Journal of Marine Systems 78 (2009) S282–S289

Contents lists available at ScienceDirect

Journal of Marine Systems

journal homepage: www.elsevier.com/locate/jmarsys

Improved ocean prediction skill and reduced uncertainty in the coastal region from multi-model super-ensembles

Michel Rixen [a,*], Jeffrey W. Book [b], Alessandro Carta [a], Vittorio Grandi [a], Lavinio Gualdesi [a], Richard Stoner [a], Peter Ranelli [a], Andrea Cavanna [a], Pietro Zanasca [a], Gisella Baldasserini [a], Alex Trangeled [a], Craig Lewis [a], Chuck Trees [a], Rafaelle Grasso [a], Simone Giannechini [a], Alessio Fabiani [a], Diego Merani [a], Alessandro Berni [a], Michel Leonard [a], Paul Martin [b], Clark Rowley [b], Mark Hulbert [b], Andrew Quaid [b], Wesley Goode [b], Ruth Preller [b], Nadia Pinardi [c], Paolo Oddo [c], Antonio Guarnieri [c], Jacopo Chiggiato [d,h], Sandro Carniel [d], Aniello Russo [e], Martina Tudor [f], Fabian Lenartz [g], Luc Vandenbulcke [g]

[a] *NATO Undersea Research Center (NURC), La Spezia, Italy*
[b] *Naval Research Laboratory, Stennis Space Center, USA*
[c] *INGV, Bologna, Italy*
[d] *ISMAR-CNR, Venice, Italy*
[e] *Università Politecnica delle Marche, Ancona, Italy*
[f] *Meteorological and hydrological service, Zagreb, Croatia*
[g] *MARE-GHER, University of Liege, Belgium*
[h] *ARPA-SIM Emilia Romagna, Italy*

ARTICLE INFO

Article history:
Received 13 September 2008
Received in revised form 30 November 2008
Accepted 22 January 2009
Available online 28 February 2009

Keywords:
Ocean prediction skill
Uncertainties
Multi model super-ensembles
Coastal environments
Kalman Filter
Particle Filter

ABSTRACT

The use of Multi-model Super-Ensembles (SE) which optimally combine different models, has been shown to significantly improve atmospheric weather and climate predictions. In the highly dynamic coastal ocean, the presence of small-scales processes, the lack of real-time data, and the limited skill of operational models at the meso-scale have so far limited the application of SE methods. Here, we report results from state-of-the-art super-ensemble techniques in which SEPTR (a trawl-resistant bottom mounted instrument platform transmitting data in near real-time) temperature profile data are combined with outputs from eight ocean models run in a coastal area during the Dynamics of the Adriatic in Real-Time (DART) experiment in 2006. New Kalman filter and particle filter based SE methods, which allow for dynamic evolution of weights and associated uncertainty, are compared to standard SE techniques and numerical models. Results show that dynamic SE are able to significantly improve prediction skill. In particular, the particle filter SE copes with non-Gaussian error statistics and provides robust and reduced uncertainty estimates.

1. Introduction

An increasing number of models are routinely providing operational (atmospheric) weather forecasts and climate predictions. The use of model ensembles has become an important method of investigating dispersion problems (Galmarini et al., 2004), tracking individual model errors (for example from initial and boundary conditions, numerical discretization, turbulence closure), increasing forecast skill, and reducing uncertainties (Lermusiaux, 1999; Lermusiaux et al., 2006) in highly dynamic and complex environments where predictability is limited (Lorenz, 1963; Roe and Baker, 2007). Model biases are challenging to remove in short-term forecasts but may be addressed by statistical tools. The multi-model Super-Ensemble (SE) technique (Krishnamurti et al., 1999), which uses an optimised combination of an ensemble of models has previously been demonstrated to improve weather, seasonal and interannual forecast skill in atmospheric (Shin and Krishnamurti, 2003a,b; Yun et al., 2005) and ocean (Rixen and Ferreira-Coelho, 2006, 2007; Rixen et al., 2008, 2009-this issue; Logutov and Robinson, 2005) models over simple-ensemble and bias-removed ensemble means. SE methods (Williford et al., 2003) have been further improved by the use of dynamic (Shin and Krishnamurti, 2003a), regularization (Yun et al., 2003), non-linear (Rixen and Ferreira-Coelho, 2007) and probabilistic (Rajagopalan et al., 2002) techniques. These methods all aim at finding a combination of models that optimally agrees with reference data over a training period (the hindcast); this combination is subsequently used to produce a SE forecast. A critical aspect for all SE methods is therefore whether the regression solution is capable of extrapolation in time and is applicable to future events. In other words, the learning should be adequate to provide generalization skills.

* Corresponding author.
E-mail address: rixen@nurc.nato.int (M. Rixen).

0924-7963/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmarsys.2009.01.014

