# Very Large ensemble ocean forecasting experiment using the Grid computing infrastructure

Nadia Pinardi[1], Alessandro Bonazzi[2], Enrico Scoccimarro[2], Srdjan Dobricic[3], Antonio Navarra[3,2], Antonia Ghiselli[4], Paolo Veronesi[4]

[1]*Università di Bologna, Centro Interdipartimentale per le Scienze Ambientali, Ravenna, Italy,*

[2]*Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy,*

[3]*Centro EuroMediterraneo per i Cambiamenti Climatici, Bologna, Italy*

[4]*Istituto Nazionale di Fisica Nucleare, Bologna, Italy*

Submitted to

*Bulletin of American Meteorological Society*

Revised version

October 12, 2007

*Corresponding author address*:

N.Pinardi, Corso di Scienze Ambientali, Via S.Alberto 163, 49100 Ravenna Italy, e-mail: n.pinardi@sincem.unibo.it

## CAPSULE SUMMARY

Ensemble ocean forecasts of exceptional size are possible using a Grid computing infrastructure and within the limitation of operational time constraints.

Atmospheric and oceanic ensemble forecasting is a way to deal with uncertainty related to inaccurate knowledge of the initial state of the atmosphere and the ocean, the lateral and vertical boundary condition errors and the model physics shortfalls (Lewis, 2005, Epstein, 1969). Since the atmosphere and the ocean are extremely non-linear systems (Lorenz, 1993, Saravanan et al., 2000) initial uncertainties can amplify and limit the predictability of short term forecasts (Kleeman and Majda, 2005).

For the ocean, ensemble forecasting is a novel field. Ensemble methods are used to compute the background error covariance matrix in data assimilation schemes (Evensen, 2003) but are not used yet to quantify the forecast uncertainty in short term ocean forecasting systems. Initial conditions uncertainty is a major source of unpredictability for ocean currents due to the limited observations available for nowcasting and the highly non-linear physics. In this study we explore the short term ensemble forecast variance generated by perturbing the initial conditions using a new computational facility, so-called Grid infrastructure (http://grid.infn.it/), distributed over the Italian territory. This infrastructure allowed us to perform several ensemble forecast experiments with 1000 members: they are completed within 5 hours of wall-clock time after their submission and the ensemble variance peaks at the mesoscales.

## ENSEMBLE FORECASTING

Ensemble forecast methods are well established in meteorology but much less in oceanography. Normally initial conditions are perturbed and several forecasts are run forming an ensemble of predictions used to study and calculate the probability distribution of the forecasts. The different initial conditions are produced by perturbation techniques, some of them very sophisticated (Cai et al., 2003). Using this methodology, the uncertainty in the predicted events of the forecast, typically jet stream intensification, synoptic events moving across the domain of interest, are quantified in terms of the ensemble variance. When the ensemble variance is high, the uncertainty in the prediction is also high, indicating a sensitivity of the system to amplify small initial perturbations.

A variety of practical implementations of ensemble techniques for weather forecasts have been proposed (Toth and Kalnay, 1993, Molteni et al., 1996, Houtekamer et al., 1996) and ensemble systems are now used operationally in weather forecasting centers. They play a crucial role in providing probabilistic information on the forecast variables of interest, especially for difficult but important state variables such as precipitation, and they have a large potential for applications (Zhu et al., 2002). The limitation in computing resources has limited the size of the ensembles considered, preventing a full exploration of the convergence properties of the ensemble and limiting the usage of the ensemble technique much outside dedicated computing centers (Buizza and Palmer, 1998).

In the ocean very little work has been done up to now to show where the unpredictability peaks. However, since it is believed that nonlinearities play a major role even in ocean short term forecasts, ensemble techniques should be used to quantify the uncertainty. One of the major drawbacks of present day ocean ensemble forecasting systems is the limitation on the number of ensemble members feasible, slowing down the understanding of the predictability limits of short term ocean forecast and its applications.

**METHODS OF STUDY**

Here we apply the methodology of ensemble forecasting to a weekly ocean forecasting system (Pinardi et al., 2003) for the entire Mediterranean Sea that uses numerical weather prediction atmospheric forcing to drive an ocean general circulation model producing 10 days forecast of three dimensional currents. To control uncertainties in the initial condition, satellite and in situ ocean observations are assimilated to produce a so-called analysis which is then used as initial condition for the ten days forecast (Dobricic et al., 2005). The error in the forecast is normally assessed by taking differences between the forecast and the observations and it is found that the error triples in the surface ocean during the 10 days of the forecast (Demirov et al., 2003).

