Assessing the Impact of Different Ocean Analysis Schemes on Oceanic and Underwater Acoustic Predictions

Andrea Storto[1] [iD], **Silvia Falchetti[1], Paolo Oddo[1], Yong-Min Jiang[1,2], and Alessandra Tesei[1]**

[1]Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy, [2]Now at School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada

Abstract Assimilating oceanic observations into prediction systems is an advantageous approach for real-time ocean environment characterization. However, its benefits to underwater acoustic predictions are not trivial due to the nonlinearity and sensitivity of underwater acoustic propagation to small-scale oceanic features. In order to assess the potential of oceanic data assimilation, integrated ocean-acoustic Observing System Simulation Experiments are conducted. Synthetic altimetry and in situ data were assimilated through a variational oceanographic data assimilation system. The predicted sound speed fields are then ingested in a range-dependent acoustic model for transmission loss (TL) predictions. The predicted TLs are analyzed for the purpose of (i) evaluating the contributions of different sources to the uncertainties of oceanic and acoustic forecasts and (ii) comparing the impact of different oceanic analysis schemes on the TL prediction accuracy. Using ensemble member clustering techniques, the contributions of boundary conditions, ocean parameterizations, and geoacoustic characterization to acoustic prediction uncertainties are addressed. Subsequently, the impact of three-dimensional variational (3DVAR), 4DVAR, and hybrid ensemble-3DVAR data assimilation on acoustic TL prediction at two signal frequencies (75 and 2,500 Hz) and different ranges (30 and 60 km) are compared. 3DVAR significantly improves the predicted TL accuracy compared to the control run. Promisingly, 4DVAR and hybrid data assimilation further improve the TL forecasts, the hybrid scheme achieving the highest skill scores for all cases, while being the most computationally intensive scheme. The optimal scheme choice thus depends on requirements on the accuracy and computational constraints. These findings foster developments of coupled data assimilation for operational underwater acoustic propagation.

Plain Language Summary Knowledge of precise maritime conditions is crucial for a variety of human activities, some of which rely on the acoustic characterization of the underwater environment. The benefits of oceanic observing networks on acoustic propagation predictions are intuitive but nontrivial, due to the sensitivity of acoustic fields to small-scale oceanic features. Using an ocean-acoustic modeling framework, this research investigates the benefits of assimilating data from oceanic observing networks on acoustic field characterization. The impact of different methodologies for data assimilation of oceanic observations is evaluated. In particular, the oceanic data generated by three data assimilation schemes with increasing complexity both in algorithm and implementation, along with the ones without synthetic data assimilation, were used as inputs for a predictive acoustic model. The impact of the schemes was then assessed against the "ground truth" at two different signal frequencies and at different ranges. It is found that assimilating oceanic observations significantly improves the acoustic propagation prediction. More sophisticated assimilation schemes further increase the accuracy of sound propagation forecasts but are more computationally intensive. It implies that one needs to balance between the modeling fidelity and computational expenses (in other words, timeliness) for operational applications.

1. Introduction

Environmental characterization of the seas relies on multidisciplinary efforts to observe, analyze, and predict the maritime environment, using together advanced observational technology, numerical models, and data assimilation methods that fuse all sources of information. All these disciplines have continuously improved along time (see, e.g., Pinardi et al., 2017, for a historical perspective). Monitoring and predicting the underwater acoustic conditions, described in terms of the sound propagation characteristics and ambient noise background, are crucial to enhance the ability to operate in the maritime domain for a variety of applications (Robinson & Lee, 1994), for example, ship-noise predictions and sonar based applications.

Sound propagation in water depends on horizontal and vertical seawater density gradients (e.g., Helber et al., 2008). However, it is acknowledged that acoustic models are sensitive to oceanic scales that are much smaller than those resolved by most regional ocean models (e.g., Castor et al., 2004), especially at high frequencies (>1 kHz). Thus, it is nontrivial to understand whether the corrections made by ocean data assimilation, while certainly improving the accuracy of the ocean predictions, are also able to improve the skills of the coupled acoustic propagation simulations.

During the last decades, a number of coupled ocean-acoustic operational prediction systems have been deployed (e.g., Lam et al., 2009; Lermusiaux et al., 2010; Wang et al., 2009), generally relying on the assimilation of oceanic observing networks. All these studies suggest and recommend, to a different extent, that the ocean variability induced by data assimilation procedures affects the underwater sound propagation prediction and provides an unambiguous motivation to further develop state-of-the-art ocean-acoustic prediction systems. Furthermore, the observed and modeled coupling between oceanic and acoustic processes has the great potential of exploiting acoustic observing networks and campaigns to improve the ocean state estimation, through ingesting such observations in coupled analysis schemes by means of inverse methods (e.g., Hursky et al., 2004). Yet the work to couple the oceanic and acoustic data assimilation systems is relatively new and explored only at a theoretical level (e.g., Ngodock et al., 2017).

Data assimilation techniques have evolved notably during the last decades, especially thanks to the increased computer power that made possible to use four-dimensional analysis schemes and/or introduce flow-dependent background-error characterization through ensemble analysis schemes. In the last 10 years or so, hybrid ensemble-variational schemes, where the robustness of variational schemes is combined to ensemble-derived flow-dependent background errors (Bishop & Satterfield, 2013; Hamill & Snyder, 2000), are the most advanced schemes implemented operationally in meteorological centers (e.g., at the UK Met Office; Clayton et al., 2013). Simplifications are often needed to allow reasonable computational demand of the analysis schemes (e.g., Storto et al., 2018).

Constructing hybrid ensemble-variational systems requires optimal ensemble generation strategies, where all the significant sources of uncertainty are accounted for. For instance, stochastic physics schemes that have been often neglected in oceanic applications have recently attracted the attention of the ocean prediction community for their relative simplicity in capturing flow-dependent error structures (Brankart et al., 2015). This is testified by their growing use in ocean models, probabilistic simulations, and air-sea coupled prediction systems (e.g., Bessières et al., 2017; Brankart, 2013; Juricke et al., 2017), although the latter studies generally focus on low-frequency variability and interannual time scales. Consequently, we devote efforts to the implementation of a new stochastic physics package for use in the hybrid ensemble-variational prediction system, thus introducing stochastic physics also in short-term oceanic prediction systems (see Appendix B for details). Ensemble approaches, as a limited-size application of Monte Carlo simulations, may also provide a useful assessment of sources of uncertainty for integrated ocean-acoustic systems (e.g., Dosso et al., 2007). This provides further motivation for developing such systems, as ranking uncertainty sources is of obvious importance to detect the reliability of the analysis system within rapid environmental assessments (Pecknold & Osler, 2012).

A way to assess the impact of data assimilation is through performing Observing System Simulation Experiments (OSSEs), which were introduced in meteorology since long (see, e.g., Arnold & Dey, 1986; Atlas, 1997, for a review of methodological concepts and applications) and recently extended to optimal planning of glider missions (Alvarez & Mourre, 2012) or impact evaluation of ocean observing networks (Halliwell et al., 2015). OSSEs are a rigorous methodology to assess the impact of current and future observing networks through design of synthetic observations and assessment of observation impact in a way similar to Observing System Experiments, where a certain observation type is withheld from the system and the decrease of accuracy in such experiment provides a measure of the impact of the observation type itself. The way the "truth" ("nature run" [NR]) is chosen and, accordingly, the way synthetic observations are drawn are crucial elements of OSSE, which generally rely on the use of an alternative model (e.g., Errico et al., 2013) or the same numerical model but with different configurations (Halliwell et al., 2014). In this work, OSSEs are used to evaluate the impact of different data assimilation schemes rather than assessing different observation types, in a way similar to the works of Ngodock and Carrier (2014) and Penny et al. (2015) for comparing different analysis schemes in coastal and global ocean data assimilation systems, respectively.
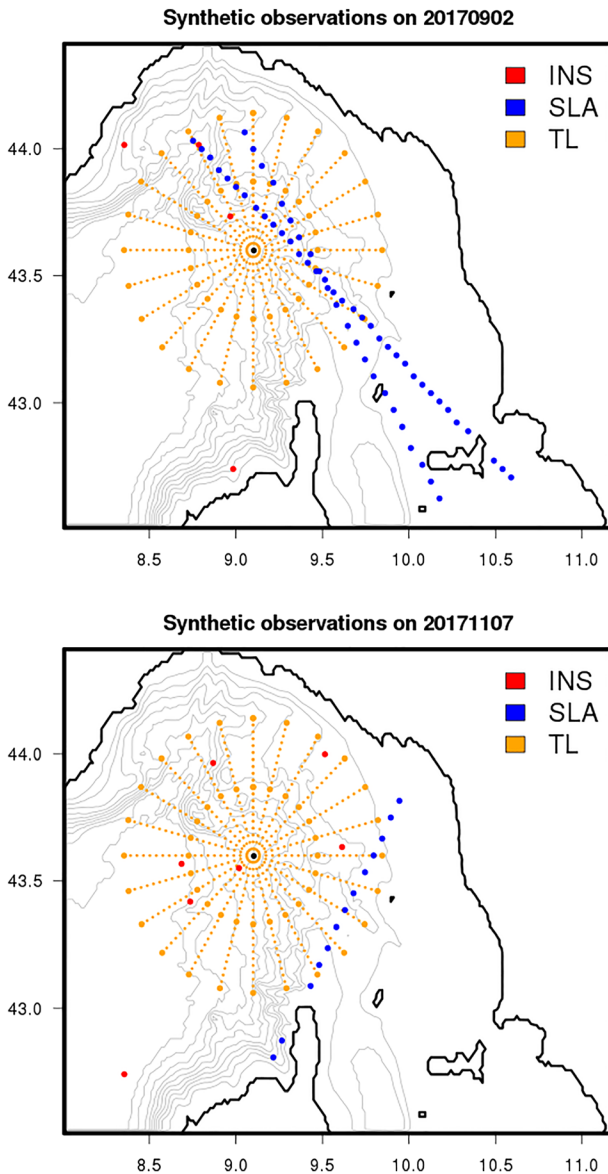
Figure 1. Examples of synthetic observations extracted from the nature run for two daily assimilation cycles during the study period. Blue and red dots identify altimetry data and in situ profiles, respectively. Orange dots indicate the geometry of the idealized acoustic experiment and synthetic observations: The small dots indicate the propagation paths, big dots the 30 and 60 km range from the sound source (black dots) located in the center (9.1°E; 43.6°N), and paths imposed radially with bearing angles every 15° and 60 km range. Gray contours indicate the bathymetry, with 250 m isobaths. The plotted area corresponds to the regional ocean model domain.