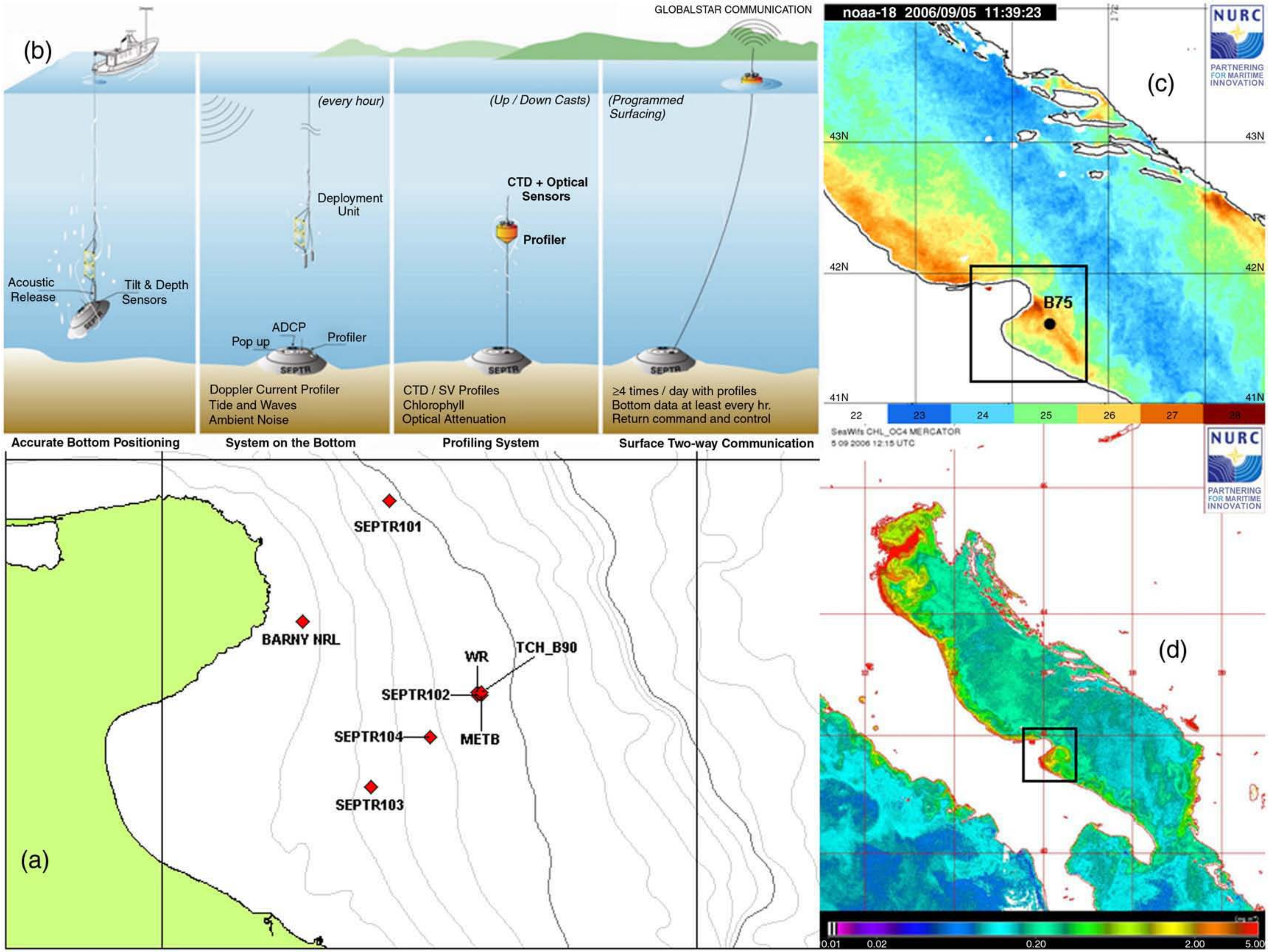

Fig. 1. Area of interest of the DART06 experiments: a) location of moorings in the Gulf of Manfredonia including SEPTR 104 b) SEPTR operations sketch: deployment, profiling and data transmission c) illustration of small-scale instabilities along the WAC as seen from an infrared NOAA AVHRR Sea Surface Temperature image (°C) on 5 September 2006 d) same as (c) but for chlorophyll A (mg m^{-3}).

Operational implementation of SE in the atmosphere has been quite straightforward due to the reliability of observational data streams and the robustness of the models. However, in the ocean, the lack of long real-time data time series – especially in heavily fished areas – and a limited suite of operational models have so far limited the application of such promising techniques.

The use of SE methods for coastal ocean prediction was explored during two Dynamics of the Adriatic in Real-Time (DART06A in winter and DART06B in summer) inter-disciplinary and multi-institutional experiments carried out in 2006 near the Gargano Peninsula and in the central Adriatic Sea (Mediterranean Sea), areas with intense navigation and fishing activity (e.g. Cushman-Roisin et al., 2001; Burrage et al., 2009-this issue). The purpose of these experiments was to assess ocean monitoring and prediction skill (Taylor, 2001; Robinson et al., 2002) in highly dynamic areas. The project specifically focused on small-scale processes resulting from instabilities of the Western Adriatic Current (WAC), which flows southward along the Italian coast. This area, like many coastal areas, is subject to intense mesoscale activity (Fig. 1). The time-scales of eddies, fronts and filaments are typically of the order of a few hours to a day. Other processes, e.g. the non-linear interactions of the instabilities with internal waves and tides, make operational forecasting in the area even more complex and challenging. Results presented hereafter refer to the B75 mooring (SEPTR 104) during the summer experiment.

2. Methods

With the exception of the simple ensemble mean method, the fusion of the different models by elaborate and reliable SE techniques require independent observations.

The delivery of real-time observations was ensured by the SEPTR (Perkins et al., 2000; Grandi et al., 2005), a bottom-mounted platform in a trawl-resistant configuration (Fig. 1), equipped with an Acoustic Doppler Current Profiler (ADCP) and a winch-controlled profiling unit fitted with Conductivity Temperature Depth (CTD), wave, and optical sensors. Data were transmitted through a GLOBALSTAR link every 6 h after profiling the water column. In practice not all of the data were successfully received in near-real time and thus the SE methods were applied to the full data set after recovery rather than the more limited and patchy real-time data set. However, future improvements to SEPTR technology will be focused on improving data transmission and hopefully make future near-real time applications more practical. The collection of models from the various home institutions and those run onboard R/V ALLIANCE was ensured through continuous mirroring of the NURC and R/V ALLIANCE FTP servers over a dedicated, high-bandwidth satellite-link system (a standard 2-way SATCOM connection complemented by a Digital Video Broadcasting System asymmetric link).

Eight different medium- to high-resolution ocean models (Fig. 2, see details of the model implementations in Appendix A) were run in the framework of DART. These models exhibited different skills, dynamic responses, and biases, as a result of their different physical assumptions and configurations (numerical discretization, initial and boundary conditions, atmospheric forcing, data assimilation, turbulence closure schemes and sub-grid scale parameterizations). This diversity offered a good opportunity to test SE methods.

The high spatio-temporal variability of prediction skill makes the application of ensemble techniques difficult in the ocean. Naïve averaging (i.e., ensemble mean, hereinafter ENSMEAN) exhibits poor skill and calls for methods with increased complexity to cope with biases (i.e., a mean of models corrected for their respective biases, hereinafter ENSMEANUNBIASED) or perform collective bias correction (i.e. a least-square linear regression between the data and the models, hereinafter LINREG) (Krishnamurti et al., 1999). To allow a dynamic evolution of model combinations, a Kalman filter (hereinafter KALMAN) (Kalman, 1960) can be used; this assumes that the weight statistics follow a normal distribution. The full Kalman filter has been implemented here without additional hypotheses. The sequential importance resampling filter (hereinafter PARTICLE) (van Leeuwen,