The forecasting model covers the whole Mediterranean Sea with a constant resolution of 1/8 x 1/8 degrees in horizontal and 31 unevenly distributed levels in vertical, so that the model is only eddy permitting not resolving. The model uses the primitive equations for the ocean and eight state variables are forecasted (temperature, salinity, density, pressure, three velocity components and sea level). The size of the problem in terms of model grid points and state variables is somewhat like $10^7$ which is equivalent to a global ocean circulation model of approximately 1 x 1.5 degrees horizontal resolution and the same vertical resolution. The initial condition of the forecast is produced by melding the model prediction with observations available during the days preceding the start of the forecast. However observations are not frequent enough to completely correct the initial conditions so uncertainty in the initial fields is still high. The atmospheric forcing is taken from the numerical weather prediction system of the European Center for Medium range Weather Forecast (ECMWF) at a horizontal resolution of 0.5 degrees. We call this a deterministic forecast system since only one initial condition and the deterministic atmospheric forecast from ECMWF is used. For this first experiment we selected a model that we would be able to fit into each cpu without resorting to parallelization of the code itself. This limitation can of course be overcome in the future.

The ensemble forecast experiment is produced by perturbing a single initial condition and producing 1000 different initial fields of temperature and salinity. The initial condition perturbation procedure is a simple one, where temperature and salinity are perturbed pseudo-randomly. The perturbed $T_p$ and $S_p$ fields are written as:

$$\begin{aligned} T_p(x,y,z) &= T_0(x,y,z) + p(x,y)\sum_{i=1}^{N} e_i f_i(z) \\ S_p(x,y,z) &= S_0(x,y,z) + p(x,y)\sum_{i=1}^{N} e_i g_i(z) \end{aligned} \qquad (1)$$

where $T_o$ and $S_o$ are the unperturbed initial conditions, $p(x,y)$ is a pseudo-random field, $f_i$ , $g_i$ are 20 vertical empirical orthogonal functions computed from model statistics and $e_i$ their eigenvalues (Demirov et al., 2003). In Fig. 1 we show the initial vertical structure of the perturbation in the temperature field. The perturbation is concentrated at the upper thermocline levels which is the site of maximum temperature variance due to seasonal water mass formation and mixing processes. The largest uncertainty in fact should be connected to misplacement of the temperature and salinity gradients in vertical. The pseudo-random field $p(x,y)$ is generated following the procedure indicated by Evensen (2003). The mean of $p$ is zero and the covariance is specified a priori in order to control the initial field horizontal smoothness. Even if the perturbation (1) is not constrained to be gravitationally stable, the final perturbation is upper thermocline intensified (not shown).

## THE ITALIAN GRID INFRASTRUCTURE

The members of the ensemble forecast experiments are run on a new distributed computing network, so-called Grid (Foster and Kesselman, 1999). The innovative aspects are related to designing and implementing the Grid Production Framework, distinguished from conventional distributed computing by its focus on large-scale resource sharing, innovative applications, and high-performance orientation. The users of a Grid infrastructure are divided into Virtual Organisations (VO), abstract entities grouping users, institutions and resources in the same administrative domain. The Italian Grid infrastructure consists of about 30 sites (Fig. 2) equipped with Computing Elements (CE) formed by 10 up to hundreds of nodes (Worker Nodes-WN), and disk-based Storage Elements (SE) from hundred of gigabytes up to hundred of terabytes (Fig. 3). Each site contains a farm composed of CEs, WNs and SEs which are dedicated resources to the Grid infrastructure, so they are used only by the VO users. The CEs are the entry points of queues managed by Local Resource Management Systems (LRMS) while the jobs are submitted to the CEs by the Resource Broker (RB) which is in direct contact with the user through a User Interface (UI). The ensemble experiments described in this study are run on a maximum number of 15 different sites.

The time needed to set up the forecast experiments and run them was about one man/year for a Ph.D. level researcher, with the consultancy of Grid experts.