Here, we use the OSSE framework to evaluate the impact of ocean data assimilation and complexity of the schemes for ingesting oceanic observations and improving one-way coupled acoustic predictions, in particular evaluating their impact in the space of the transmission loss (TL) synthetic observations. This work represents the first step toward fully coupled oceanic-acoustic data assimilation, whereas acoustic observations are not yet assimilated here. Later, this kind of exercise is expected to be repeated with real oceanographic and acoustic data. The article is structured as follows: Section 2 introduces numerical models, ensemble systems, and OSSE configuration. Section 3 presents results regarding the main sources of uncertainties in both the oceanic and acoustic domains. Section 4 shows the impact of the oceanic data assimilation on the acoustic simulations. Section 5 discusses and concludes.

2. Data and Methods

In this section, the ocean-acoustic modeling system and the oceanographic data assimilation scheme are briefly introduced, and then the ensemble system and the ensemble generation strategy are described. Finally, the OSSE methodology is recalled, and its implementation in this research is specified.

2.1. The Oceanographic Prediction System

The ocean modeling area (Figure 1) is a portion of the Long-Term Glider Mission for Environmental Characterization 2017 (LOGMEC17) sea trial area of operations and extends from 8.24°E to 9.66°E and from 43.64°N to 44.46°N. LOGMEC17 is an acoustical and oceanographic campaign aimed to study the variability and predictability of the oceanographic and acoustic environments on different temporal and spatial scales. The sea trial was conducted in the Ligurian Sea from 14 September to 14 November 2017.

The analysis and forecast system implemented along with LOGMEC17 sea trial is based on a regional implementation of the Nucleus for European Modelling of the Ocean (NEMO; Madec & the NEMO team, 2012) primitive equation model, Version 3.6. The NEMO domain covers the Ligurian Sea with two open boundaries at the western and southern sides of the domain, located at 8°E and 42.5°N, respectively. The horizontal resolution of the model is about 1.8 km, and the vertical grid is discretized using 91 depth levels with partial steps. Surface fluxes are calculated iteratively every model time step (hundreds) by means of bulk formulas (Pettenuzzo et al., 2010) using the air temperature, humidity, winds, and cloud cover from the atmospheric operational analyses provided of the National Centers for Environmental Predictions (NCEP) at 1/4° horizontal resolution.

The NEMO model is one-way nested with the Copernicus Marine Environment Monitoring Service (CMEMS) Mediterranean Sea operational model (Oddo et al., 2014), also called Mediterranean Forecasting System (MFS). The sea surface temperature (SST) from the regional CMEMS near-real-time data set produced from satellite data (Buongiorno Nardelli et al., 2013) is assimilated using a nudging scheme by which the SST information is ingested into the model through the adjustments of air-sea net heat fluxes.

The formulation of the data assimilation scheme is detailed in Appendix A, while the operational configuration is described in Storto and Oddo (2019) and Storto et al. (2019). Here we provide only a concise description. The variational scheme is based on the classical incremental formulation with the control variable

transformation (Courtier, 1997). Background-error covariances are modeled through a sequence of operators that account for vertical covariances through multivariate Empirical Orthogonal Functions (EOFs) and horizontal correlations through the application of a first-order recursive filter. EOFs were calculated from a data set of anomalies from a control run experiment (see Appendix A for details). Length scales of horizontal background-error correlations were defined homogenously (from about 6 to 12 km and increasing with depth, decreasing to 1 km in the proximity of the coasts and domain boundaries). The balance operator partly follows the formulation of Weaver et al. (2005), namely, the sea level is given by the sum of a balanced component, using the dynamic height formula, and an unbalanced component, which is derived from the multivariate EOFs (from temperature, salinity, and unbalanced sea level, the latter being the residuals of the balance). In situ observation error profiles were scaled from misfit statistics collected during a preliminary control run simulation.

In order to achieve the variational solution, minimization of the cost function (see Appendix A, Equation A2) is achieved through the L-BFGS minimizer (Byrd et al., 1995), which is an efficient limited-memory quasi-Newton algorithm for large optimization problems. Convergence of the minimizer is reached when the infinity norm of the cost function gradient (maximum of the absolute value of the gradient components) is smaller than 3% of the initial cost function gradient. In the current setting, the minimizer employs on the average 17 iterations to converge, reducing the initial cost function by about five times on the average. Constrained by the synthetic glider and altimetry observations (see section 2.4.2), the variational minimization produces three-dimensional analyses of temperature, salinity, and sea surface height.

Both sea level anomaly and in situ data undergo a background-quality check, namely, observations with a too large departure from the background, that is, 3 times the square root sum of observational and background error variances, are rejected. Additionally, a variational quality control scheme is applied to in situ profile data, in order to reject as few observations as possible, while giving small weight in the analyses to the observations with large departures (Storto, 2016). Note that, although the prediction system is used here in an OSSE context, all the features proper of the real-time system (i.e., observation preprocessing and error characterization; see Storto et al., 2019) are retained, in order to mimic an operational environment.

2.2. The Acoustic Propagation Model Coupled to the Oceanic Predictions

The acoustic simulations were carried out using the acoustic propagation modeling in use at CMRE. The propagation modeling engine of the acoustic prediction system is Range-dependent Acoustic Model (RAM; Collins, 1993). RAM is a 2-D range-dependent (with respect to bathymetry and sound speed profile) acoustic model that uses split-step Padé solution for the parabolic equation (PE) method (Collins, 1994). It assumes a flat sea surface. The main advantage of this model is that it is as accurate as higher-order PE solutions and approximately 2 orders of magnitude faster than other finite-difference solutions of the PE (Collins et al., 1996). RAM is especially appropriate for low-frequency sound propagation in the spatially varying ocean environment such as the one that this study focuses on. The wave equation solved by RAM is derived in cylindrical coordinates. The bathymetry used with both NEMO and RAM was extracted from the General Bathymetric Chart of the Oceans database (Weatherall et al., 2015).

A set of nominal values was used to represent the seabed geoacoustic properties in the area of the experiment. The geoacoustic properties of the seabed were adopted from the NOAA Deck41 database (Bershad & Weiss, 1976) for this area, that is, with specified compressional wave sound speed (c_p = 1,500 m/s), attenuation (α = 0.1 dB/λ), and density (ρ = 1,500 kg/m^3), corresponding to the clayey silt sediment type (p. 116, Lurton, 2010). The noise source information, the bathymetry, and the sound speed profiles predicted by the ocean model are passed to the range-dependent acoustic model RAM to compute the TL from the source to the receiver at defined frequencies. In particular, we investigate two configurations, two sources at 10 m of depth and frequencies equal to 75 Hz, representative of ship noise propagation, and 2,500 Hz, representative of active sonar applications. For both sources, the receiver depth is fixed at 60 m (see also section 2.4 for further details). Temperature and salinity fields from NEMO outputs are converted to sound speed fields and used by RAM at the nominal NEMO resolution. Thus, we consider RAM one-way coupled to NEMO (i.e., no feedback from the acoustic underwater propagation model to NEMO).

2.3. The Ensemble System

We introduce here the ensemble variational data assimilation system (EDA), which is built upon the ocean-acoustic prediction system described above, with a twofold aim of (i) investigating and ranking the sources of uncertainty in ocean-acoustic predictions and (ii) providing a reliable ensemble spread to feed the flow-dependent component of background-error covariances in the hybrid ensemble-variational data assimilation system. The ensemble system is based on independent three-dimensional variational (3DVAR) realizations with perturbed data sets. These realizations are used in the hybrid ensemble-variational data assimilation experiments (see Appendix A) to estimate the time-varying parameters (EOFs and correlation length scales) used to build the flow-dependent ensemble-derived component of the background error covariance matrix.