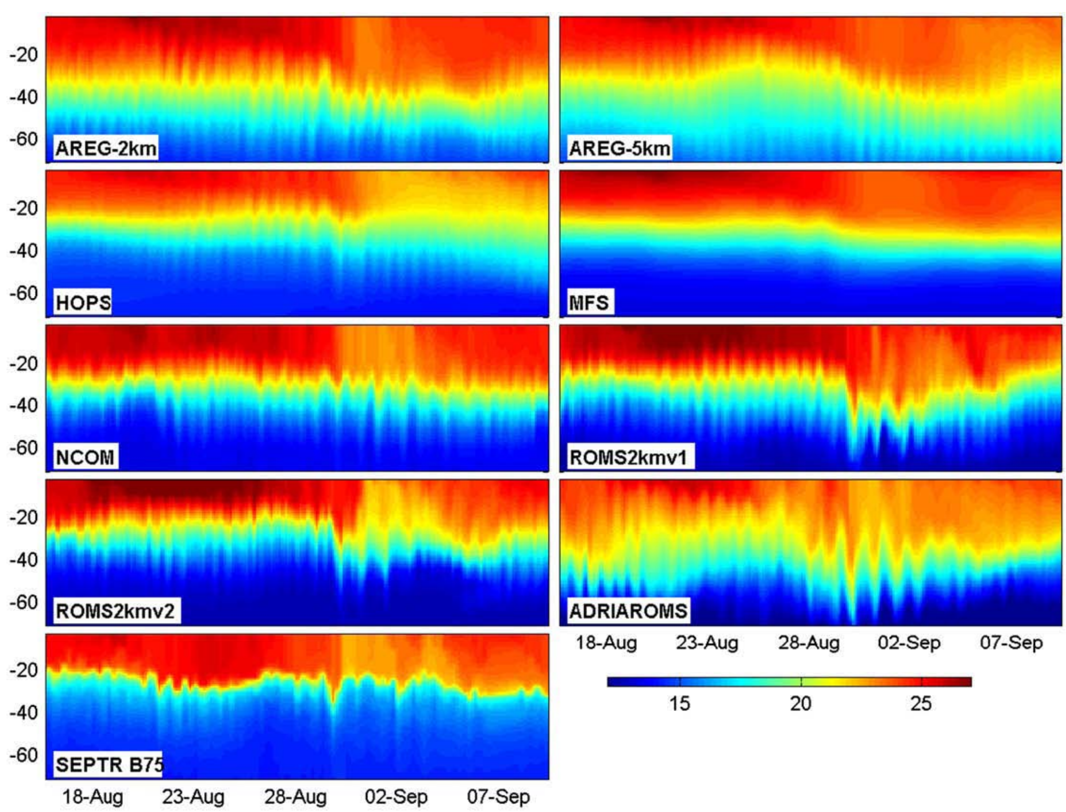

Fig. 2. Time series of temperature (°C) versus depth (m) at mooring B75 from the eight ocean models (see Appendix A) and SEPTR 104 real-time data.

2003) was also tested to challenge the assumption of the weights having a Gaussian distribution (see additional details in Appendix B). For both KALMAN and PARTICLE, the *a priori* error statistics were optimised by cross-validation, yielding observational errors of 0.2 °C (including sensor errors and unmeasured scales) and SE model weight errors of 0.05 (this corresponds roughly to a change in weights of 20%/day).

All these methods may be complemented with additional 'tricks' (Rixen and Ferreira-Coelho, 2007): 1) normalizing the models and data for better numerical conditioning (NORM suffix), 2) adding a synthetic model (predicting 1 all the time) to allow for a constant or dynamic bias to be removed (for all SE methods: LINREG, KALMAN, PARTICLE), and 3) a regularization using empirical orthogonal functions (EOF suffix) by retaining only the dominant modes represented in the models (95% of the variance in the present study). Method 3) avoids collinearities between the models (which may generate numerical problems); hence it usually significantly improves generalization skills.

All the methods are computed at every single depth, yielding different weights at different depths. Additionally, dynamic methods use time-evolving weights, initialized with the LINREG solution. It should be noted that observations that are assimilated by numerical models could be assimilated in the SE as well, although this was not considered in the present work.

3. Results

The results focus on 24-hour predictions of temperature with a learning period starting 16 August 2006 and an evaluation period from 4 to 9 September 2006. For a given 'present' time, past data are used for the learning phase (linear regression, Kalman filter or particle filter data assimilation). Future data are used for verification.

The time series of temperature from SEPTR 104 at mooring B75 show a gradual cooling of the surface and deepening of the thermocline during a major cooling event occurring in early September (Fig. 2). Qualitatively, all models were able to reproduce the general patterns identified in the SEPTR data. Quantitatively (Figs. 3 and 4), the models exhibit various kinds of errors including systematic biases, amplitude and phase errors, offsets in the thermocline depth, strength and response of the thermocline, discrepancies in the penetration of mixing events, anomalous under or over-heating at the surface, and weak temporal variability.

It can be noticed from Fig. 3 that errors for the different Super-Ensemble solutions decrease with increased complexity from the classic ENSMEAN and ENSMEANUNBIASED to LINREG, KALMAN, and PARTICLE complemented by regularization. The ENSMEAN, although reducing somewhat the errors at the bottom and surface is unable to correct the large errors at the level of the thermocline because it is incapable of integrating the data information. The ENSMEANUNBIASED allows for a substantial correction of the errors. The errors are further reduced by the LINREG and by dynamic methods as KALMAN_NORM_EOF and PARTICLE_NORM_EOF. The EOF regularization, when retaining 95% of the variance, typically 'compresses' the 8 models down to 2 to 6 normal modes, thus highlighting the fact that some of the models are closely correlated. The resulting number of modes varies with depth and is usually higher near the thermocline. RMS errors (Fig. 4) have been reduced from a range of 1.35–3.21 °C for the individual models to less than 0.51–0.53 °C for the dynamic methods with regularization. The corresponding biases (0.02–2.15 °C) have, in most cases, been reduced to less than 0.14 °C and the correlations increased from 0.86–0.96 to 0.99. The signal energy discrepancies have similarly reduced from 0.12–1.95 to 0.23. The overall skill (Taylor, 2001), which takes into account both the difference in the signal standard deviation and the correlation of the signal (0 being the lowest skill and 1 the highest) has been increased from 0.53–0.90 to 0.98, i.e., by at least 8% as a conservative estimate. This suggests that the signal energy of the prediction and the truth are in good agreement and that the prediction is also better correlated to the truth.