## ENSEMBLE FORECAST EXPERIMENTS

A total of 67 ensemble forecast experiments have been carried out at different hours and week days over a period of 20 days in order to test the Grid efficiency through its normal workload cycle. No special arrangement has been made to the Grid configuration and operation policies for this experiment. Each ensemble forecast experiment is designed to launch 1000 jobs within a total time of five hours, after which the jobs are deleted without paying any attention to their status. In summary the ensemble forecast experiments were done following a 3 phases procedure (Fig.3):

Phase 1: Replication. The input files and executable code are uploaded to the INFN-CNAF CE in Bologna (Fig. 2). The input files are replicated over all the available CE of

15 sites.  This procedure takes about 10 minutes, and only after its successful completion will the job execution phase start.

Phase 2: Execution. All the jobs are submitted by the INFN-CNAF (Fig. 2) RB that looks for the best available CE to execute the jobs. To do so, it interrogates an Information Service to query the status of computational and storage resources and the File Catalogue to find the location of the required data. At this point, the LRMS handles a quasi-parallel submission of 1000 jobs on the WNs. Only the CEs belonging to the same farm of the SE, where the data are stored,  are used.

Phase 3: Downloading. If a job is finished successfully, a procedure for the downloading of the model output files is activated. After five hours, all the jobs are cancelled..

The wall-clock-time for the 67 ensemble forecast experiments is shown in Fig.4: the results indicate that al least 200 jobs are accomplished in 2 hours and at least 450 in five hours. If we look at the dispersion around the mean, we realize that  it is very common that 300 jobs will be accomplished in 2 hours. The requirements imposed to the CE were that at least one CPU is free and that the job queue has a job time limit set to a minimum of 80 minutes. With this simple requirement policy we obtained the largest usage of Grid CE and a reasonable efficiency. The 67 ensemble experiments were launched on the Grid at different day time hours giving a daily distribution of the work load that is rather uniform between the 15 sites used.  During each forecast experiment only about 20% of the Grid computing resources were used. About 60% of the jobs reached above 900 members in the five hours due to different problems in the WN availability.

**SUMMARY OF RESULTS**

One important aspect of the ensemble forecast experiment is the ensemble spread which we take to be represented by the standard deviation around the ensemble mean. In the analysis below we concentrate on the Sea Surface Height (SSH) field only: this is an interesting field to study because it is practically equivalent to the integral of the density from surface to bottom, thus giving an overall measure of the growth of the initial perturbation. The results are shown for one particular run of the 67 produced, the one for the forecast from November 16 to 25, 2005 which consisted of exactly 1000 members.

The first day perturbation amplitude, in terms of standard deviation amplitude is of the order of few tens of centimeters and it is shown in Fig. 5a. The perturbation is randomly but uniformly distributed in the deep parts of the basin (yellow-red areas). It is important to notice that this ensemble standard deviation (done at the equivalent of the analysis time or initial condition time) is ten times smaller than the analysis error standard deviation that is estimated to be around 5 cm for a three years period (Dobricic et al., 2005). At the tenth day of the forecast the SSH standard deviation (Fig. 5b) is limited in extent and it is concentrated along strong current jets, eddy borders and frontal structures. While the initial perturbation in SSH is at rather large scales and it has no precise connection with the dynamical structures of the circulation (Pinardi and Masetti, 2000), the final ensemble standard deviation has large amplitude in relatively small areas. The final amplitude is ten times the initial one in those areas, it is about 10% of the ensemble mean signal which is of the order of 10-30 cm (not shown) and is comparable to the analysis error standard deviation already mentioned. The regions with high standard deviation values in Fig. 5b are representative of areas where the errors in the initial condition grow largest, indicating a limited predictability of the flow field or a large uncertainty of the forecast.

The amplitude of the SSH standard deviation did not sensibly change using 300 or 500 members, indicating a saturation of the ensemble variance or spread. Such a number could clearly depend from the kind of perturbation technique used, from the modeling system and from the specific geographical area of our study. However, recent work on storm surge ensemble forecasting in the Gulf of Gascogne (Lamouroux et al., 2006) indicates again that a few hundred members would saturate the growth of the ensemble spread.

In order to detail the perturbation growth, we show in Fig. 6 the 10 days growth of the ensemble standard deviation in areas of the basin having different standard deviations on the last day of the forecast. We can see that the growth in the large standard deviation areas is almost linear and it has the fastest growth rate. An adjustment phase occur in the first two days. Within the first forecast day, the perturbation grows very fast, probably due to geostrophic adjustment. In the second day, where the dynamics is more advective, there is a stabilization after which the perturbation starts a linear growth.

