Improving Accuracy and providing Uncertainty Estimations: Ensemble Methodologies for Ocean Forecasting

Ibrahim Hoteit[1], Eric Chassignet[2], Mike Bell[3]

[1]Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
[2]Center for Ocean-Atmospheric Prediction Studies, Florida State University, United States
[3]MetOffice, Exeter, UK

Correspondence to: Ibrahim Hoteit (ibrahim.hoteit@kaust.edu.sa)

10

Abstract. Ensemble forecasting has emerged as an essential approach for addressing the uncertainties inherent in ocean prediction, offering a probabilistic framework that enhances accuracy of both short-term and long-range forecasts. By more effectively addressing the intrinsic chaotic nature of mesoscale and sub-mesoscale variability, ensemble methods offer critical insights into forecast errors and improve the reliability of predictions. This paper reviews the ensemble methodologies currently used in ocean forecasting, including techniques borrowed from weather prediction like virtual ensembles and Monte Carlo methods. It also explores the latest advancements in ensemble data assimilation, which have been successfully integrated into both ocean general circulation models and operational forecasting systems. These advancements enable more accurate representation of forecast uncertainties (error-of-the-day) by sampling perturbations conditioned on available observations. Despite the progress made, challenges remain in fully realizing the potential of ensemble forecasting, particularly in developing tools for analyzing results and incorporating them into decision-making processes. This paper highlights the crucial role of ensemble forecasting in improving ocean predictions and advocates for its wider adoption in operational systems.

1. Introduction to Ensemble Forecasting

Forecasts of the ocean state generated by numerical models are inherently uncertain owing to the nonlinear chaotic nature and imperfect internal physics of the ocean models, and inevitable uncertainties in their inputs such as initial and boundary conditions, atmospheric forcing, bathymetry, etc. (e.g., Lorenz, 1996; Pinardi et al., 2008; Sandery et al., 2014; Vandenbulcke and Barth, 2015; Kwon et al., 2016; Sanikommu et al., 2020). Thus, the future ocean cannot be completely described by a single forecast model run, and is better described by a set, or ensemble, of forecasts that provides an indication of the range of possible future ocean states and that represents the uncertainty in the forecasts, also known as errors-of-the-day (Houtekamer and Zhang, 2016; Hoteit et al., 2018) (Figure 1).

Ensemble forecasting has increasingly become a key aspect of weather and climate predictions – see Du et al. (2019) for a review – as it provides a basis to communicate forecasts confidence to end users for better decision making. Similarly, it should

become an integral part of ocean forecasts. Ensemble forecasting was indeed proven to provide extended ocean prediction skills compared to deterministic forecasts, especially for extended time-scale predictions (Mullen and Buizza 2002; Ryan et al., 2015). This ensemble probabilistic framework is also needed for short-range forecasting to better describe the intrinsic

35   chaotic nature of the mesoscale and sub-mesoscale variability resolved by the new generation high-resolution ocean models (Thoppil et al., 2021). Information about forecast uncertainty can be used in many ways. For instance, the probabilistic information that ensembles provide are particularly valuable for early warnings of hazardous conditions in the ocean and can be integrated into the decision-making process based on economic values (Richardson, 2000; Du and Deng, 2010). On short timescales, the probabilistic information is useful to trigger the deployment of environment protection measures in the event

40   of an oil spill (Barker et al., 2020), to advise fishermen about the most probable regions of fishing zones, to help coast-guards on the probable areas to focus for search and rescue operations (Melsom et al., 2012), or to advise on path planning for autonomous marine vehicles (e.g., Albarakati et al., 2021), etc. On climate time scales, ensemble forecasting is useful for providing probabilistic information on climate indices such as El Nino and the Indian Ocean Dipole (Schiller et al., 2020).
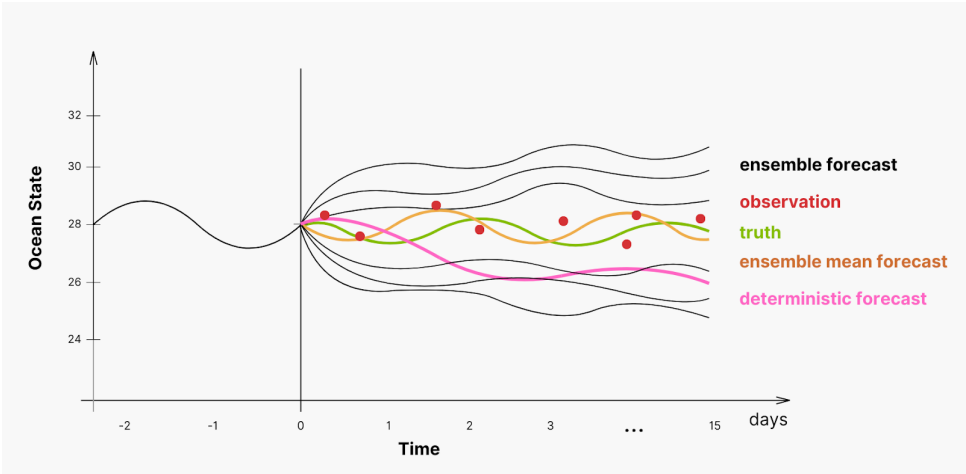