Ensemble methods may provide improved prediction skill but also offer as a by-product an estimate of the associated uncertainty (i.e. the confidence interval) at marginal cost (see additional details in Appendix B). For operational purposes, overestimation or

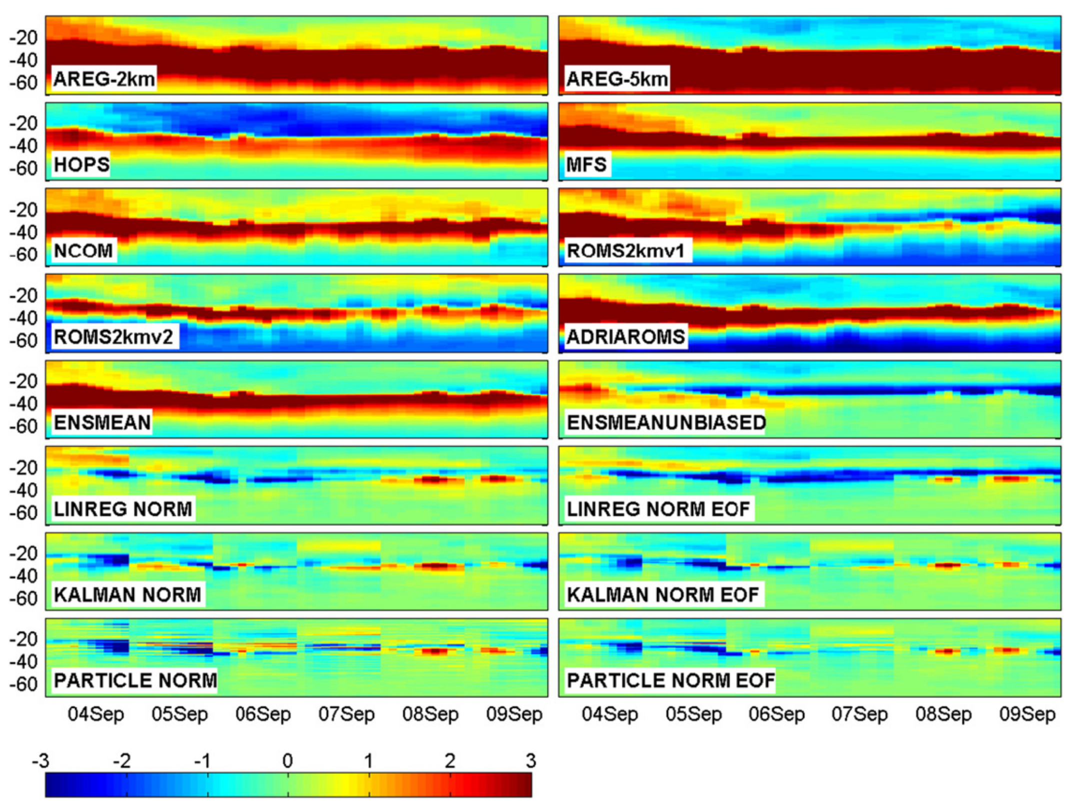

Fig. 3. Time series of 24-hour temperature (°C) forecast errors versus depth (m) for 6 consecutive days of individual models and super-ensemble predictions. From left to right and top to bottom: the 8 individual models, ENSMEAN, ENSMEANUNBIASED, LINREG_NORM, LINREG_NORM_EOF, KALMAN_NORM, KALMAN_NORM_EOF, PARTICLE_NORM and PARTICLE_NORM_EOF.

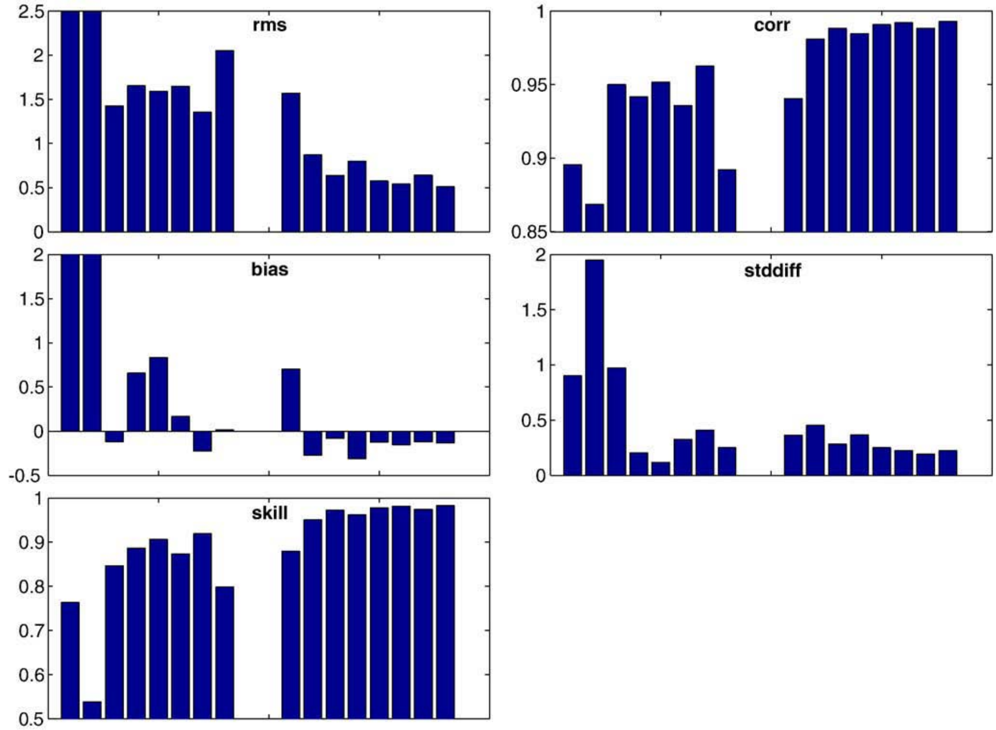

Fig. 4. Error statistics of 24-hour temperature (°C) forecast. From left to right and top to bottom: root mean square error, correlation, bias, signal standard deviation difference and skill for the 8 different individual models (left) and SE models (right) of Fig. 3.

underestimation of the uncertainty can be a serious issue. Decision makers could ignore a valuable prediction that is assigned too much uncertainty or be over-confident of a prediction that is assigned too little uncertainty. Narrowing down the confidence interval for the predictions while minimizing these types of failures is hence of great value to the end-user.