## DISCUSSION

In this study we have shown the largest ocean forecast ensemble experiments done hitherto with an operational forecasting system. The forecast are carried out for the whole Mediterranean Sea with a primitive equation, eddy permitting, state-of-the-art model. We have shown that the variance of the ensemble saturates at about 300 members and that the ensemble standard deviation is concentrated at the mesoscales. Given a pseudo-random horizontal, large scale but thermocline intensified initial perturbation, the ensemble forecast standard deviation in 10 days will concentrate at smaller spatial scales, near frontal structures and eddies. If a sufficient number of members can be produced, this method could be useful to estimate the possible areas and periods of larger prediction errors in the ocean.

Last but not least, we have demonstrated that this large ensemble experiment is realizable under the best effort Grid computing conditions. Approximately 500 members can be executed within ocean predictions operational wall clock time, i.e. within five hours after the ensemble forecast experiment is submitted. This result promises to be valid for any very large computing application that requires jobs to run quite independently from each other, using the resource sharing protocols now developed for Grid infrastructures.

*References*

Buizza, R. and T. N. Palmer. 1998: Impact of Ensemble Size on Ensemble Prediction. *Monthly Weather Review*, **126**, 2503–2518.

Cai M., E. Kalnay and Z. Toth. 2003: Bred Vectors of the Zebiak–Cane Model and Their Potential Application to ENSO Predictions. *Journal of Climate*: Vol. 16, No. 1, pp. 40–56.

Demirov, E. N. Pinardi, C. Fratianni, M. Tonani, L. Giacomelli and P. De Mey, 2003: Assimilation scheme of the Mediterranean Forecasting System: operational implementation" *Annales Geophysicae*, 21: 189-204.

Dobricic, S. N. Pinardi, M. Adani, A. Bonazzi, C. Fratianni and M. Tonani, 2005: "Mediterranean Forecasting System: a new assimilation scheme for Sea Level Anomaly and its validation". *Quarterly Journal of the Royal Meteorological Society*, 131, pp. 3627-3642.

Epstein, E. S., 1969: Stochastic dynamic predictions. Tellus, 21, 739–759.

Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53, 343–367

Foster, I. and Kesselman, C. (eds.), 2004: The Grid: Blueprint for a New Computing Infrastructure, 2nd Edition, Morgan Kaufmann. ISBN: 1-55860-933-4.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.

Kleeman, R., and A. Majda. 2005: Predictability in a Model of Geophysical Turbulence. *Journal of the Atmospheric Sciences*:, **62**, 2864–2879.

Lamouroux, J., P. De Mey, F.Lyard and E. Jeansou, 2006: Control of a barotropic model of the Bay of Biscay in presence of atmospheric forcing errors. *J. Geophys. Res*. In press

Lewis, J.M,  2005: Roots of Ensemble Forecasting**.** *Monthly Weather Review*: **133**, 1865–1885.

Lorenz, E., 1993: *The Essence of Chaos*. University of Washington Press, 319 pp.

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

Pinardi, N. and E. Masetti, 2000: Variability of the large-scale general circulation of the, Mediterranean Sea from observations and modelling: a review", *Palaeogeography, Palaeoclimatology*, 158, pp. 153-173.

Pinardi, N., I. Allen , E. Demirov, P. De Mey, G.Korres, A.Lascaratos, P-Y. Le Traon, C.Maillard, G. Manzella, and C. Tziavos, 2003: The Mediterranean ocean Forecasting System: first phase of implementation (1998-2001), *Annales Geophysicae*, 21, 3-20.

Saravanan, R, G. Danabasoglu, S. C. Doney and J. C. McWilliams, 2000: Decadal Variability and Predictability in the Midlatitude Ocean–Atmosphere System. Journal of Climate: 13,1073–1097.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330

Zhu, Y., Z. Toth, R. Wobus, D. Richardson and K. Mylne. 2002: The Economic Value Of Ensemble-Based Weather Forecasts**.** *Bulletin of the American Meteorological Society*:  83, 73–83.