45   **Figure 1: Schematic illustration of deterministic hindcast (black line at the forecast date 0) and forecast (pink line after day 0), and ensemble forecasts (black lines after day 0) of the ocean state. The ensemble forecasts were driven by various sources of uncertainties (including initial conditions, atmospheric forcing, model physics, bathymetry, …). The ensemble forecast mean and the unknown truth are respectively represented by the orange and green lines. Solid red dots denote observations.**
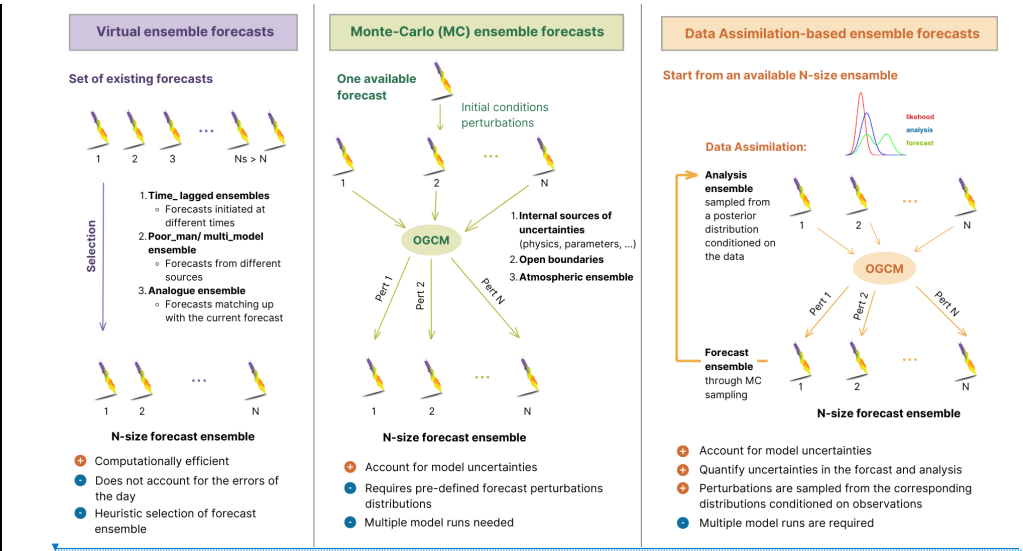
Figure 2: Schematic diagram illustrating the steps involved in various ensemble forecasting methods. Characteristics of each method are also listed.

2. Methods

Ensemble forecasts find their roots in weather forecasting and can be generated (i) as virtual ensembles whose members are selected from deterministic forecasts and/or historical runs, (Hoffman and Kalnay, 1983; Ebert, 2001; Du, 2004; Schwartz and Sobash, 2017), (ii) or by applying some form of Monte Carlo (MC) analysis in which a set of forecasts are produced by perturbing the model physics and/or inputs, as a way to account for their inherent uncertainties (Martin et al., 2015; Houtekamer and Zhang, 2016; Hoteit et al., 2018). Ensemble forecasts may also be generated following a multi-model approach as the forecasts of different ocean models, or from their combination with MC forecasts (Figure 2). Ideally, the actual future oceanic state should fall within the predicted ensemble range.