Fig. 5 illustrates the time series of the different uncertainty predictions for a 99.7% confidence interval (or 3 standard deviations) and an *a posteriori* verification of the quality of these estimates. The mean ENSMEAN uncertainty is 4.09 °C, as a result of the wide spread of the individual models. By injecting SEPTR data, the SE methods narrow down the uncertainty estimates to mean values of 2.21 °C for the ENSMEANUNBIASED, 3.64 °C for LINREG_NORM, 2.05 °C for LINREG_NORM_EOF, 1.67 °C for the KALMAN_NORM_EOF and 1.01 °C for the PARTICLE_NORM_EOF. Compared to the KALMAN_NORM_EOF (which assumes Gaussian error statistics), the non-Gaussian capability of the particle filter allows error estimate reduction by a further 40%. Uncertainty within each day increases for dynamic methods as a result of the intrinsic uncertainty of the forecast-model weights (hence the discontinuity at the end of the 24 h forecast period because of the running 'present time' window). *A posteriori* verification as to whether the ground truth is within the prediction-uncertainty range demonstrates that the failure rate is around 3% for ENSMEAN and ENSMEANUNBIASED. For linear regression methods and KALMAN_NORM, the failure rate is 0% due to the overestimation of the error estimate and less than 1% for the remaining SE methods.

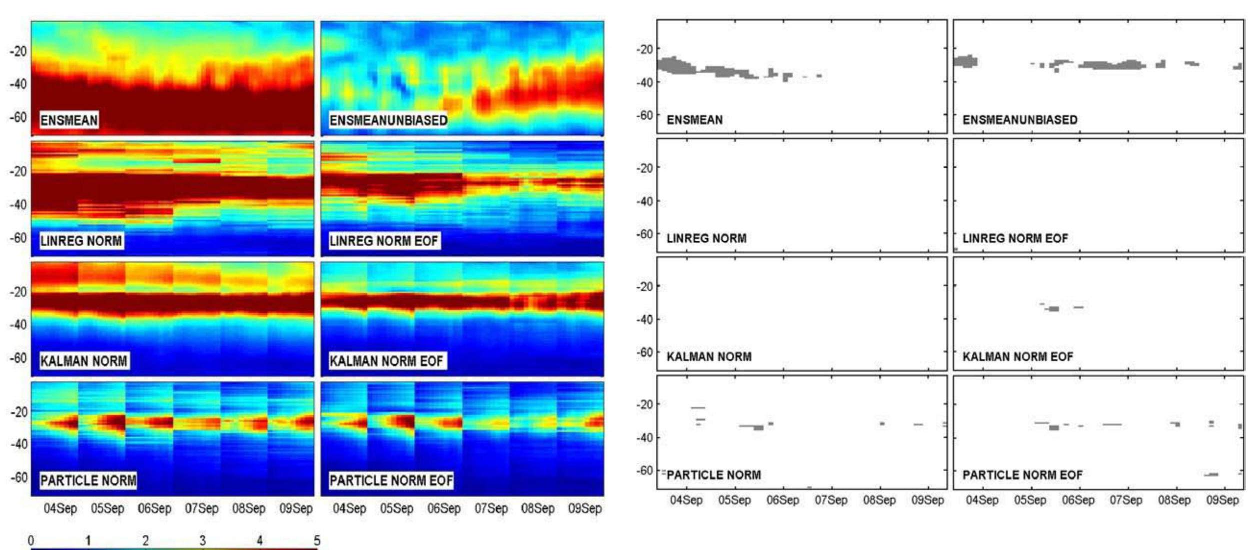

Fig. 5. Time series of uncertainty and *a posteriori* verification of 24-hour temperature forecast (°C) versus depth (m): (left) uncertainty (99.7% confidence interval); (right) corresponding *a posteriori* check if observed values is not within the 99.7% confidence interval (black).

For longer lead times (not shown), the advantage of complex dynamic methods becomes less obvious. For a 36-hour forecast, the particle filter and Kalman filter methods still show better error statistics. At 48 h and beyond, the linear regression and unbiased ensemble mean provide better results. It should be stressed however that for a forecast range of 24–72 h almost all the SE methods show better statistics than any of the individual models. It is expected that the SE will directly benefit from continuous improvements on individual models skills.

4. Conclusions

Our results support the concept of 'self-modifying' models (Dee, 1995; Lermusiaux, 2007). The SE methods outperform the individual models on several error measures. Skill improves with increased method complexity on 24-hour forecasts. Dynamic, non-Gaussian and regularized SE techniques exhibit better skill and lower uncertainty. Accurate predictions and reliable uncertainty estimates are equally valuable products for decision makers. Setting aside the SEPTR data transmission issue, operational implementation of the various methods is straightforward. We have shown that, at marginal cost, the unique approach fusing operational predictions and real-time data from the SEPTR bottom-mounted platform in the trawl-resistant configuration offers a new paradigm for improved predictions and reliable error estimates for a potentially wide range of environmental parameters in shallow waters.

Acknowledgments

This work was performed by the NURC funded by the North Atlantic Treaty Organization in the framework of the DART Joint Research Program with the Naval Research Laboratory at Stennis Space Center, and in collaboration with 33 partner institutions whose contribution is gratefully acknowledged. Funding for the Naval Research Laboratory work was provided by the Office of Naval Research under Program Element Number 0602435N.

Appendix A. Description of the numerical models used in the super-ensemble

A.1. AdriaROMS, ROMS2kmV1 and ROMS2kmV2

AdriaROMS is the operational ocean forecast system for the Adriatic Sea running at ARPA-SIM (http://www.arpa.emr.it/sim/?mare). It is based on the Regional Ocean Modelling System (ROMS, detailed kernel description is in Shchepetkin and McWilliams, 2005). This Adriatic configuration has a variable horizontal resolution, ranging from 3 km in the north Adriatic to ~10 km in the south, with 20 s-coordinates levels. Surface forcing is provided by the Limited Area Model Italy (LAMI, local implementation of the model LM, Steppeler et al., 2003), a non-hydrostatic numerical weather prediction model with 7 km horizontal resolution. MFS data (Tonani et al., 2008) are used at the open boundary to the south with superimposed four major tidal harmonics (S2, M2, O1, K1), from the work of Cushman-Roisin and Naimie (2002) following Flather (1976). Forty-eight rivers (and springs) are included, using monthly climatological value from Raicich (1994). Persistence of the daily discharge measured one day previously is used for the Po river. Additional details can be found in Chiggiato and Oddo (2008).