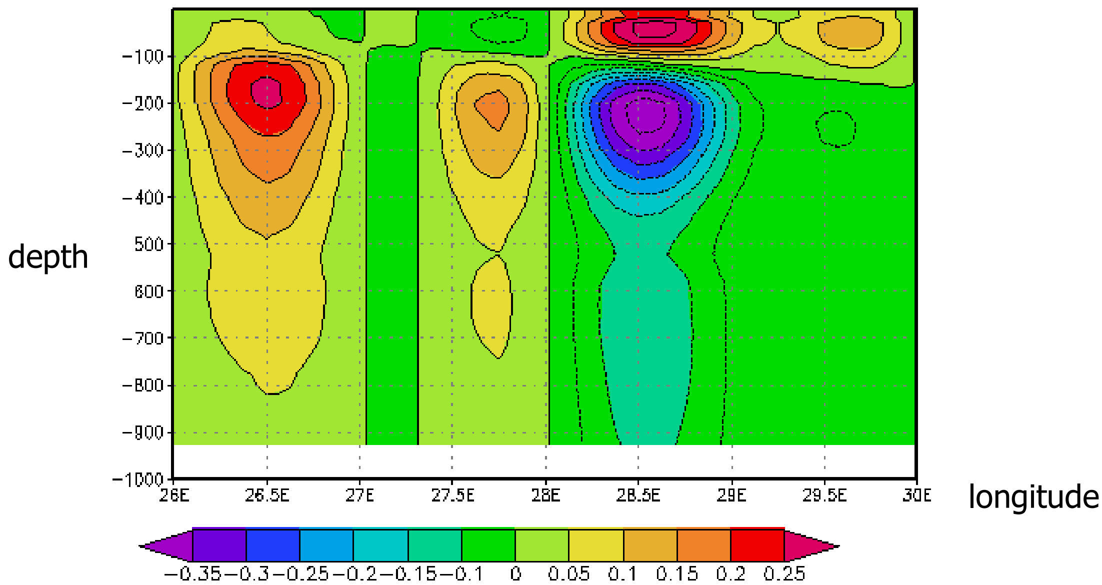


Fig. 1 The vertical structure of the temperature initial perturbation along a longitudinal and vertical cross section located at 34 degrees N (see Fig. 5). The units are $^{o}$C.


Fig.2  The distribution of Computing nodes of the Italian Grid. Each node contains a farm with from 10 to several hundreds PC, so-called Worker Nodes.