- *Virtual Ensemble Forecasts.* The lower-cost virtual ensembles can be used to quantitatively estimate forecast uncertainties based on existing forecasts through various techniques including for instance: (a) the time-lagged ensemble, which automatically creates a forecast ensemble by pulling multiple forecasts that have been initiated at different times, (b) the poor-man ensemble, which gathers single-model forecasts from different sources and is thus a multi-model ensemble from existing forecasts, and (c) the analogue ensemble, made of past forecasts matching up with the current forecast. These methods are straightforward but may result in restricted ensembles due to the limited
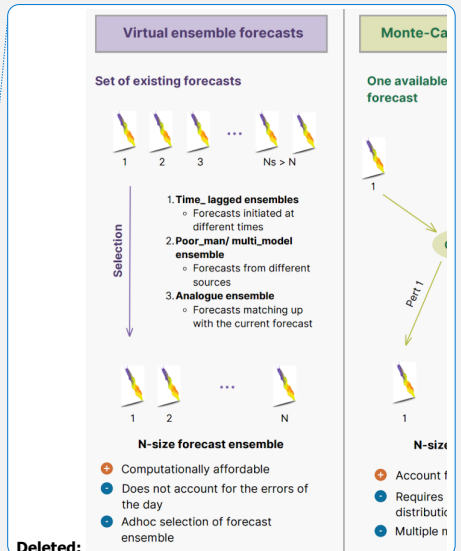
3

available sources of existing forecasts. They are also not designed to capture the flow-dependent error-of-the-day (Du et al., 2019).

- *Monte Carlo (MC) Ensemble Forecasts*. This kind can be generated by perturbing the ocean model physics and/or inputs (Du et al., 2019). Uncertainties in the ocean model could be accounted for by perturbing its internal sources of uncertainties which could come from the missing physics, parameterization schemes, and numerical errors. Different approaches were suggested such as (a) the multi-physics approach which uses a different parameterization scheme for each ensemble member (Sanikommu et al., 2020), (b) the perturbed parameters approach of a selected parameterization scheme, and (c) the stochastic parameterizations approach which injects stochastic perturbations into the physical parameterization schemes (Brankart et al., 2015; Storto and Andrioopoulos, 2021). Additionally, given that the short-term predictability of the atmosphere and the ocean is dominated by their initial conditions (ICs), various methods to perturb the initial model state have been proposed to generate ensembles. These include (i) random perturbations sampled from some available error statistics, (ii) the singular vectors and their variants designed to represent the perturbations with the fastest error growth, and (iii) the vector breeding approach, which computes the initial perturbations as the differences between a pair of past concurrent forecasts. Different approaches were also suggested to perturb the bathymetry, open boundaries and the atmospheric and river forcing (Lima et al., 2019; Storto and Yang 2023; Zedler et al., 2023), but ensembles of atmospheric and oceanic forecasts are now available from the global operational prediction centers and can be readily used to generate ocean forecasts ensembles.

- *Data Assimilation-Based Ensemble Forecasts*. Ensemble forecasts in data assimilation are typically generated by introducing multiple, slightly different estimates of the current system state to capture uncertainties in observations and model parameters while accounting for the "error-of-the-day". For example, in an Ensemble Kalman Filter (EnKF), observations can be perturbed (or not) (Whitaker and Hamill, 2002; Hoteit et al., 2015), and the model is then integrated from these perturbed initial states, sampled according to the estimated initial-state statistics derived from previous forecasts and the most recent observations, resulting in an ensemble of forecasts. Additional perturbations may be introduced to the model physics or inputs to represent other sources of uncertainty, as demonstrated in Monte Carlo (MC) ensemble forecast methods (Whitaker and Hamill, 2012; Hoteit et al., 2018; Sanikommu et al., 2020). This collection of forecasts provides a probabilistic picture of future conditions, reflecting both initial condition and model uncertainties.

Virtual ensemble forecasts were traditionally more common for operational purposes as they do not require major extra computations, although their large ensemble spread was perceived as a disadvantage. The multi-model approach involves the tedious task of running and maintaining different ocean general circulation models (OGCMs), but it can be facilitated by combining the forecasts from different operational centers (e.g., Ren et al., 2019). Ensemble forecasts generated by a MC approach are increasingly adopted operationally. Despite their demonstrated skill, the MC ensemble forecasts require that the ensemble truly represent the probability distribution of the underlying dynamical system (Leith, 1974). Designing perturbation

schemes that accurately capture all sources of uncertainty (e.g., initial conditions, forcing, model physics) remains a significant challenge, as does determining how to vary these perturbations in time.

Recent advances in ensemble data assimilation approaches now provide robust frameworks to represent the error-of-the-day, for both initial conditions and inputs or parameters, by sampling perturbations directly from (approximate) error distributions conditioned on observations (Hoteit et al., 2018; Carrassi et al., 2022). Nevertheless, obstacles persist, particularly in high-dimensional ocean forecasting systems, where the ensemble size is often limited by computational costs. Methods such as localization and inflation are commonly used to mitigate sampling errors and maintain adequate ensemble spread (Brankart et al., 2015; Storto and Andriopoulos, 2021). Hybrid ensemble–variational approaches and other advanced techniques can further alleviate these issues by blending flow-dependent ensemble covariances with multi-year or climatological statistics (Buehner, 2010). However, each solution carries its own computational demands and assumptions, highlighting the ongoing need to balance accuracy, efficiency, and complexity in operational ocean forecasting systems.