ROMS has been run also in hindcast mode, with a different configuration and a finer grid compared to AdriaROMS. The horizontal resolution is 2 km overall the basin, with 20 s-coordinate levels on the vertical. The model has been spun up from a rest state defined by objective analysis of URANIA CTD casts collected in January 2006 (courtesy of GOS CNR-ISAC, Rome) alone (ROMS2kmV2 - version 2), or an objective analysis of URANIA data merged with winter climatological data (Artegiani et al., 1997) as background (ROMS2kmV1 version 1). ROMS2km makes use of an MPDATA family advection scheme (Margolin and Smolarkiewicz, 1998) for tracers, third order upstream scheme (Shchepetkin and McWilliams, 1998) for momentum, with weak background diffusivity and no viscosity. The generic length scale turbulence closure model is used for vertical mixing (as implemented by Warner et al., 2005). A density jacobian with spline reconstruction of the vertical profiles is used for the pressure gradient (Shchepetkin and McWilliams, 2003). Surface forcing is provided by LAMI with turbulent fluxes computed following Fairall et al. (2003) and evaporation-precipitation flux included. At the southern open boundary, tracers coming from the climatological dataset MEDATLAS are prescribed with relaxation-radiation and four superimposed major tidal harmonics (S2, M2, O1, K1), from the work of Cushman-Roisin and Naimie (2002) following Flather (1976). Forty-eight rivers (and springs) are included, using monthly climatological value from Raicich (1994), except for the Po river for which daily observed discharge values were used.

A.2. HOPS

The Harvard Ocean Prediction System (HOPS) (Lozano et al., 1996) implementation has a resolution of 3 km and 21 sigma vertical levels. Air-sea fluxes are taken from the Limited Area Model Italy (LAMI). Open boundaries are set for Otranto and the Po river (considered as a channel); Orlanski radiation conditions for temperature, salinity and velocity, and specific boundary conditions (Spall and Robinson, 1990) for transport stream function and vorticity were used. A constant flux of 1000 m^3/sec was set for the Po river. The model uses a rigid lid and does not include tides. The turbulence closure follows Pacanowski and Philander (1981). Initial conditions are derived from the AREG-5 km (Oddo et al., 2006) hindcast daily average on 2 August 2006. Temperature and salinity data collected by R/V Dallaporta and R/V Alliance (14–27 August) and derived geostrophic velocities were intermittently assimilated via Optimal Interpolation (Robinson et al., 1998) integrated with the MED2 summer climatology in areas not covered by data in the Southern Adriatic Sea.

A.3. NCOM

NCOM and its setup for the Adriatic are described in Martin et al. (2006). The domain consists of the entire Adriatic Sea and includes the Strait of Otranto and a small part of the northern Ionian Sea. The horizontal grid resolution is 1020 m. The vertical grid consists of 32 total layers, with 22 sigma layers used from the surface down to a depth of 291 m and level coordinates used below 291 m. Daily boundary conditions were taken from hindcasts and forecasts of a global model (Barron et al., 2004). Tidal forcing was provided for eight constituents using tidal elevation and depth-averaged normal and tangential velocities at the open boundaries from the Oregon State University tidal databases. Tidal potential forcing was used in the interior. Atmospheric forcing was obtained from the Aire Limitee Adaptation Dynamique development InterNational (ALADIN) atmospheric model run by the Croatian Meteorological and Hydrological Service. The NCOM sea surface temperature (SST) was relaxed towards a satellite SST analysis. River and runoff inflows for the Adriatic were taken from the monthly climatological database of Raicich (1994), except for the Po river for which daily observed discharge values were used (courtesy of ARPA-SIM Emilia Romagna).

A.4. MFS, AREG-5 km, AREG-2 km

The Italian National Institute for Geophysics and Volcanology (INGV) provided data from three operational ocean forecasting systems, namely the Adriatic REGional forecasting system, AREG, with horizontal resolution of 5 and 2 km (http://gnoo.bo.ingv.it/afs/) and data from the Mediterranean ocean Forecasting System, MFS (http://gnoo.bo.ingv.it/mfs/). Table A1 shows the major features of the three forecasting systems. Additional details can be found in Oddo et al. (2005, 2006) and Tonani et al. (2008).

Table A1
Implementation details of the set of three INGV forecasting systems used in the framework of the DART06 experiment.

Name	AREG ~5 km	AREG-2 km	MFS
Code	POM	POM	OPA
Resolution dx,dy,dz	~2 km (31 sigma)	~2 km (31 sigma)	~6.5 km (72 zeta)
Air–sea boundary conditions	Interactively computed from operational ECMWF 0.5° 6 h		
Lateral boundary conditions	From MFS operational output Climatological river Raicich (1994), daily observed Po (courtesy of ARPA-SIM)		Atlantic closed No rivers
Tides included	No tide-free surface		No tide – filtered free surface
Turbulence closure	Mellor and Yamada (1982)		KPP (Large et al., 1994)
Initial conditions and/or data assimilation	1-1-1999 from climatological run. No assimilation	1-1-1999 from MFS interannual run No assimilation	1-1-1997 climatology SOFA data assimilation De Mey et al. (2002)

Appendix B. The filtering problem

Filtering is the problem of making optimal estimates of a system as observations become available, taking into account process and measurement noises. The best known algorithm to solve the problem under the assumption of Gaussian error statistics is the Kalman Filter (Kalman, 1960). For non-linear systems and non-Gaussian error statistics, the superiority of Particle Filters (or Sequential Monte-Carlo techniques) has been demonstrated in numerous studies (e.g. Doucet et al., 2001). Fig. B1 illustrates the non-Gaussian nature of the probability distribution functions of weight statistics and observations in the present study. These distributions clearly show significant departure from normal distributions with multiple peaks, evident skewness or kurtosis.

Uncertainties are estimated from the commonly used 'law of propagation of uncertainty', which in the case of super-ensembles yields that the uncertainty is the square mean root of model uncertainties, weighted by the corresponding model forecasts. In the case of the Kalman filter, an estimate of the weight errors is given by the diagonal elements of the error covariance matrix P. In the case of the particle filter, they are estimated by the variance of the population of particles (variance of weights). For the simple linear regression, they are obtained by cross-validation.

References

Artegiani, A., Bregant, D., Paschini, E., Pinardi, N., Raicich, F., Russo, A., 1997. The Adriatic Sea general circulation, part II: Baroclinic circulation structure. J. Phys. Oceanogr. 27, 1515–1532.

Barron, C.N., Kara, A.B., Hurlburt, H.E., Rowley, C., Smedstad, L.F., 2004. Sea surface height predictions from the global navy coastal ocean model during 1998–2001. J. Atmos. Oceanic Technol. 21, 1876–1893.

Burrage, D.M., Book, J.W., Martin, P.J., 2009. Eddies and filaments of the Western Adriatic Current near Cape Gargano: analysis and prediction. J. Mar. Syst. 78, S205–S226 (this issue).