The generation of the ensemble is a crucial task in order to account for the main sources of uncertainty that underpin the ensemble spread and consequently for the uncertainties spanned by the hybrid variational-ensemble system. Indeed, all the sources of the uncertainty shall be accounted for, so that the ensemble system realistically mimics the true error evolution, including the uncertainty linked to boundary conditions, forcing fields, and input data sets (Storto & Randriamampianina, 2010). The generation of our EDA system includes six different sources of uncertainties, which are briefly discussed below.

2.3.1. Perturbation of the Atmospheric Forecast Data

In order to account for uncertainty in the atmospheric data used to force the ocean simulations, we use either European Centre for Medium-Range Weather Forecasts (ECMWF) or NCEP operational analysis and forecast fields.

2.3.2. Perturbation of the Bulk Formulas

The uncertainty in the air-sea flux exchanges (i.e., transfer coefficients) is accounted for through using either the Coordinated Ocean Reference Experiments bulk formulas (Large & Yeager, 2004) or the MFS bulk formulas (Pettenuzzo et al., 2010).

2.3.3. Perturbation of the LBCs

Lateral boundary conditions (LBCs) uncertainty is accounted for through the use of two state-of-the-art operational oceanographic analyses and forecasts to provide the LBCs at the southern and western boundaries of the Ligurian Sea computational domain: the CMEMS MED-MFC analyses (at 1/24° of horizontal resolution; Oddo et al., 2014) and the CMEMS GLO-MFC (at 1/12° of horizontal resolution; Lellouche et al., 2018).

2.3.4. Perturbation of the SST Analysis Data Set

Similarly, the uncertainty in the SST daily analyses used within the surface nudging scheme is imposed through combining either the CMEMS MED SST regional product (from CNR/Italy; Buongiorno Nardelli et al., 2013) or the CMEMS GLO SST global product (OSTIA, from UK Met Office; Donlon et al., 2011).

2.3.5. Perturbation of the Ocean Model Physics

Subgrid-scale parameterizations have been perturbed by means of a novel implementation in the NEMO ocean model of the Stochastically Perturbed Parameterization Tendency (SPPT) scheme, which was originally developed for medium- to long-range atmospheric prediction systems (Buizza et al., 1999). Formulation and details of the scheme are reported in Appendix B. Here, we summarize that the scheme produces perturbations collinear to the tendencies and bounded in order not to reverse the sign of the tendency itself. In the current configuration, horizontal and vertical diffusion and solar penetration tendencies are perturbed.

2.3.6. Perturbation of the Oceanic Observations

Observations are perturbed by adding a random realization to each observation, with variance equal to the observational error variance of each observation (e.g., Burgers et al., 1998). The observation perturbations are assumed uncorrelated in space and time as usually adopted in ensemble simulations, consistently with the assumption of uncorrelated observational errors in the data assimilation scheme. In the future, correlations between, for example, along-track altimetry data or in situ measurements from the same profiles might be introduced to increase the realism of the perturbations.

While the perturbation of oceanic observations is added to each ensemble member, the other sources of uncertainties are combined together. Two sources of atmospheric forcing, times two types of bulk formulas, times two sets of LBCs, times two SST data sets provide $2^4 = 16$ realizations. We have used three realizations of the stochastic physics and a set of members with unperturbed ocean model physics. Combined together,

these form an ensemble of 64 members (48 with perturbed physics and 16 without), which is the basis of the EDA system. These members use stationary background-error covariances (the same ones as in the standard 3DVAR experiment introduced later), as they are run off-line and independently from each other for the sake of simplicity and computational resources. In the future, online EDA systems that include recentering of the ensemble mean to avoid systematic offset between the EDA ensemble and deterministic systems will be investigated, as in Lei and Hacker (2015) and Houtekamer et al. (2018).

Note that when running the deterministic EDA member with hybrid background-error covariances, the nominal set of input data is used (described in section 2.1, i.e., regional high-resolution products for SST and LBC, NCEP atmospheric forcing, bulk formulas from Pettenuzzo et al., 2010, and unperturbed physics and observations).

2.4. Design of the OSSEs

We have adapted the OSSE methodology to our goal, that is, to study the impact of different physical ocean analysis schemes to the coupled acoustic fields. Following the suggestions of Hoffman and Atlas (2016) for coupled OSSEs, we use a coupled ocean-acoustic NR to extract both oceanographic and acoustic observations from the two model components. Oceanographic observations are then assimilated into the variational data assimilation system, and the one-way coupled TL maps are finally validated against the synthetic acoustic observations, previously extracted from the NR. Doing this, we are able to project the impact of the different analysis scheme onto the TL results. Note that the acoustic observations are indeed used only for validation in the present study.

2.4.1. Choice of the NR

Choosing a NR whose error characteristics resemble the typical system accuracy with respect to a nonassimilative run is a delicate task. A popular methodology, followed in ocean applications by Halliwell et al. (2014), consists of using a modified version of the ocean model (different parameterizations and numerical schemes) to form the NR. Here, we follow a similar approach, where however rather than changing the numerical schemes of the OGCM, we use different input data sets. In particular, the NR has different lateral and boundary condition data (taken from CMEMS GLO MFS and ECMWF, respectively) compared to the analysis setup used in the OSSE experiments (with LBC and surface boundary condition [SBC] from CMEMS MED MFC and NCEP, respectively). Moreover, the stochastic physics scheme previously introduced for the ensemble generation is applied in order to provide a completely different realization of the NR. In practice, the NR is a member of the ensemble system previously discussed, while OSSEs are performed with a configuration from another member of the same ensemble. Being one out of 64 members, the dependence between the NR and the EDA system is assumed to be negligible.

The choice of the NR has been checked against the Halliwell et al. (2014) criteria, namely, the statistics of differences between the NR and control OSSE experiment (CTR, without synthetic data assimilation) resemble the statistics of the OSSE. For instance, we found for sea level that the mean squared difference between the NR and CTR is equal to 1.6 cm, while statistics of a data assimilation experiment versus the control run are equal to 1.8–2.0 cm in real observation space (depending on the specific configuration, not shown).

2.4.2. Synthetic Oceanic Observations

Synthetic oceanic observations drawn from in situ profiles of temperature and salinity and altimetry have been produced for each day of the study period (1 August to 15 November 2017) by randomly selecting a number of profiles and altimetry tracks and randomly choosing their location and time. Vertical sampling of in situ profiles follows the actual vertical resolution of glider profiles during the LOGMEC17 campaign; similarly, altimetry data belonging to each track follow the 7 km spatial resolution of actual L2 satellite altimetry products (e.g., from CMEMS), although their orientation may be different from actual satellite orbit tracks. In particular, the daily number of in situ profiles and altimetry tracks varies from 0 to 12 (6 on average) and from 0 to 4 (2 on average), respectively, and two random uniform distributions are chosen to sample the daily number of profiles and tracks. These numbers were chosen in order to be close on average to the observational sampling during the LOGMEC17 observational campaign, although the day-by-day sampling might be different from the original one. Examples for the observation scenarios on 2 September and 7 November 2017 are shown in Figure 1. In such examples, two or one altimetry tracks, respectively, occurred for each day, while 4 or 12 profiles were measured, indicating the diverse observation sampling among different days of simulations. The observations are obtained from the NR fields spatially interpolated onto

values at observation location, plus a value perturbation that accounts for the observational errors and is generated through a Gaussian random deviate whose variance is equal to the observational error variance used in the data assimilation scheme. Observational errors were retuned during preliminary experiments and chosen from assimilation output diagnostics (Storto et al., 2019). Quality checks introduced in section 2.1 are not relevant for these OSSE applications due to the synthetic nature of the observations, although occasionally the random error added to the NR may increase the observation misfit and lead to rejection.

2.4.3. Synthetic Acoustic Observations

In order to validate the ocean-acoustic experiments, a number of synthetic acoustic observations were extracted through coupling the RAM acoustic model with the NR previously introduced. TLs at 30 and 60 km ranges were simulated from the NR for a source located at 9.1°E, 43.6°N, along radial paths at 15° interval. Such idealized configuration allows us to assess the sound propagation which is representative of the entire area of study. The geometry of the synthetic observations is shown in the panels of Figure 1 (orange dots). In particular, the different paths cover several scenarios, for example, propagation toward deep areas (westward) or toward shallow areas (eastward and southward) and with an abrupt decrease of bathymetry (southward, toward off northern Corsica). Similarly, the sections exhibit different circulation regimes and modeling challenges, linked for instance to the vicinity of LBCs (eastern sections) or coastal processes. The choice of two different ranges for the TL evaluation responds also to the possibility of partly disentangling boundary condition versus bathymetric influence on the acoustic propagation.

As mentioned earlier, two acoustic scenarios are considered, a 75 Hz signal with the source at a depth of 10 m, which is representative of ship noise sound propagation, and a 2,500 Hz signal with the source at 10 m of depth, representative of active sonars. For both scenarios, sound speed fields from the oceanic OSSEs were used by RAM, and the resulting acoustic TL is compared to the synthetic acoustic observations. Note that, since acoustic observations are not assimilated, the synthetic TL data are not perturbed as they represent the truth in the TL space. The receiver depth is chosen at 60 m depth for both scenarios, which is generally located just below the thermocline, thus limiting any spurious impact of the highly variable near-surface ocean. In the future, more acoustic scenarios and sound-receiver geometries can be considered.