3. Probabilistic assessment

Forecast ensembles are evaluated through their sample statistics, mainly the ensemble mean and its spread (the standard deviation with respect to the ensemble mean). The mean can be directly compared with available observations, while the spread indicates the confidence in the forecast: a smaller spread implies lower uncertainty, and vice versa. High-order moments, such as skewness and Kurtosis, help characterize the shape of the ensemble distribution (Groeneveld and Meeden, 1994). In addition, probabilistic validation and verification methods, including reliability, resolution, sharpness and rank histograms, are frequently employed (Johnson and Bowler, 2009). An ensemble is deemed reliable if the predicted probability of an event aligns with the observed frequency. Resolution assesses how far the forecast deviates from the climatological event frequency; increasing this deviation enhances the reliability of the forecast. In the same context, sharpness measures the ability of an ensemble forecast to spread away from the climatological average. Ideally, an ensemble forecast needs to be reliable, with as many forecasts as possible away from the climatological average. Rank histograms, which tally the position of the observation among sorted ensemble values, are used to test reliability and diagnose errors in the ensemble mean or spread (Hamill, 2001). Another commonly used metric is the Continuous Ranked Probability Score (CRPS), which evaluates both accuracy and reliability by comparing the forecast distribution with the observed value across all possible outcomes. A lower CRPS indicates a closer match to reality and thus better overall probabilistic forecasts (Leutbecher and Haiden, 2021).

4. Current status of ensemble forecasts in Operational Ocean Forecasting Systems (OOFSs)

Despite the early establishment of ensemble methods for ocean data assimilation and forecasting (Evensen, 1994), ensemble forecasts, particularly the global systems, only recently found their ways to the operational centers. This is mostly because the centers prioritized using the available computational resources to increase the resolution of ocean models. This was due to the need to resolve the mesoscale to sub-mesoscales processes to better describe the energy cascade in the ocean, and to meet user demands for higher resolution forecasts (e.g. D'addezio et al., 2019; Davidson, 2021). Recent developments in ocean ensemble

forecasting followed the improved coverage in ocean observations that provided increased information to accurately constrain the initial ocean state for extended forecast horizons, the better coordination between ocean forecasting groups, the ease of access to atmospheric ensembles, and the ever-increasing availability of computational power (Metzger et al., 2010; Smith et al., 2011; Strohmaier et al., 2015; Bauer et al., 2021). Ocean ensemble forecasts are now routinely generated at several

195 operational ocean centers on both global and regional scales to cater to different needs as summarized in Table 1.

Table 1: Summary of selected operational ensemble forecasting systems worldwide.

Institution	Forecasting System	Domain (resolution)	Ensemble Perturbations (Size)	Type of Forecast	Reference
Met Office, UK	FOAM	Global (9km)	Observations + internal physics + Atmosphere (36)	Short-range ocean state	**Lea et al. (2022)**
NRL, USA	Navy-ESPC	Global (9km)	Observations (16)	Days to subseasonal ocean state	**Barton et al. (2021)**
Bluelink, Australia	OceanMAPS	Global (10km)	Initial conditions + Time-lagged (48)	Short-range ocean state	**Brassington al (2023)**
ECMWF	NEMO	Global (25km)	Initial conditions + Forcing + Observations (5)	Near Real time ocean state	**Zuo et al. (2019)**
NERSC, Norway	TOPAZ5	North Atlantic and Arctic (6km)	Atmosphere (100)	Short-range ocean state	**Nakanowatari et al (2022)**

6

KAUST, Saudi Arabia	MITgcm	Red Sea (4km)	Atmosphere + Internal physics (50)	Short-range ocean state	**Sanikommu et al. (2020)**
INCOIS, India	ROMS	Indian Ocean (8km)	Atmosphere + Internal physics (80)	Short-range ocean state	**Balaji et al. (2019)**
Bureau of Meteorology, Australia	ACCESS-S	Global (4km)	Internal physics + Time-lagged (30)	Multi-week to seasonal ElNino/IOD	**Wedd et al. (2022)**
CMA, China	CMMEv1	Global (100km)	Multi-model + Initial conditions (90)	Multi-week to seasonal ElNino/IOD	**Ren et al. (2019)**
CMCC	CMCC-SPS3.5	Global (25km)	Initial conditions + Model physics (50)	184 days	**Gualdi et al. (2020)**
ECMWF	SEAS5	Global (25km)	Initial conditions + Model physics + Observations (51)	6 months	**Johnson et al. (2019)**
Meteo-France	Meteo-France System 8	Global (25km)	Model dynamics (51)	7 months	**Pianezze et al. (2022)**