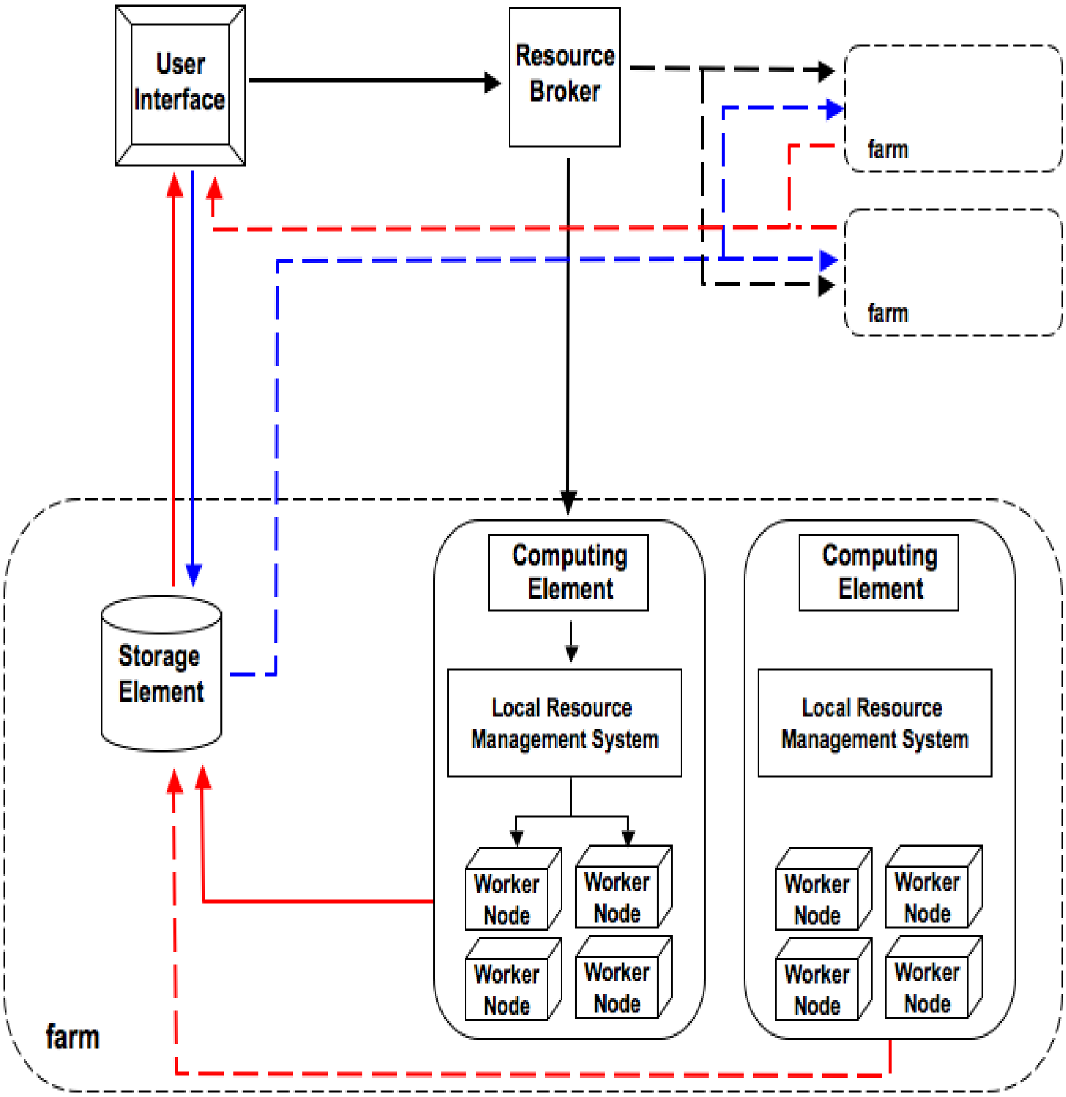


Fig. 3 The Grid infrastructure schematic components with the information flow: the blue lines indicate flow of files and instructions for the Replication phase 1, the black lines the Execution phase 2 and the red lines the Downloading phase 3, as described in the text. The dashed lines are the repetition of the continuous lines information flow.