3. Error Budget Analyses From the Ensemble System

In this section, we review and rank the effect of different sources of uncertainty (viz., perturbation types) on the resulting ensemble spread, for both the oceanic and acoustic ensemble systems, after having discussed the reliability of the ensemble. In particular, we use the oceanic ensemble system (section 2.3) that will be later used in the hybrid ensemble-variational scheme, to infer information on the sources of uncertainty. Acoustic simulations were one-way coupled to each member of the oceanic ensemble system. As limited-size Monte Carlo simulations, ensemble prediction systems provide indications on the sensitivity of uncertainties (Pecknold & Osler, 2012). The main diagnostics we rely on is the "explained ensemble variance," which is built by clustering the different ensemble members depending on the perturbation used. In particular, in order to assess the effect of a certain uncertainty/perturbation component (e.g., the LBCs), the diagnostics represents the percentage with which the subsampled ensemble variance decreases, with respect to the full-size ensemble variance (e.g., $EV = 100 \, (ES_t - ES_i)/ES_t$, where ES_t refers to the full-size ensemble variance and ES_i to the ith subsample ensemble variance). The resulting diagnostics is the arithmetic average of the quantity of interest, in the LBC example over the two LBC subsamples; their dispersion, in turn, quantifies the confidence in the diagnostics. In this way, the reduction of the ensemble dispersion due to the withholding of a certain perturbation quantifies the impact that such perturbation has in the ensemble system.

3.1. Ensemble Reliability

Ensemble reliability needs to be assessed prior to utilizing the ensemble spread for the diagnostics presented later. A common way to assess the ensemble spread is through comparison with the root mean square error (RMSE) of the ensemble mean versus observations, in the observation space. Ideally, the spread should match the error of the ensemble mean in terms, for example, of mean values and temporal variability (see Fortin et al., 2014; Hopson, 2014). Unbiased independent observations should be used for such type of validation (Rodwell et al., 2016). Here, we assess the spread-RMSE relationship against the observations before

their ingestion in the prediction system, namely, they can be reasonably assumed to be independent if one neglects the temporal correlation of the observational errors.

Results of this assessment are summarized in Figure 2, which shows the spread-RMSE relationship from daily assimilation time windows, separately for temperature and salinity within different vertical regions. Correlation values identify their degree of temporal agreements, while the ratio of standard deviations quantifies the different amplitude of their temporal variations. It turns out that correlations are all significant (at 99% confidence level), which indicates that the spread-error relationship is well-posed. The ratio for temperature is between 1.1 and 1.6, suggesting a slight underdispersion of the ensemble spread with respect to the ensemble mean error. The underdispersion for salinity is larger near the surface (2.2), although correlations are high. The spread is underestimated for particularly large error events (i.e., visible as outliers in the panels of Figure 2).

3.2. Spread Diagnostics in the Oceanographic System

Explained ensemble variances for the different perturbation types and several oceanic parameters (mixed layer depth, sea surface height, and temperature and salinity at different vertical levels) are shown in Figure 3, together with their associated range of confidence. These diagnostics are computed for a 1 month period from mid-October to mid-November, and uncertainty sources are ranked in the panels in decreasing order of explained ensemble variance. The perturbation of LBC is the most impacting perturbation for the entire set of parameters investigated. In particular, LBC perturbations dominate the sea surface height diagnostics, due to the effect of the barotropic transports on sea level. LBCs also dominate the explained ensemble variance for the salinity at both 100 and 800 m water depth. However, mixed layer depth and, consistently, temperature at 100 m of depth are also impacted by the stochastic physics (OSTO). At the sea surface, for both temperature and salinity, perturbation of SBCs becomes the second most impacting source of perturbation, and in particular, the perturbation of the bulk formulas impacts more than the atmospheric data set (both shown in the panels). Small impact is provided by the SST data set perturbation. At depth (800 m), the impact of LBC perturbations increases significantly, likely due to the different vertical discretization and physics of the two LBC data sets, which in turn amplify the differences in the implied bottom currents in the limited area model experiments. To conclude, while LBC perturbations generally represent the largest source of uncertainty, the model subgrid physics (i.e., the stochastic physics scheme perturbing horizontal and vertical mixing and shortwave radiation penetration) provide the second most important source of uncertainty between the mixed layer depth and thermocline, while at the sea surface, SBC perturbation is the second most important one.

Figure 4 shows the explained ensemble variance of the range-averaged mixed layer depth and SST for the entire period, as a function of the cross section identified by the bearing angle with respect to the acoustic source, consistently with synthetic acoustic observation configuration introduced in section 2.4 (Figure 1). The mixed layer depth is a fundamental quantity for underwater sound propagation, roughly coinciding with the sonic layer depth (viz., the depth where the sound speed is maximum, which characterizes the acoustic ducts) at midlatitudes (see, e.g., Helber et al., 2008; Sutton et al., 1993). The plots show only the largest three sources of uncertainty for sake of clarity, that is, LBC, OSTO, and SBC. Interestingly, these three sources of uncertainty exhibit a geographically distinct behavior for the mixed layer depth explained variance, where LBC shows the largest impact to the west, OSTO northwest, and SBC south of the source, consistently with the open ocean boundaries of the regional domain. The SST explained ensemble variance provides a rather similar picture, with the LBC perturbation dominating toward all bearings except northeast where OSTO prevails, likely due to the enhanced effect of stochastic physics in shallow areas. For temperature in deeper regions (below 50 m of depth, not shown), the LBC explained ensemble variance is much higher than any other perturbation for all bearings.

Such a sketch underpins the need for implementing multiple sources of uncertainty in an operational oceanographic ensemble system to span different uncertainties that may impact differently depending on the vertical region and specific location.