DWD	GCFS 2.1	Global (25km)	Initial conditions + Model physics (50)	215 days	Frohlich et al. (2021)
ECMWF	IFS	Global (10km)	Internal physics (51)	Short-range waves	Browne et al. (2019)
NCEP	GWES	Global (25km)	Wind (30)	Short-range waves	Penny et al. (2015)
UK Metoffice	Wavewatch-III	Atlantic-UK (3km)	Wind (22)	Short-range waves	Bunny and Saulter (2015)
MET-Norway	Barotropic version of ROMS	Norway (4km)	Atmosphere (51)	Short-range storm surge	Kristensen et al. (2022)

5. Role of ensemble forecasts in next generation OOFSs

Recognizing the importance of representing uncertainties in ocean forecasts to meet the need of future demands in probabilistic predictions, ensemble forecasts are expected to become a standard output of any operational ocean product. Although high-resolution observations of some surface variables are now more accessible, the lack of dense, three-dimensional coverage, especially at subsurface levels, still leaves mesoscale and submesoscale processes poorly constrained by ocean analysis systems. Uncertainties from the unconstrained scales might lead to larger forecast errors due to growing dynamical instabilities (Sandery et al., 2017), which limits the forecasting skills of high-resolution ocean models (e.g., Thoppil et al., 2021). Ensemble forecasting has been proven efficient to extend ocean forecasting horizons when model uncertainties in the initial conditions, inputs, and physics are accounted for (Mullen and Buizza 2002; Ryan et al., 2015; Sanikommu et al., 2020). Ensemble forecasts are also essential for providing the error statistics required by ocean analysis systems, thereby enabling better use of high-density observations from recently launched and upcoming satellite missions, such as Surface Water and Ocean Topography (SWOT) (Fu and Ubelmann, 2014). Long delayed by the desire of the community to increase the resolution of the ocean models to improve their realism, the ever-increasing computing resources will provide more and more power to integrate these within ensemble forecasting frameworks.

Ocean forecasts have long been produced by data assimilation (DA) systems and are now routinely used operationally. Ensemble forecasts could be generated from deterministic DA systems, which produce one single-forecast, by simply perturbing the observations (or other parameters of the assimilation system), or during the forecasting step using an ensemble

forecasting method. Ensemble DA methods, on the other hand, readily produce ensemble ocean perturbations that (approximately) represent the error-of-the-day and can be directly used to generate ensemble forecasts. These could be also combined with standard ensemble forecasting methods to further represent the missing information about the error growth in the computationally restricted DA ensembles. To fully exploit the benefits from ocean ensemble forecasts, new tools to analyze and visualize, and also integrate these probabilistic products in decision making and management of ocean services need to be developed and made available for the end users.

References

Albarakati, S., Lima, R., Theußl, T., Hoteit, I., and Knio, O., 2021. Multi-objective risk-aware path planning in uncertain transient currents: an ensemble-based stochastic optimization approach. IEEE Journal of Oceanic Engineering, 46, 1082–1098.

Barker, C. H., Kourafalou, V. H., Beegle-Krause, C., Boufadel, M., Bourassa, M. A., Buschang, S. G., Androulidakis, Y., Chassignet, E. P., Dagestad, K.-F., Danmeier, D. G., et al., 2020. Progress in operational modeling in support of oil spill response. Journal of Marine Science and Engineering, 8(9), 668. https://doi.org/10.3390/jmse8090668.

Barton, N., Metzger, E. J., Reynolds, C. A., Ruston, B., Rowley, C., Smedstad, O. M., Ridout, J. A., Wallcraft, A., Frolov, S., Bauer, P., Dueben, P.D., Hoefler, T., Quintino, T., Schulthess, T.C., and Wedi, N. P., 2021. The digital revolution of Earth-system science. Nature Computational Science, 1, 104–113.

Browne, P. A., De Rosnay, P., Zuo, H., Bennett, A., and Dawson, A., 2019. Weakly coupled ocean–atmosphere data assimilation in the ECMWF NWP system. Remote Sensing, 11(3), 234. https://doi.org/10.3390/rs11030234.

Brankart, J.-M., Candille, G., Garnier, F., Calone, C., Melet, A., Bouttier, P.-A., Brasseur, P., and Verron, J., 2105. A generic approach to explicit simulation of uncertainty in the NEMO ocean model, Geosci. Model Dev., 8, 1285–1297.