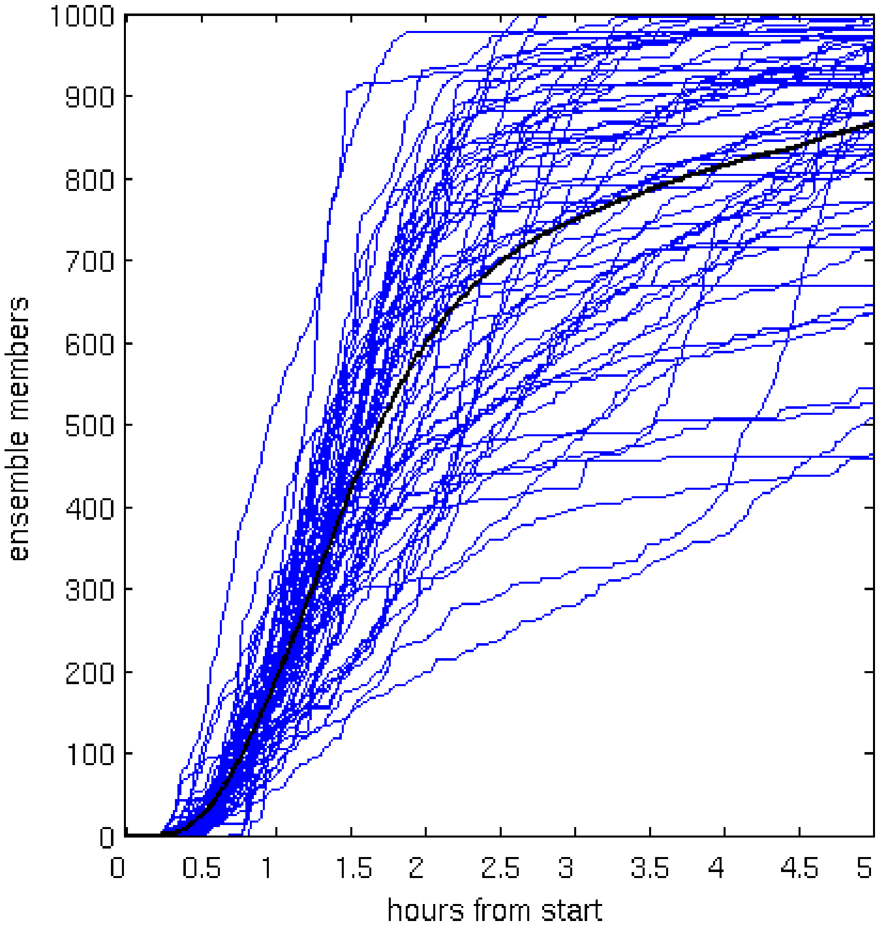


Fig. 4  Number of member jobs successfully carried out from 67 ensemble experiments launched on the Grid as a function of time. Each experiment is set to last maximum 5 hours and should run as much as 1000 jobs. The central black curve is the average.

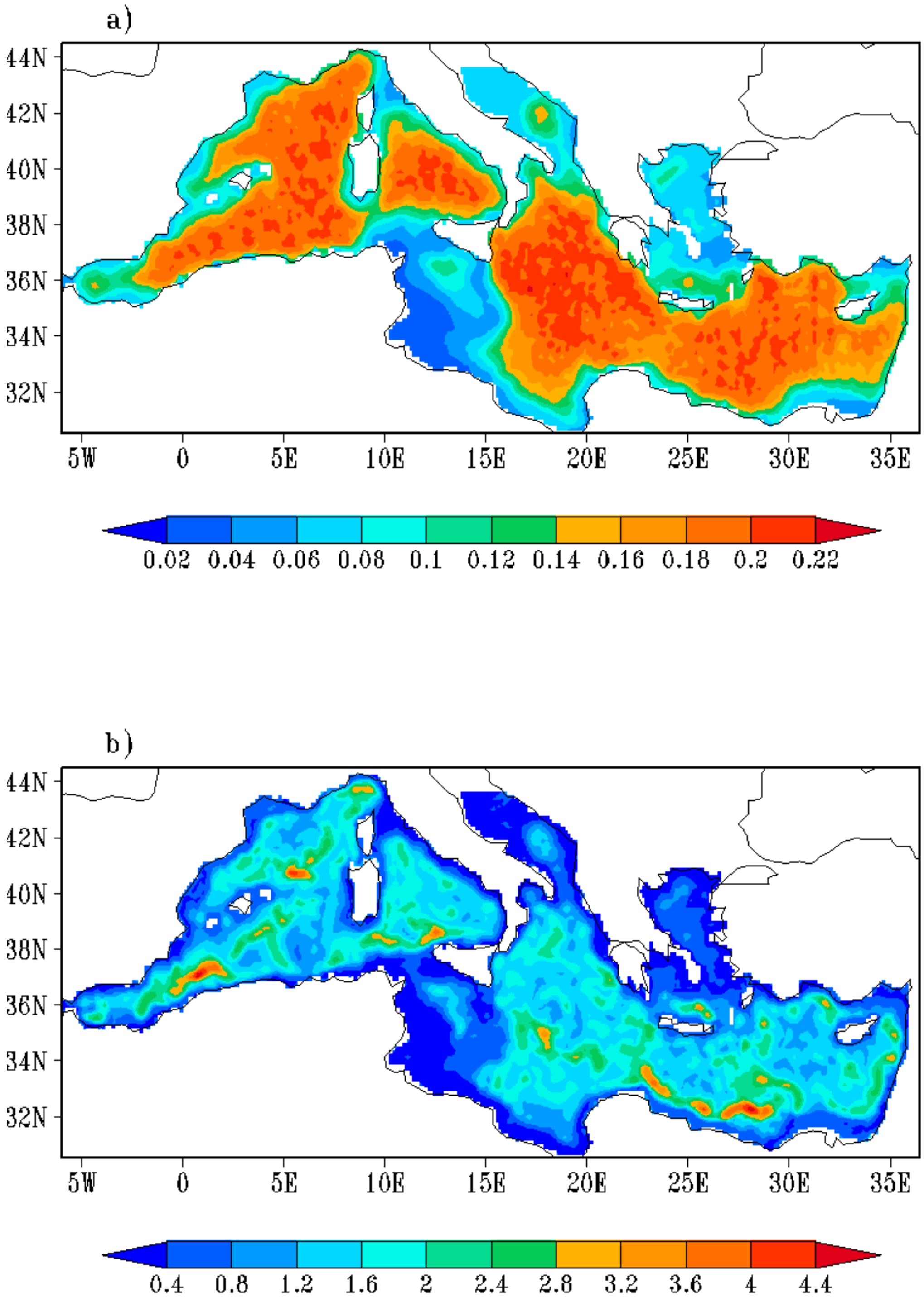


Fig. 5 The amplitude and structure of the standard deviation for SSH at forecast day 1 (a) and 10 (b). The 500 members ensemble mean has been subtracted and the units are cm. Please note the different scale of the two pictures

