like to thank the personnel working at the different institutions who have collaborated in providing the sources to ENSURF. We thank also Ronan Crèach (from SHOM, France), Vibeke Huess (from DMI, Denmark), Elizabeth Bradshaw (from POL, UK) and Gonzalo Crisostomo (from IGN, Portugal), for their support and cooperation for making the tide gauge data available.

Edited by: S. Cailleau

References

Alvarez F. E., Pérez-Gómez, B., and Rodríguez, I.: A description of tides in the Eastern North Atlantic, Prog. Oceanogr., 40, 217–244, 1997.

Alvarez F. E., Pérez, B., and Sánchez-Arévalo, I. R.: Nivmar: A storm surge forecasting system for the Spanish Waters, Sci. Mar., 65, 145–154, 2001.

Chézy, A.: Formule pour trouver la vitesse de l'eau conduit dan une rigole donnée, Dossier 847 (MS 1915) of the manuscript collection of the École National des Ponts et Chaussées, Paris, 1776, Reproduced, in: Histoire d'une formule d'hydraulique, edited by: Mouret, G. and Chézy, Antoine, Annales des Ponts et Chaussées, 61, Paris, 1921.

Mourre, B., De Mey, P., Mènard, Y., Lyard, F., and Le Provost, C.: Relative performance of future altimeter systems and tide gauges in constraining a model of North-Sea high-frequency barotropic dynamics, Ocean Dynam., 56, 473–486, 2006.

Backhaus, J. O.: A semi-implicit scheme for the shallow water equations for application to shelf sea modeling, Cont. Shelf Res., 2, 243–254, 1983.

Beckers J. V. L., Sprokkereef, E., and Roscoe, K. L.: Use of Bayesian model averaging to determine uncertainties in river discharge and water level forecasts, Proc. 4th International Symposium on Flood Defence: Managing Flood Risk, Reliability and Vulnerability, Toronto, Ontario, Canada, 6–8 May, 2008.

Charnok, H.: Wind stress on a water surface, Q. J. Roy. Meteor. Soc., 81, 639–640, 1955.

Egbert, G. D. and Erofeeva, S. Y.: Efficient inverse modeling of barotropic ocean tides. Journal of Atmospheric and Oceanic Technology, 19, 183–204, 2002.

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., and Young, G. S.: Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere Coupled-Ocean Atmosphere Response Experiment, J. Geophys. Res., 101, 3747–3764, 1996.

Flather, R. A.: Practical surge prediction using numerical models, in: Floods Due to High Winds and Tides, edited by: Peregrine, D. H., London, Academic Press, 109 pp., 1981.

Flather, R. A.: Estimates of extreme conditions of tide and surge using a numerical model of the north-west European continental shelf, Est. Coast. Shelf Sci., 24, 69–93, 1987.

Foreman, F. G. G.: Manual for tidal heights analysis and predictions, Institute of Ocean Sciences, Patricia Bay, Pacific Marine Report no. 77, 101 pp., 1977.

GETECH: Geophysical Exploration TECHnology internal report, Dept. Earth Sciences, University of Leeds, UK, 1995.

Hamill, T. M., Snyder, C., and Morss, R. E.: A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles, Mon. Weather Rev., 128, 1835–1851, 2000.

Hoeting, J. A., Madigan, D. M., Raftery, A. E., and Volinsky, C. T.: Bayesian model averaging: A tutorial (with discussion), Stat. Sci., 14, 382–401, [A corrected version is available on-line at: www.stat.washington.edu/www/research/online/hoeting1999.pdf.], 1999.

Kass, R. E. and Raftery, A. E.: Bayes factors, J. Amer. Stat. Assoc., 90, 773–795, 1995.

Large, W. G. and Pond, S.: Open ocean momentum flux measurements in moderate to strong winds, J. Phys. Oceanogr., 11, 324–336, 1981.

Leamer, E. E.: Specification searches, Wiley, NY, 370 pp., 1978.

Leith, C. E.: Theoretical skill of Monte-Carlo forecasts, Mon. Weather Rev., 102, 409–418, 1974.

Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T.: The ECMWF ensemble system: Methodology and validation, Q. J. Roy. Meteor. Soc., 122, 73–119, 1996.

Pérez, B., Vela, J., and Alvarez-Fanjul, E.: A new concept of multi-purpose sea level station: example of implementation in the REDMAR network, in: Proceedings of the Fifth International Conference on EuroGOOS, May 2008: Coastal to global operational oceanography: achievements and challenges, Exeter, UK, 2008.

Poole, D. and Raftery, A. E.: Inference for Deterministic Simulation Models: The Bayesian Melding Approach, J. Am. Stat. Assoc., 95, 1244–1255, 2000.

Raftery A. E., Gneiting, T., Balabdaoui, F., and Polakowsky, M.: Using Bayesian Model Averaging to Calibrate Forecast Ensembles, Mon. Weather Rev., 133, 1155–1174, 2005.

Rodríguez, I., Alvarez, E., Krohn, J., and Backhaus, J.: A mid-scale tidal analysis of waters around the north-western corner of the Iberian Peninsula, Proceedings Computer Modelling Ocean Eng., 91, 568 pp., Balkema, 1991.

Shchepetkin, A. F. and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Model., 9, 347–404, 2005.

Sotillo, M. G., Jordi, A., Ferrer, M. I., Conde, J., Tintoré, J., Alvarez-Fanjul, E.: The ESEOO regional ocean forecasting system, in: Proceedings of the ISOPE-2007, The 17th International Offshore Ocean and Polar Engineering Conference, Lisbon, Portugal, 2007.

Sotillo, M. G., Alvarez-Fanjul, E., Castanedo, S., Abascal, A. J., Menéndez, J., Emelianov, M., Olivella, R., García-Ladona, E., Ruiz-Villarreal, M., Conde, J., Gómez, M., Conde, P., Gutiérrez, A. D., and Medina, R.: Towards an operational system for oil-spill forecast over Spanish waters: Initial developments and implementation test, Mar. Pollut. Bull., 56, 686–703, 2008.

Toth, Z. and Kalnay, E.: Ensemble forecasting at the NMC: The generation of perturbations, B. Am. Meteorol. Soc., 74, 2317–2330, 1993.

Undén, P., Rontu, L., Järvinen, H. Lynch, P., Calvo, J., Cats, G.,Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., Jones, C., Lenderlink, G., McDonald, A., McGrath, R., Navascues, B., Woetman Nielsen, N., Odegaard, V., Rodríguez, E., Rummukainen, M., Rõõm, R., Sattler, K., Hansen Sass, B., Savijärvi, H., Schreur, B. W., Sigg, R., The, H., and Tijm, A.: HIRLAM-5 Scientific Documentation, HIRLAM-5 Project, c/o Per Undën, SMHI, SE-601 76, Norrkping, Sweden, 2002.

Wu, J.: Wind-stress coefficients over sea surface from breeze to hurricane, J. Geophys. Res., 87, 9704–9706, 1982.