3.3. Spread Diagnostics in the Acoustic System

The assessment and ranking of prediction uncertainty sources are repeated here for the one-way coupled acoustic simulations, in particular for the TL at 75 Hz and 30 or 60 km range from the sound source. For
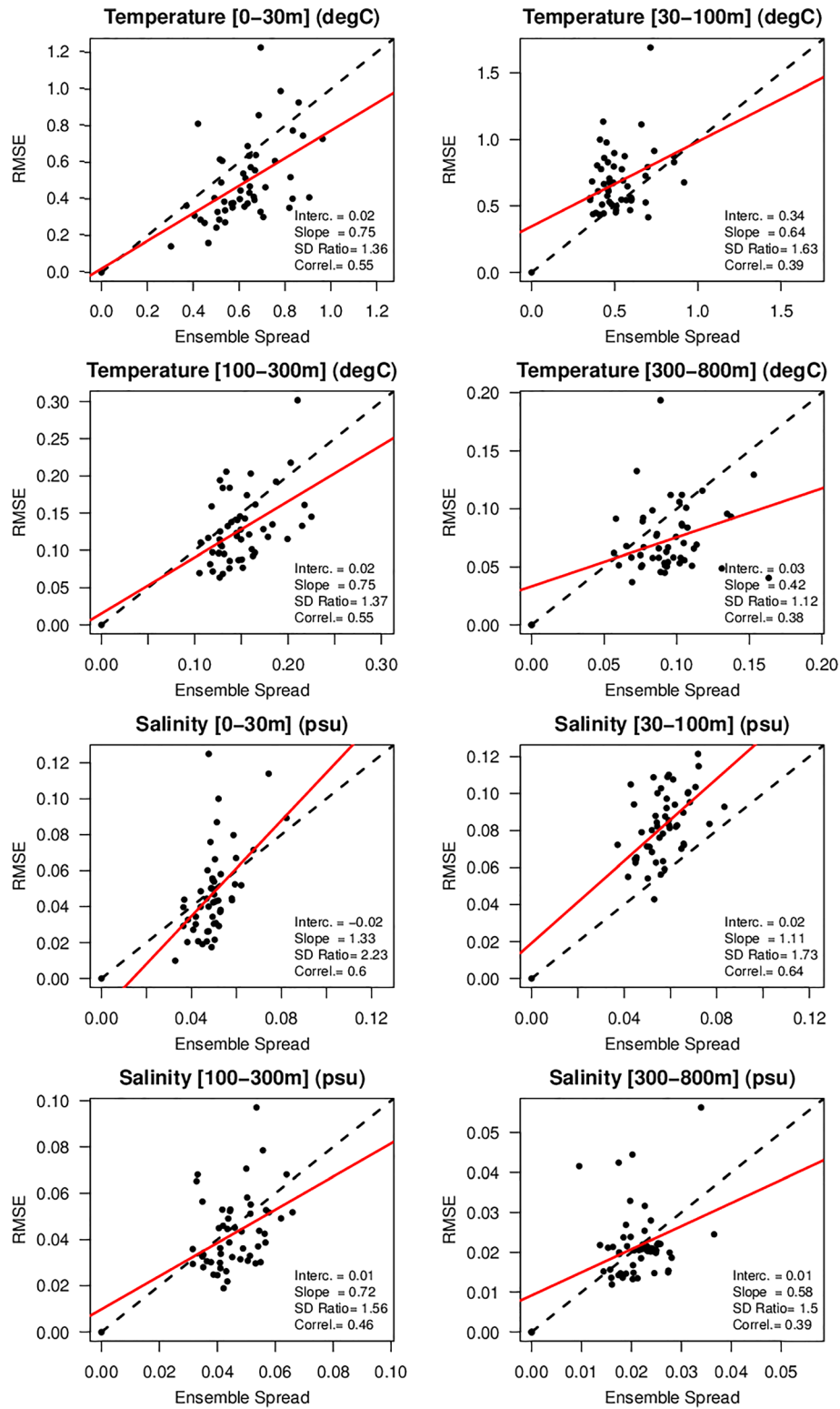
Figure 2. Scatter plot of the ensemble spread versus ensemble mean RMSE in observation space. Dots correspond to RMSE and spread pairs for daily time windows, for temperature and salinity in several vertical regions. The interpolating line is shown in red, while the perfect case line is dashed. The panels report also the intercept and slope of the interpolating line, the ratio between the standard deviations of the two time series (SD ratio), and their correlation (Correl.). Correlations are statistically significant for values greater than 0.36 (based on *t* test at 99% confidence level).
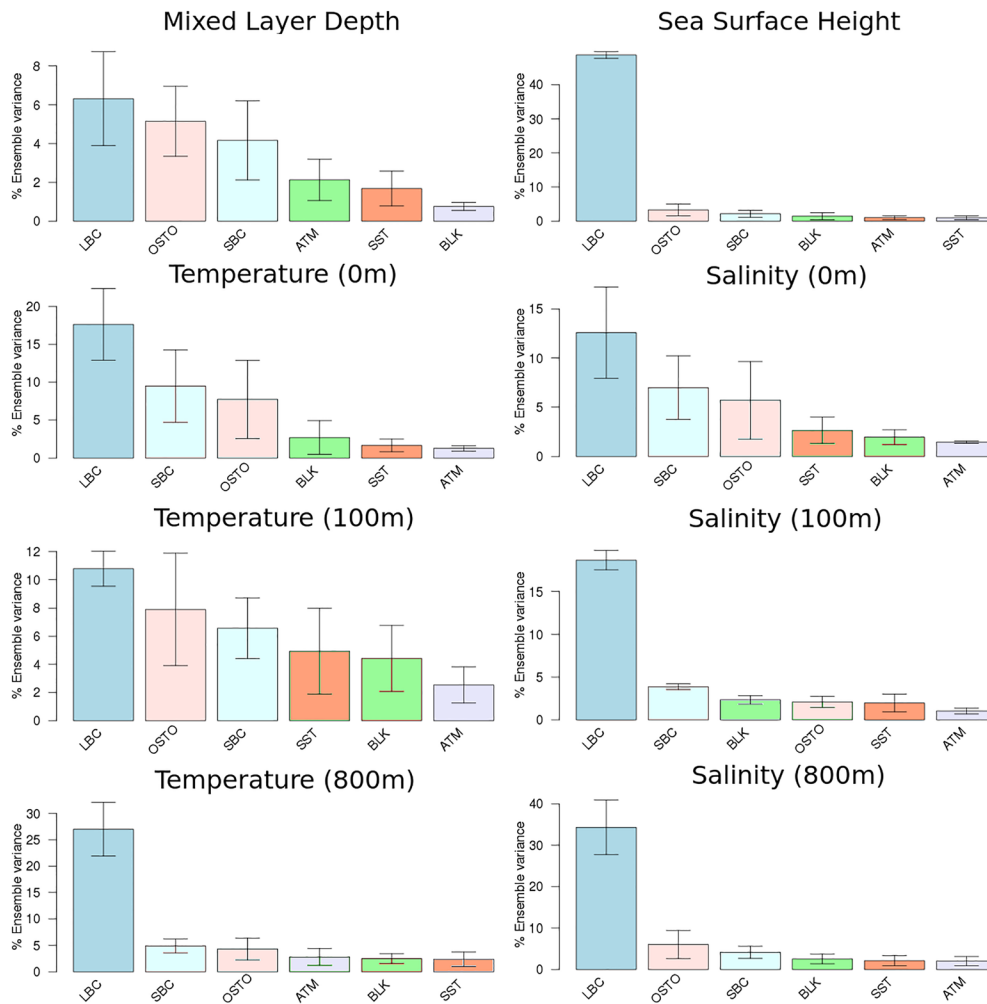
Figure 3. Explained ensemble variance and range of confidence for different perturbation types and ocean parameters, sorted in decreasing order. Note that SBC merges ATM and BLK perturbations. The thin bars indicate the range of confidence of the diagnostics. LBC = lateral boundary condition perturbations; OSTO = ocean stochastic physics; SBC = surface boundary condition perturbations; ATM = atmospheric input data perturbation; SST = SST data perturbation; BLK = bulk formulas perturbation.

the time being, no other frequencies are investigated in terms of ensemble spread, due to the massive computational resources required to run the 1 month ensemble simulations (being the computational costs approximately proportional to the frequency investigated). In order to compare the oceanographic sources of uncertainty with the acoustic ones, we augment the ensemble by introducing the perturbation of the geoacoustic properties, in a way similar to Rouseff and Ewart (1995). Each oceanic ensemble member is combined with five realizations of the geoacoustic properties. The TL ensemble is thus formed by 320 members. In particular, we consider uncertainty of 30 m/s for compressional wave sound speed, 0.05 dB/λ for attenuation, and 30 kg/m^3 for density, and the five combinations come from a random generation of the acoustic properties following a Gaussian distribution with mean the nominal value (see section 2.2) and standard deviation equal to these uncertainties. These values of uncertainty were estimated as standard deviation compiling several estimates for clayey silt to silty clay seabed, as found in the LOGMEC area (Bershad & Weiss, 1976), from textbooks and articles referring to such geological sea bottom composition (Hamilton & Bachman, 1982; Jensen et al., 2011; Lurton, 2010).

Explained ensemble variance diagnostics, similarly to section 3.1, are presented in Figure 5: The top panels show the bearing-averaged TL explained ensemble variance, while the bottom panels show the explained ensemble variance as a function of the bearing angle for the four most impacting perturbations. It turns out that LBC perturbation is the most impacting source of uncertainty at both 30 and 60 km range,
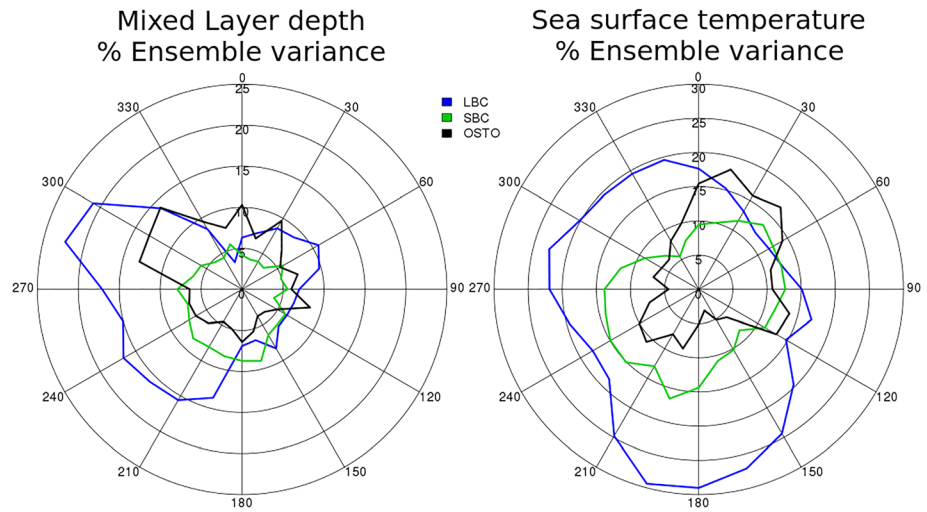
Figure 4. Mixed layer depth and sea surface temperature explained ensemble variance as a function of the bearing angle of the cross section with respect to the sound source, for the three leading perturbation types. LBC = lateral boundary condition perturbations; OSTO = ocean stochastic physics; SBC = surface boundary condition perturbations.