Brassington, G. B., Sakov, P., Divakaran, P., Aijaz, S., Sweeney-Van Kinderen, J., Huang, X., and Allen, S., 2023. OceanMAPS v4.0i: A global eddy resolving EnKF ocean forecasting system, in OCEANS 2023-Limerick, IEEE, 1–8, https://doi.org/10.1109/OCEANSLimerick52467.2023.10244383.

Carrassi, A., Bocquet, M., Demaeyer, J., Grudzien, C., Raanes, P., and Vannitsem, S., 2022. Data assimilation for chaotic dynamics. Data Assimilation for Atmospheric. Oceanic and Hydrologic Applications IV, 1–42. https://doi.org/10.48550/arXiv.2010.07063.

D'Addezio, J. M., Smith, S., Jacobs, G. A., Helber, R. W., Rowley, C., Souopgui, I. and Carrier, M. J., 2019. Quantifying wavelengths constrained by simulated SWOT observations in a submesoscale resolving ocean analysis/forecasting system. Ocean Modelling, 135, 40–55.

Davidson, M., 2021. Forecasting coastal evolution on time-scales of days to decades. Coastal Engineering, 168, 103928. https://doi.org/10.1016/j.coastaleng.2021.

Du, J., 2004. Hybrid Ensemble Prediction System: A New Ensembling Approach. Preprints. In Symp. 50th Anniversary Operational Numerical Weather Prediction.

9

Ebert, E. E., 2001. Ability of a poor man's ensemble to predict the probability and distribution of precipitation. Monthly Weather Review, 129(10), 2461–2480.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans, 99(C5), 10143–10162. https://doi.org/10.1029/94JC00572

Fu, L. L., and Ubelmann, C., 2014. On the transition from profile altimeter to swath altimeter for observing global ocean surface topography. Journal of Atmospheric and Oceanic Technology, 31(2), 560–568.

Groeneveld, R. A., and Meeden, G., 1984. Measuring skewness and kurtosis. Journal of the Royal Statistical Society Series D: The Statistician, 33, 391–399.

Hamill, T. M., 2001. Interpretation of rank histograms for verifying ensemble forecasts. Monthly Weather Review, 129(3), 550–560.

Hoffman, R. N., and Kalnay, E., 1983. Lagged average forecasting, an alternative to Monte Carlo forecasting. Tellus A: Dynamic Meteorology and Oceanography, 35(2), 100–118.

Hoteit, I., Pham, D. T., Gharamti, M. E., and Luo, X., 2015. Mitigating observation perturbation sampling errors in the stochastic EnKF. Monthly Weather Review, 143, 2918–2936.

Hoteit, I., Luo, X., Bocquet, M., Kohl, A., and Ait-El-Fquih, B., 2018. Data assimilation in oceanography: Current status and new directions. New Frontiers in Operational Oceanography, Editors: E. P. Chassignet, A. Pascual, J. Tintore, and J. Verron. GODAE Ocean View, 815pp.

Houtekamer, P. L., and Zhang, F., 2016. Review of the ensemble Kalman filter for atmospheric data assimilation. Monthly Weather Review, 144(12), 4489–4532.

Johnson, C. and Bowler, N., 2009. On the reliability and calibration of ensemble forecasts. Monthly Weather Review, 137(5), 1717–1720.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., et al., 2019. SEAS5: the new ECMWF seasonal forecast system. Geoscientific Model Development, 12, 1087–1117.

Kristensen, N. M., Røed, L. P., and Sætra, Ø., 2022. A forecasting and warning system of storm surge events along the Norwegian coast. Environ. Fluid Mech., 23, 307–329.

Kwon, K. M., Choi, B.-J., Lee, S.-H., Kim, Y. H., Seo, G.-H., and Cho, Y.-K., 2016. Effect of model error representation in the Yellow and East China Sea modeling system based on the ensemble Kalman filter. Ocean dynamics, 66, 263–283.

Lea, D. J., While, J., Martin, M. J., Weaver, A., Storto, A., and Chrust, M., 2022. A new global ocean ensemble system at the met office: Assessing the impact of hybrid data assimilation and inflation settings. Quarterly Journal of the Royal Meteorological Society, 148(745), 1996–2030.

Leith, C. E., 1974. Theoretical Skill of Monte Carlo Forecasts. Monthly Weather Review, 102(6), 409–418.