Chiggiato, J., Oddo, P., 2008. Operational ocean models in the Adriatic Sea: a skill assessment. Ocean Sci. 4, 61–71.

Cushman-Roisin, B., Naimie, C.E., 2002. A 3D finite-element model of the Adriatic tides. J. Mar. Syst. 37, 279–297.

Cushman-Roisin, B., Gacic, M., Poulain, P.M., Artegiani, A., 2001. Physical Oceanography of the Adriatic Sea: Past, Present and Future. Kluwer Academic Publishers, Dordrecht.

Dee, D., 1995. On-line estimation of error covariance parameters for atmospheric data assimilation. Mon. Weather Rev. 123, 1128–1145.

De Mey, P., Benkiran, M., 2002. A multivariate reduced order optimal interpolation method and its application to the Mediterranean basin-scale circulation. In: Pinardi, N., Woods, J.D. (Eds.), Ocean Forecasting: conceptual basis and applications. Spinger-Verlag, pp. 281–306.

Doucet, A., de Freitas, N., Gordon, N., 2001. Sequential Monte Carlo Methods in Practice. Springer978-0387951461.

Fairall, C.W., Bradley, E.F., Hare, J.E., Grachev, A.A., Edson, J.B., 2003. Bulk parameterisations of air–sea fluxes: updates and verification for the COARE algorithm. J. Climate 16, 571–591.

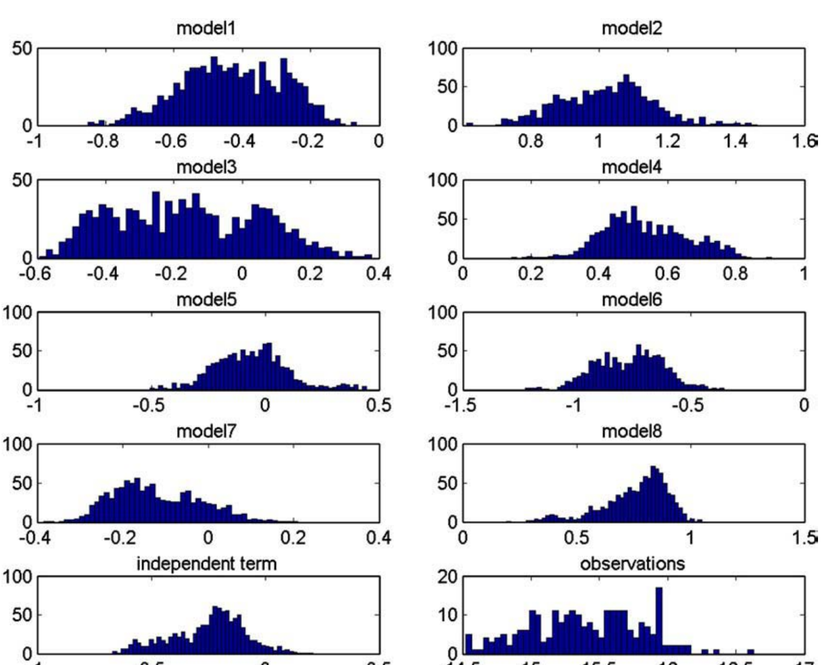

Fig. B1. Probability distributions functions (PDFs) of weights at the end of the particle filtering on 4 September 9:00 at 40 m depth for 24-hour forecast. From top to bottom and left to right: PDFs on the 8 individual models and independent term (1000 particles), and typical PDF on temperature (°C) observations for the whole period.