followed by SBC or ASTO (geoacoustic property perturbation) in the 30 or 60 km range, respectively. ASTO in particular exhibits large impact locally. For instance, in southeast direction (bearing equal to 150°), corresponding to a section with bathymetry abruptly rising, the perturbation of geoacoustic properties dominates over the other sources of uncertainty, while its contribution to the spread is generally small in other directions compared, for example, to the LBC perturbation. This feature highlights the impact of sea-bottom reverberation for such geometry. LBC shows a large impact on the western and southeastern direction, in analogy with the diagram presented in Figure 4. Results at 60 km range are obviously less
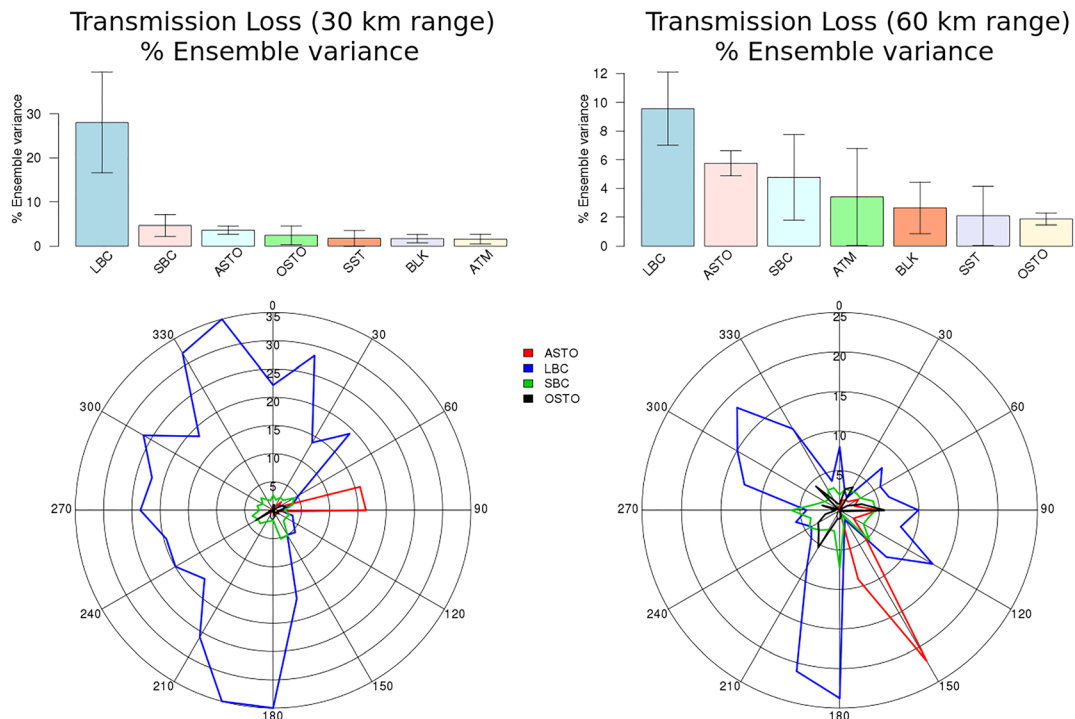

Figure 5. The same as Figures 3 and 4 but for the explained ensemble variance calculated at 30 and 60 km range transmission loss at 75 Hz. TL data are averaged along the last km of the transmission path (29–30 and 59–60 km range). ASTO = acoustic stochastic parameters.
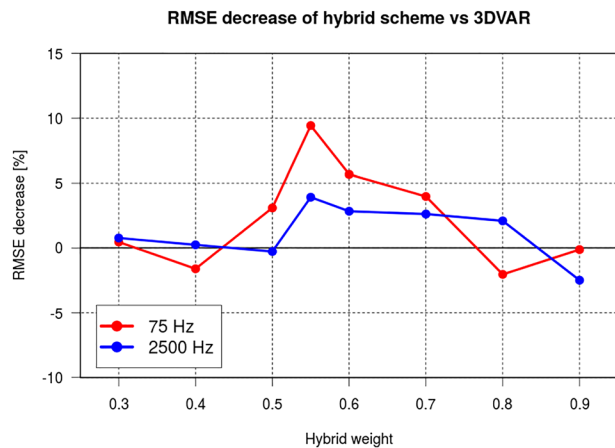
Figure 6. Bearing-mean 60 km transmission loss RMSE decreases with respect to 3DVAR, as a function of the hybrid weight, for the 75 and 2,500 Hz frequencies. Values are percentage relative to the 3DVAR RMSE; positive values indicate hybrid outperforming the standard 3DVAR. The weight equal to 0 (1) means fully ensemble-derived (stationary) background-error covariance matrix. TL data are averaged along the last km of the transmission path (59–60 km range).

neat than those at 30 km range due to the stronger influence that the bathymetry supposedly brings to long paths. The other sources are of less importance and exhibit less preferential directions.

Generalization of the previous results must, however, be taken cautiously. For instance, the impact of SBC on the spread of both ocean and acoustic fields may have a pronounced seasonal dependence. On the other hand, the impact of LBC depends obviously on the vicinity of the investigated fields on any domain boundary and to the impact of the transports through such boundary on the inner dynamics. Finally, the results suggest that the effect of geoacoustic uncertainty varies a lot for different bathymetry and geometries.

4. Results of the OSSEs

4.1. Experiments

In this section, we present the results from the OSSE experiments testing the impact of different ocean data assimilation schemes. CTR is the experiment without data assimilation; 3DVAR with the standard formulation of the three-dimensional variational analysis scheme; HYBRID with hybrid ensemble-variational background-error covariances, whose time-varying component is calculated from the ensemble system presented in section 2.3 and whose optimal weight is discussed below in

section 4.1; and finally, 4DVAR is the experiment with simplified four-dimensional variational data assimilation. Appendix A summarizes the formulations of the analysis schemes mentioned above. The schemes are implemented similarly to the global ocean system described in Storto et al. (2018), except for the sea level balance operator and for the regional tuning of error covariance matrices and tangent linear and adjoint operators in 4DVAR. The experiments 3DVAR and 4DVAR use the same background-error covariances, which are the same as the stationary component in HYBRID, used also within the members of the ensemble system.

The 4DVAR scheme is based upon the simplification that only the temperature and salinity evolve within the assimilation time window, while other state variables (in this case only the sea level) are diagnosed through the 3DVAR balance components (Storto et al., 2018). Such a simplification implies that only processes describing dynamical and thermodynamical evolution of temperature and salinity (diffusion, advection, and air-sea exchanges) are included in the tangent-linear and adjoint models, thus leads to a significant saving of computational costs compared to the classical 4DVAR problem, where also momentum tendencies need to be included and evaluated within the tangent-linear and adjoint models. 4DVAR is the scheme where background-error covariances evolve implicitly along with the assimilation time window, which is daily in our current setup. It does provide a more sophisticated and computationally expensive framework, which in principle may lead to analysis increments more dynamically balanced than 3DVAR with respect to the fully nonlinear ocean model, with potential improvements for the subsequent forecasts (e.g., Lorenc, 2003). Only one outer loop is used in the current configuration of 4DVAR.

Compared to 3DVAR, the HYBRID experiment was more expensive by a factor equal to the ensemble size (64 in this case), while the 4DVAR experiment resulted more expensive by a factor of approximately 15, in terms of CPU hours utilized. Note, however, that the parallelization of hybrid schemes, involving multiple members, is straightforward, while 4DVAR parallelization is more demanding from an algorithmic point of view and may easily suffer from poor scalability issues (Trémolet & Le Dimet, 1996).

4.2. Sensitivity to the Stationary/Ensemble Weight in the Hybrid Scheme

The hybrid scheme implemented in this study makes use of the augmented control vector formulation (see, e.g., Buehner, 2005), which is, however, equivalent to linearly combine stationary and ensemble-derived background error covariance matrices (Wang et al., 2007). Thus, although theoretical frameworks exist for optimal hybrid weight estimation (e.g., Ménétrier & Auligné, 2015), it is primarily important to identify the optimal hybrid weight that leads to the best forecast skill scores. Here, we investigate the optimal
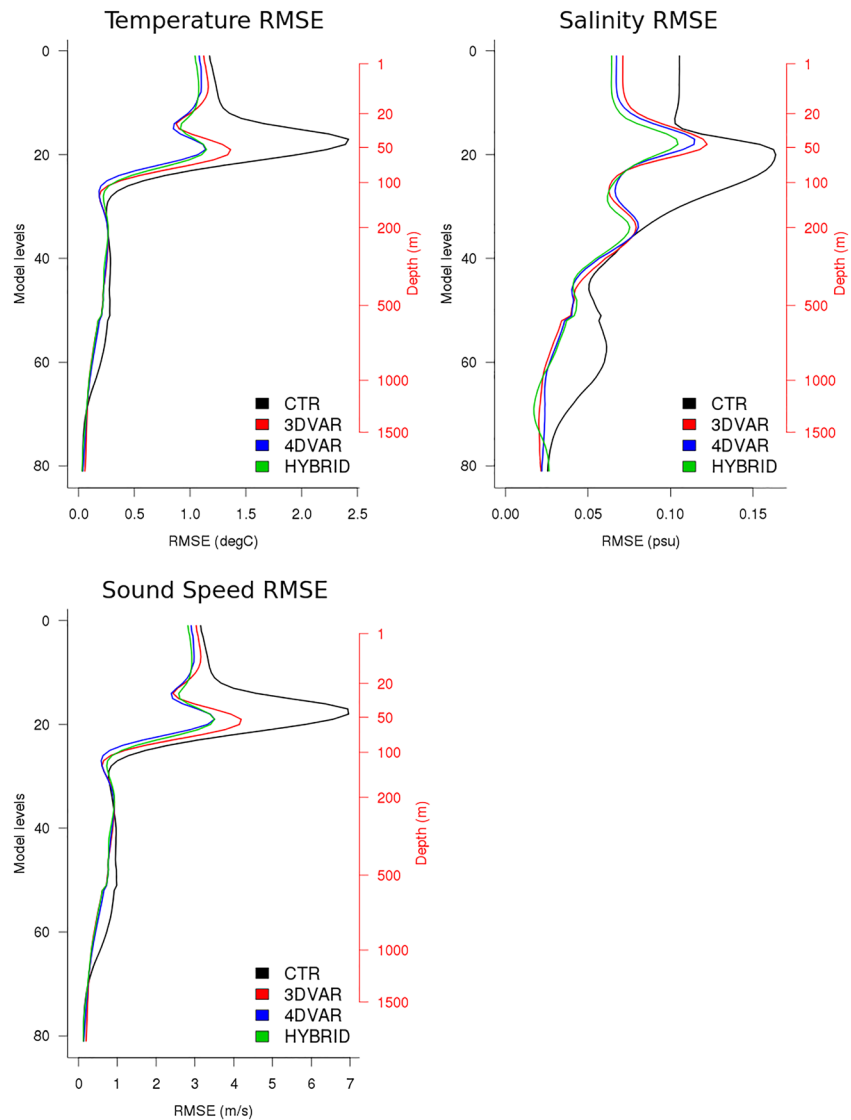
Figure 7. RMSE profiles versus the nature run over the round area corresponding to the acoustic synthetic observation scenario (60 km radius around the sound source, see Figure 1) for temperature, salinity, and sound speed and for the CTR, 3DVAR, HYBRID, and 4DVAR experiments. The period is from 15 October to 14 November 2017.