Leutbecher, M., and Haiden, T., 2021. Understanding changes of the continuous ranked probability score using a homogeneous Gaussian approximation. Quarterly Journal of the Royal Meteorological Society, 147, 425–442.

Deleted: Du, J. et al.: Ensemble Methods for Meteorological Predictions. In: Duan, Q., Pappenberger, F., Wood, A., Cloke, H., Schaake, J. (eds) Handbook of Hydrometeorological Ensemble Forecasting. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39925-1_13 https://doi.org/10.1007/978-3-642-39925-1_13, 2019.
Du, J., and Deng, G.: The utility of the transition from deterministic to probabilistic weather forecasts—verification and application of probabilistic forecasts. Meteor. Mon, 36, 10–18. (in Chinese). [...4]

Deleted: .: [...5]

Deleted: https://doi.org/10.1175/1520-

Deleted: .:

Deleted: https://doi.org/10.1029/94JC00572, 1994

Deleted: .

Deleted: . :

Deleted: -568. https://doi.org/10.1175/JTECH-D-13-00109. [...6]

Deleted: .:

Deleted: https://doi.org/10.1175/1520- [...7]

Deleted: .

Deleted: .:

Deleted: monte carlo

Deleted: https://doi.org/10.1111/j.1600-0870.1983.tb00189. [...8]

Deleted: .:

Deleted: , 2018

Deleted: .

Deleted: .:

Deleted: https://doi.org/10.1175/MWR-D-15-0440.1 [...9]

Moved up [6]: C.,

Moved down [7]: G.,

Deleted: Hudson, D., Alves, O., Hendon, H. H., Lim, E.-P [...10]

Deleted: Marshall, A. G., Shi, L., Wang,

Deleted: et al. (2017)..: Access-s1 the new bureau of mete [...11]

Deleted: .:

Deleted: https://doi.org/10.1175/2009MWR2715.1 [...12]

Deleted: . &

Deleted: ...:

Deleted: https://doi.org/10.1007/s10652-022-09871-4, 2022-

Deleted: .:

Deleted: https://doi.org/10.1007/s10236-015-0909-8, 2016.

Deleted: .:

Deleted: https://doi.org/10.1002/qj.4292, 2022.

Deleted: .:

Deleted: https://doi.org/10.1175/1520- [...13]

Lima, L. N., Pezzi, L. P., Penny, S. G., and Tanajura, C. A. S., 2019. An investigation of ocean model uncertainties through ensemble forecast experiments in the Southwest Atlantic Ocean. Journal of Geophysical Research: Oceans 124, 432–452.

Lorenz, E. N, 1996. The essence of chaos. Pure and Applied Geophysics, 147(3), 598–599.

Martin, M. J., Balmaseda, M., Bertino, L., Brasseur, P., Brassington, G., Cummings, J., Fujii, Y., Lea, D., Lellouche, J.-M., Mogensen, K., et al., 2015. Status and future of data assimilation in operational oceanography. Journal of Operational Oceanography, 8(sup1), s28–s48.

Melsom, A., Counillon, F., LaCasce, J. H., and Bertino, L., 2012. Forecasting search areas using ensemble ocean circulation modeling. Ocean Dynamics, 62, 1245–1257.

Metzger, E. J., Hurlburt, H. E., Wallcraft, A.J., Shriver, J. F., Townsend, T. L., Smedstad, O.M., Thoppil, P. G., Franklin, D. S., and Peggion, G., 2010. Validation Test Report for the Global Ocean Forecast System V3. 0-1/12 deg HYCOM-NCODA: Phase II. Naval Research Lab Stennis Space Center MS Oceanography Div.

Mullen, S. L., and Buizza, R., 2002. The impact of horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF Ensemble Prediction System. Wea. Forecasting, 17, 173–191.

Nakanowatari, T., Xie, J., Bertino, L., Matsueda, M., Yamagami, A., and Inoue, J., 2022. Ensemble forecast experiments of summertime sea ice in the Arctic Ocean using the TOPAZ4 ice-ocean data assimilation system. Environmental Research, 209, 112769.

Penny, S. G., Behringer, D. W., Carton, J. A., and Kalnay, E., 2015. A hybrid global ocean data assimilation system at NCEP. Monthly Weather Review, 143(11), 4660–4677.

Pinardi, N., Bonazzi, A., Scoccimarro, E., Dobricic, S., Navarra, A., Ghiselli, A., and Veronesi, P., 2008. Very large ensemble ocean forecasting experiment using the grid computing infrastructure. Bulletin of the American Meteorological Society, 89(6), 799–804

Ren, H.-L., Wu, Y., Bao, Q., Ma, J., Liu, C., Wan, J., Li, Q., Wu, X., Liu, Y., Tian, B., et al., 2019. The China multi-model ensemble prediction system and its application to flood-season prediction in 2018. Journal of Meteorological Research, 33(3), 540–552.