Flather, R.A., 1976. A tidal model of the northwest European continental shelf. Mem. Soc. R. Sci. Liege 6 (10), 141–164.
Galmarini, S., Bianconi, R., Addis, R., Andronopoulos, S., Astrup, P., Bartzis, J.C., Bellasio, R., Buckley, R., Champion, H., Chino, M., D'Amours, R., Davakis, E., Eleveld, H., Glaab, H., Manning, A., Mikkelsent, T., Pechinger, U., Polreich, E., Prodanova, M., Slaper, H., Syrakov, D., Terada, H., Van Der Auwera, L., 2004. Ensemble dispersion forecasting, Part II: application and evaluation. Atmos. Environ. 38 (28), 4619–4632.
Grandi, V., Carta, A., Gualdesi, L., de Strobel, F., Fioravanti, S., 2005. An overview of SEPTR: shallow water environmental profiler in a trawl-safe real-time configuration, current measurement technology. Proceedings of the IEEE/OES Eighth Working Conference.
Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Trans. ASME - J. Basic Eng. 82, 35–45.
Krishnamurti, T., Kishtawal, C., LaRow, T.E., Bachiochi, D.R., Zhang, Z., Williford, C.E., Gadgil, S., Surendran, S.S., 1999. Improved weather and seasonal climate forecasts from multimodel superensemble. Science 285 (5433), 1548–1550. doi:10.1126/science.285.5433.1548.
Large, W.G., McWilliams, J.C., Doney, S.C., 1994. Oceanic vertical mixing: a review and a model with nonlocal boundary layer parameterisation. Rev. Geophys. 32, 363–403.
Lermusiaux, P.F.J., 1999. Data assimilation via error subspace statistical estimation. Part II: Middle Atlantic Bight shelfbreak front simulations and ESSE validation. Mon. Weather Rev. 127 (7), 1408–1432.
Lermusiaux, P.F.J., 2007. Adaptive modeling, adaptive data assimilation and adaptive sampling. In: Jones, C.K.R.T., Ide, K. (Eds.), Refereed invited manuscript. Special issue on “Mathematical Issues and Challenges in Data Assimilation for Geophysical Systems: Interdisciplinary Perspectives”. Physica D, vol. 230, pp. 172–196.
Lermusiaux, P.F.J., Chiu, C.S., Gawarkiewicz, G.G., Abbot, P., Robinson, A.R., Miller, R.N., Haley, P.J., Leslie, W.G., Majumdar, S.J., Pang, A., Lekien, F., 2006. Quantifying uncertainities in ocean predictions. In: Paluszkiewicz, T., Harper, S. (Eds.), Oceanography, Special issue on “Advances in Computational Oceanography”, vol. 19, 1, pp. 92–105.
Logutov, O.G., Robinson, A.R., 2005. Multi-model fusion and error parameter estimation. Q. J. R. Meteorol. Soc. 131 (613), 3397–3408. doi:10.1256/qj.05.99.
Lorenz, E.N., 1963. Deterministic nonperiodic flow. J. Atmos. Sci. 20, 130–141.
Lozano, C.J., Robinson, A.R., Arango, H.G., Gangopadhyay, A., Sloan, Q., Haley, P.J., Anderson, L., Leslie, W.G., 1996. An interdisciplinary ocean prediction system: assimilation strategies and structured data models. In: Malanotte-Rizzoli, P. (Ed.), Modern Approaches to Data Assimilation in Ocean Modeling. Elsevier Oceanography Series, vol. 61. Elsevier, pp. 413–452.
Margolin, L., Smolarkiewicz, P.K., 1998. Antidiffusive velocities for multipass donor cell. SIAM J. Sci. Comput. 20, 907–929.
Martin, P.J., Book, J.W., and Doyle, J.D., 2006. Simulation of the northern Adriatic circulation during winter 2003, J. Geophys. Res. 111, C03S12. doi:10.1029/2006JC003511.
Mellor, G.L., Yamada, T., 1982. Development of a turbulence closure model for geophysical fluid problems. Rev. Geophys. Space Phys. 20, 851–875.
Oddo, P., Pinardi, N., Zavatarelli, M., 2005. A numerical study of the interannual variability of the Adriatic Sea (2000–2002). Sci. Total Environ. 353, 39–56.
Oddo, P., Pinardi, N., Zavatarelli, M., Coluccelli, A., 2006. The Adriatic Basin Forecasting System. Acta Adriat. 47 (Suppl.), 169–184.
Pacanowski, R.C., Philander, S.G.H., 1981. Parameterization of vertical mixing in numerical models of tropical oceans. J. Phys. Oceanogr. 11, 1443–1451.
Perkins, H., de Strobel, F., Gualdesi, L., 2000. The BARNY Sentinel Trawl-Resistant ADCP Bottom Mount: design, testing, and application. IEEE J. Oceanic Eng. 25, 430–436.
Raicich, R., 1994. Note on the flow rates of the Adriatic rivers. Tech. Rep. RF 02/94, 8 pp., CNR. Ist. Sper. Talassografico, Trieste, Italy.
Rajagopalan, B., Lall, U., Zebiak, S.E., 2002. Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. Mon. Weather Rev. 130, 1792–1811.
Rixen, M., Ferreira-Coelho, E., 2006. Operational prediction of acoustic properties in the ocean using multi-model statistics. Ocean Model. 11 (3-4), 428–440. doi:10.1016/j.ocemod.2005.02.002.
Rixen, M., Ferreira-Coelho, E., 2007. Operational surface drift prediction using linear and non-linear hyper-ensemble statistics on atmospheric and ocean models. J. Mar. Syst. 65, 105–121.
Rixen, M., Ferreira Coelho, E., Signell, R., 2008. Surface drift prediction in the Adriatic using hyper-ensembles statistics on atmospheric, ocean and wave models: uncertainties and probability distribution areas. J. Mar. Syst. 69, 86–98.
Rixen, M., Le Gac, J.C., Hermand, J.P., Peggion, G., Ferreira-Coelho, E., 2009. Super-ensemble forecasts and resulting acoustic sensitivities in shallow waters. J. Mar. Syst. 78, S290–S305 (this issue).
Robinson, A.R., Lermusiaux, P.J.F., Sloan, N.Q., 1998. Data assimilation. In: Brink, K.H., Robinson, A.R. (Eds.), The Sea, vol. 10. John Wiley & Sons, pp. 541–594.
Robinson, A.R., Lermusiaux, P.F.J., Haley, P.J., Leslie, W.G., 2002. Predictive Skill, predictive capability and predictability in ocean forecasting. Proceedings of “The OCEANS 2002 MTS/IEEE” conference. Holland Publications, pp. 787–794.
Roe, G.H., Baker, B.M., 2007. Why is climate sensitivity so unpredictable? Science 318 (5850), 629. doi:10.1126/science.1144735.
Shchepetkin, A., McWilliams, J.C., 1998. Quasi-monotone advection schemes based on explicit locally adaptive dissipation. Mon. Weather Rev. 126, 1541–1580.
Shchepetkin, A.F., McWilliams, J.C., 2003. A method for computing horizontal pressure-gradient force in an oceanic model with a non-aligned vertical coordinate. J. Geophys. Res. 108 (C3), 3090. doi:10.1029/2001JC001047.
Shchepetkin, A.F., McWilliams, J.C., 2005. The regional ocean modelling system: a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Model. 9, 347–404.
Shin, D., Krishnamurti, T., 2003a. Short- to medium-range superensemble precipitation forecasts using satellite products: 1. Deterministic forecasting. J. Geophys. Res. 108 (8), 8383. doi:10.1029/2001JD001510.
Shin, D., Krishnamurti, T., 2003b. Short- to medium-range superensemble precipitation forecasts using satellite products: 2. Probabilistic forecasting. J. Geophys. Res. 108 (8), 8384. doi:10.1029/2001JD001511.
Spall, M.A., Robinson, A.R., 1990. Regional primitive equation studies of the Gulf Stream meander and ring formation region. J. Phys. Oceanogr. 20 (7), 985–1016.
Steppeler, J., Doms, G., Shatter, U., Bitzer, H.W., Gassmann, A., Damrath, U., Gregoric, G., 2003. Meso-gamma scale forecasts using the nonhydrostatic model LM. Meteorol. Atmos. Phys. 82, 75–96.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in single diagram. J. Geophys. Res. 106 (D7), 7183–7192.
Tonani, M., Pinardi, N., Dobricic, S., Pujol, I., Fratianni, C., 2008. A high resolution free surface model of the Mediterranean Sea. Ocean Sci. 4, 1–14.
van Leeuwen, P.J., 2003. A variance-minimizing filter for large-scale applications. Mon. Weather Rev. 131, 2071–2084.
Warner, J.C., Sherwood, C.R., Arango, H.G., Signell, R.P., 2005. Performance of four turbulence closure models implemented using a generic length scale method. Ocean Model. 8, 81–113.
Williford, C.E., Krishnamurti, T., Torres, R.C., Cocke, S., 2003. Real-time multi-model superensemble forecasts of Atlantic tropical systems of 1999. Mon. Weather Rev. 131, 1878–1894.
Yun, W., Stefanova, L., Krishnamurti, T., 2003. Improvement of the multimodel superensemble technique for seasonal forecasts. J. Climate 16 (22), 3834–3840.
Yun, W., Stefanova, L., Mitra, A.K., Kumar, T., Dewar, W., Krishnamurti, T., 2005. A multi-model superensemble algorithm for seasonal climate prediction using DEMETER forecasts. Tellus 57 (3), 280–289. doi:10.1111/j.1600-0870.2005.00131.x.