hybrid weight showing the TL RMSE decrease with respect to the standard 3DVAR case, for a discrete range of the weight from 0.3 to 0.9, where 0 corresponds to the fully ensemble covariance case and 1 to the fully stationary covariance case. See the parameter a in Equation A8 of Appendix A for the exact definition of the hybrid weight. Calibration of the weight based on RMSE scores is performed for the 60 km range acoustic TL, as for this range, the sensitivity to the data assimilation scheme amplifies and provides clearer indications than shorter ranges.

The statistics are shown in Figure 6 as bearing-averaged RMSE decrease, estimated during the period mid-October to mid-November. Interestingly, at both frequencies, the optimum hybrid weight is found for $a = 0.55$. In the a range between 0.5 and 0.7, the hybrid scheme provides an improvement with respect to 3DVAR for both frequencies. Note that we have tested a reasonable range of values between 0.3 and 0.9, assuming that values of a beyond this range provide worse results. RMSEs indicate a large consistency for the skill score improvements at the two frequencies, suggesting that the hybrid scheme is capable to impact the multiple scales that the sound speed propagation at different frequencies depends upon. In particular, optimal results are found near the equal partitioning of the stationary and ensemble-derived covariances.
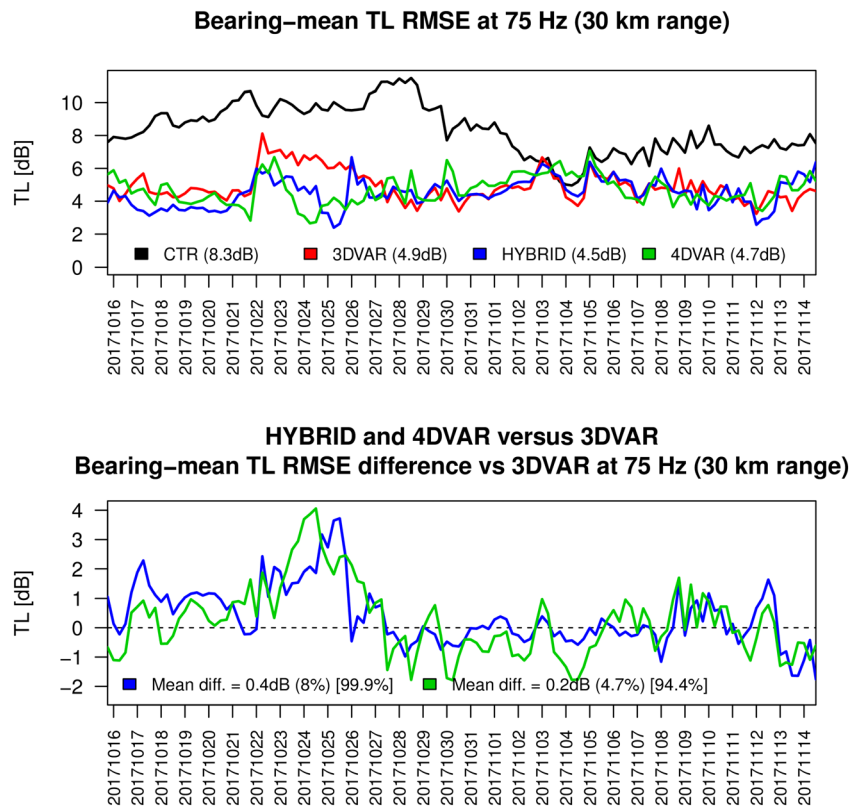
Figure 8. Bearing mean 30 km transmission loss RMSE against the synthetic acoustic observations for the 75 Hz configuration, for the four experiments CTR, 3DVAR, HYBRID, and 4DVAR (top panel). Values in parentheses are the time-averaged RMSE. The bottom panel shows the RMSE differences between 3DVAR and HYBRID (blue) and 4DVAR (green). Positive values indicate that HYBRID or 4DVAR outperforms 3DVAR. Values in parentheses are the time-averaged RMSE difference; values into squared brackets are the confidence level of the t test on the squared departures, for the null hypothesis that the two RMSE series belong to the same distribution, that is, when the values are greater than 95%, then HYBRID or 4DVAR RMSE differences with respect to 3DVAR are statistically significant.

In the remainder of this article, we will refer to the HYBRID experiment as the one with hybrid weight equal to 0.55, namely, the configuration leading to the best skill scores for the 60 km range TL.

It is worth noting that withholding certain perturbations from the ensemble system from which the flow-dependent covariances are computed provides detrimental results (not shown). For instance, without stochastic physics, the improvement with respect to 3DVAR for $a = 0.55$ decreases from 12% to 4% at 75 Hz and from 9% to 4% at 2,500 Hz. Similar results apply when LBC perturbations are withheld. It is however not possible to distinguish if this different performance is due to the enhanced ensemble spread that stochastic physics leads to, or to the different ensemble sizes itself, and dedicated experiments are required for a definitive answer, as future work.

4.3. Impact of the Schemes on Physical Ocean Variables

We review here the impact of the different schemes on physical ocean variables. This is shown in Figure 7 in terms of RMSE profiles of temperature, salinity, and sound speed over the validation period, for the circle region through which the sound propagates in the experiments. RMSEs are computed in the model space. Within the temperature scores, 3DVAR beats the CTR experiments in the top 150 m and between 200 and 1,000 m of depth. Compared to 3DVAR, both 4DVAR and HYBRID outperform 3DVAR in the top 20 m, with HYBRID providing the best skill scores. Between 20 and 100 m of depth, 4DVAR is the experiment that leads to the lowest RMSE. Around the mixed layer depth, corresponding to the maximum of the CTR experiment error, the errors decrease from 2.4°C (CTR) to 1.4°C (3DVAR) to 1.1°C (4DVAR and HYBRID). Below 100 m of depth, the impact of 4DVAR or HYBRID with respect to 3DVAR is negligible. The salinity RMSE profile is characterized by the largely positive impact of HYBRID, while 4DVAR's improvement is less marked,
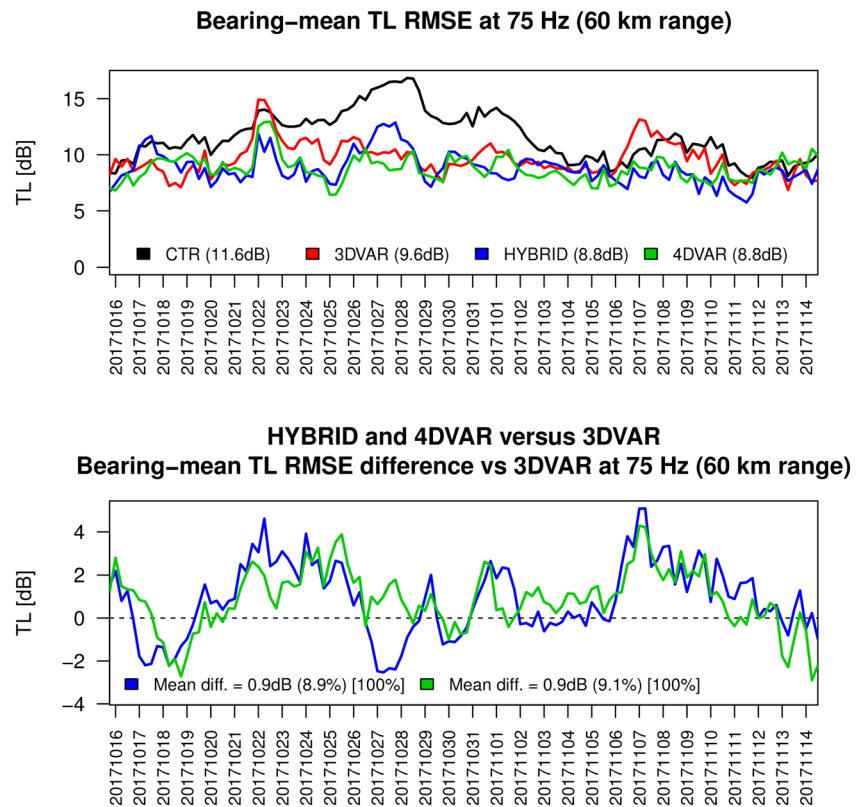
Figure 9. The same as Figure 8 but for the TL RMSE at 60 km range.

sometimes behaving worse than 3DVAR (near 100 m of depth) and suggesting that the physics embedded in the tangent linear and adjoint operators may be oversimplified. Finally, the RMSE profiles for sound speed resemble closely those of temperature, so they are not discussed further.