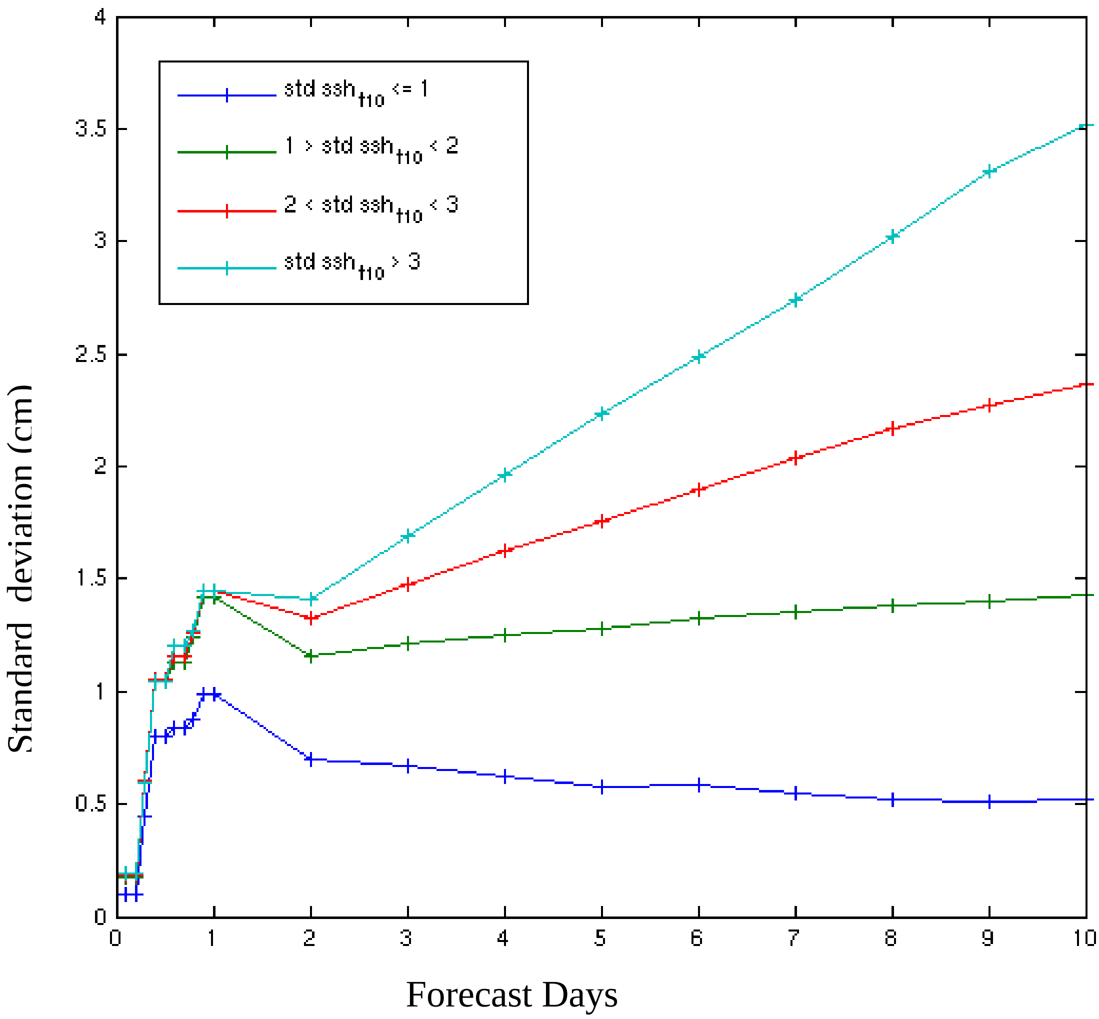


Fig. 6 The growth in amplitude of the standard deviation for the 10 days ensemble forecast experiment with 500 members. Different curves are averages done in regions of Fig. 5 with different standard deviations at day 10.