Richardson, D. S., 2000. Skill and relative economic value of the ECMWF ensemble prediction system. Quarterly Journal of the Royal Meteorological Society, 126(563), 649–667.

Ryan, A. G., Regnier, C., Divakaran, P., Spindler, T., Mehra, A., Smith, G. C., et al., 2015. GODAE OceanView Class 4 forecast verification framework: global ocean inter-comparison. Journal of Operational Oceanography, 8(sup1), s98-s111

Sandery, P. A., Sakov, P., and Majewski, L., 2014. The impact of open boundary forcing on forecasting the east australian current using ensemble data assimilation. Ocean Modelling, 84, 1–11.

Sandery, P.A. and Sakov, P., 2017. Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the submesoscale. Nature Communications, 8(1), 1566.

485 Sanikommu, S., Toye, H., Zhan, P., Langodan, S., Krokos, G., Knio, O., and Hoteit, I., 2020. Impact of atmospheric and model physics perturbations on a high-resolution ensemble data assimilation system of the red sea. Journal of Geophysical Research: Oceans, 125(8), e2019JC015611.

Schiller, A., Brassington, G. B., Oke, P., Cahill, M., Divakaran, P., Entel, M., Freeman, J., Griffin, D., Herzfeld, M., Hoeke, R., et al., 2020. Bluelink ocean forecasting Australia: 15 years of operational ocean service delivery with societal, economic

490 and environmental benefits. Journal of Operational Oceanography, 13(1), 1–18.

Schwartz, C. S. and Sobash, R. A., 2017. Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations. Monthly Weather Review, 145(9), 3397–3418.

Smith, S., Cummings, J.A., Rowley, C., Chu, P., Shriver, J., Helber, R., Spence, P., Carroll, S., Smedstad, O.M. and Lunde, B., 2011. Validation Test Report for the Navy Coupled Ocean Data Assimilation 3D Variational Analysis (NCODA-VAR)

495 System, Version 3.43. NRL Memo. Report.

Storto, A., and Andriopoulos, P., 2021. A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. Q. J. R. Meteorol. Soc., 147,1691–1725.

Storto, A., and Yang, C., 2023. Stochastic schemes for the perturbation of the atmospheric boundary conditions in ocean general circulation models. Front. Mar. Sci., 10, 1155803.

500 Strohmaier, E., Meuer, H. W., Dongarra, J., and Simon., H. D., 2015. The top500 list and progress in high-performance computing. Computer, 48, 42–49.

Thoppil, P. G., Frolov, S., Rowley, C. D., Reynolds, C.A., Jacobs, G. A., Metzger, J. E., Hogan, P. J., Barton, N., Wallcraft, A. J., Smedstad, O. M. and Shriver, J. F., 2021. Ensemble forecasting greatly expands the prediction horizon for ocean mesoscale variability. Communications Earth and Environment, 2(1), 89.

505 Vandenbulcke, L., and Barth, A., 2015. A stochastic operational forecasting system of the black sea: Technique and validation. Ocean Modelling, 93, 7–21.

Wedd, R., Alves, O., de Burgh-Day, C., Down, C., Griffiths, M., Hendon, H. H., Hudson, D., et al., 2022. ACCESS-S2: the upgraded Bureau of Meteorology multi-week to seasonal prediction system. Journal of Southern Hemisphere Earth Systems Science, 72, 218–242.

510 Whitaker, J. S., and Hamill, T. M., 2002. Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130, 1913–1924.

Whitaker, J. S., and Hamill, T. M., 2012. Evaluating methods to account for system errors in ensemble data assimilation. Monthly Weather Review, 140, 3078–3089.

Zedler, S. E., Polton, J. A., King R. R., Wakelin S. L., 2023. The effect of uncertain river forcing on the thermohaline properties

515 of the northwest European shelf seas. Ocean Model., 183, Article 102196.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M., 2019. The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: A description of the system and assessment. Ocean Sci., 15, 779–808.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Data and/or code availability

545   This can also be included at a later stage, so no problem to define it for the first submission.

Authors contribution

This can also be included at a later stage, so no problem to define it for the first submission.

Acknowledgements

The authors greatly acknowledge the contributions of Dr. Siva Reddy and Dr. Naila Raboudi to this article and for their

550   invaluable insights.

Deleted: and support throughout the writing process.

13