4.4. Impact of the Schemes on Acoustic TL

We evaluate here the impact of data assimilation of oceanic observations and the potential of the HYBRID and 4DVAR schemes for improving the TL accuracy. The assessment in terms of TL skill score metrics is presented as bearing-averaged RMSE versus time for the two study frequencies 75 and 2,500 Hz and the two ranges 30 and 60 km (Figures 8 to 11, respectively).

For the lowest frequency at 30 km range (Figure 8), ocean data assimilation improves the TL RMSE from 8.3 to 4.9 dB (CTR vs. 3DVAR), namely, the 3DVAR almost halves the errors, during all period and more pronouncedly during the first 20 days of the analysis period. There also occurs a significant improvement (at more than 99% confidence level) using HYBRID instead of 3DVAR, while the improvement of 4DVAR compared to 3DVAR is smaller. At 60 km range (Figure 9), TL RMSE suggests that further to a large improvement borne by data assimilation, especially visible until the beginning of November, 4DVAR and HYBRID further improve the RMSE statistics compared to 3DVAR, leading to a decrease of time-averaged RMSE from 9.6 to 8.8 dB (from the 11.6 dB of the observation-blind experiment). The difference of RMSE between 3DVAR and 4DVAR and 3DVAR and HYBRID are both of about 1.2 dB and statistically significant at a 99% confidence level, while differences between 4DVAR and HYBRID are not significant. The error decrease appears rather constant during the study period, although daily variability exists. During the last 10 days of the verification period (i.e., from 4 November 2017), when the 3DVAR data assimilation does not give large improvements compared to the control experiment, 4DVAR and HYBRID are able to improve the skill scores.

The same diagnostics are reported in Figures 10 and 11 for the 2,500 Hz experiments (as an example of sonar operating frequency). Also, in this case, 4DVAR and HYBRID outperform 3DVAR on the average for the 30 km range, with a time-mean error that reduces from 12.8 (3DVAR) to 11.9 dB (4DVAR) and 11.6 dB
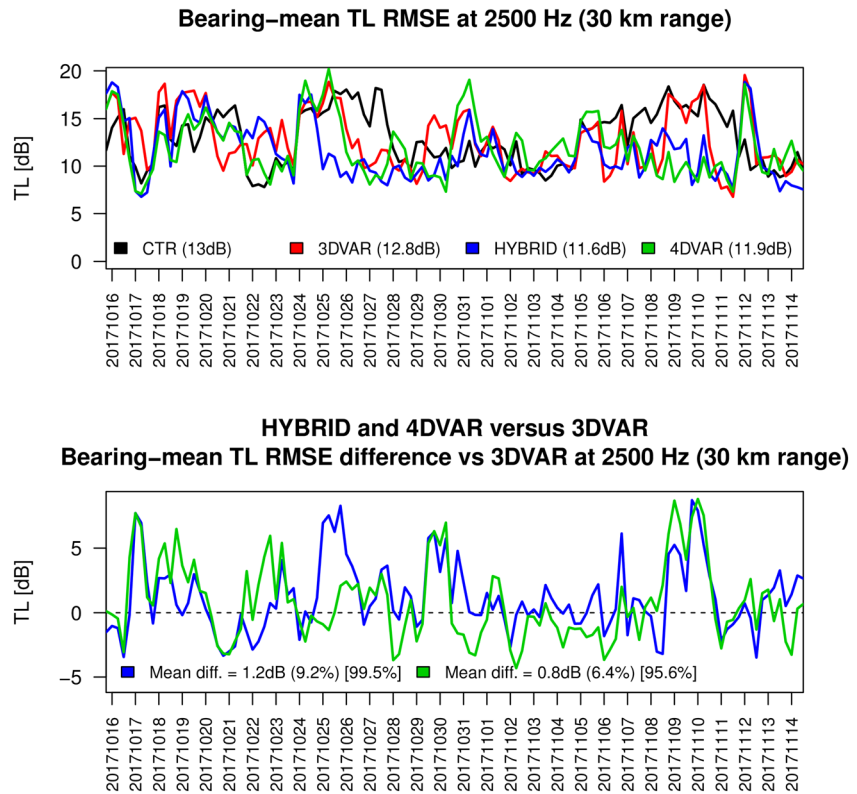
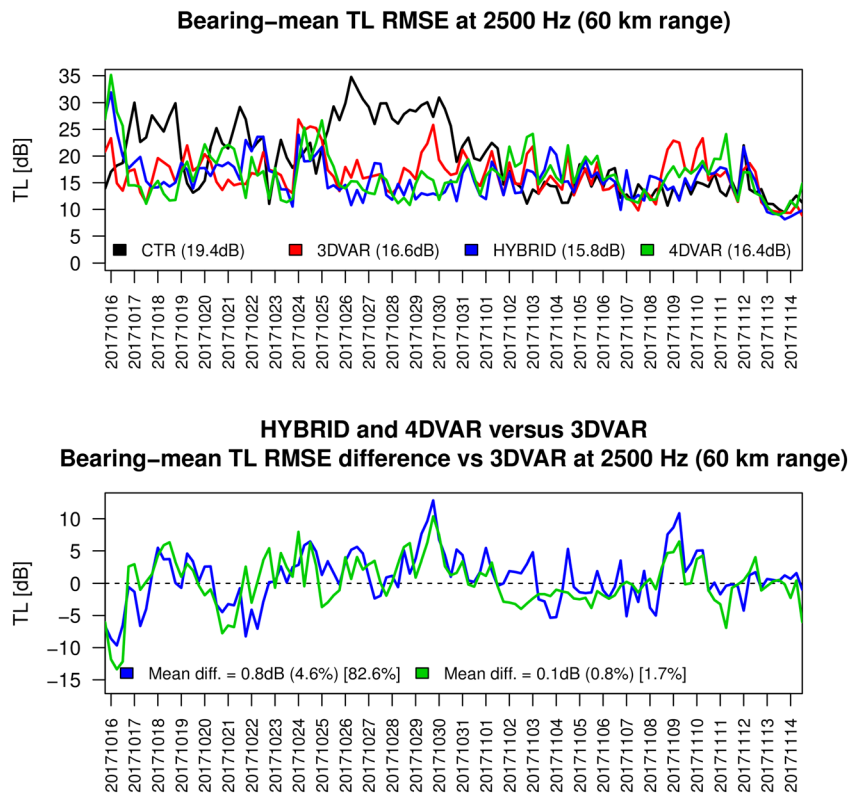**Figure 10.** The same as Figure 8 but for the 2,500 Hz configuration.

Figure 11. The same as Figure 8 but for the 2,500 Hz configuration and the TL RMSE at 60 km range.

Table 1
NRMSE for the Experiments Presented in the Text and the Physical and Acoustic Quantities Assessed

Experiment		CTRL	3DVAR	4DVAR	HYBRID
Variable					
Sound speed	**Depth (m)**				
	0–30	1.79 (2.0 m/s)	1.54 (13%)	1.48 (17%)	1.47 (18%)
	30–100	5.33 (1.0 m/s)	3.57 (33%)	2.71 (49%)	3.08 (42%)
	100–300	2.46 (0.4 m/s)	2.17 (12%)	2.08 (16%)	2.27 (8%)
	300–800	2.47 (0.4 m/s)	1.72 (30%)	1.79 (27%)	1.71 (31%)
Transmission loss (Hz)	**Range (km)**				
75	30	3.24 (2.6 dB)	1.91 (41%)	1.82 (44%)	1.76 (46%)
	60	4.52 (3.6 dB)	3.75 (17%)	3.41 (24%)	3.42 (24%)
2,500	30	5.04 (10.3 dB)	4.97 (2%)	4.65 (8%)	4.51 (11%)
	60	7.56 (15.8 dB)	6.45 (15%)	6.39 (15%)	6.15 (19%)

Note. The normalized root mean square error (NRMSE) is defined as the RMSE (against the validating synthetic observations) divided by the observed interquartile range, that is, the difference between upper and lower quartiles (75th and 25th percentiles) of the observation data. Values between parentheses report the observed interquartile range (within the CTRL experiment column) and NRMSE decrease with respect to the CTRL experiment (within the columns relative to the other experiments). For sound speed, the RMSE is computed over all the 60 km radial paths (see Figure 1), while for transmission loss is computed over the last kilometer at ranges 30 or 60 km, at the receiver depth (60 m).
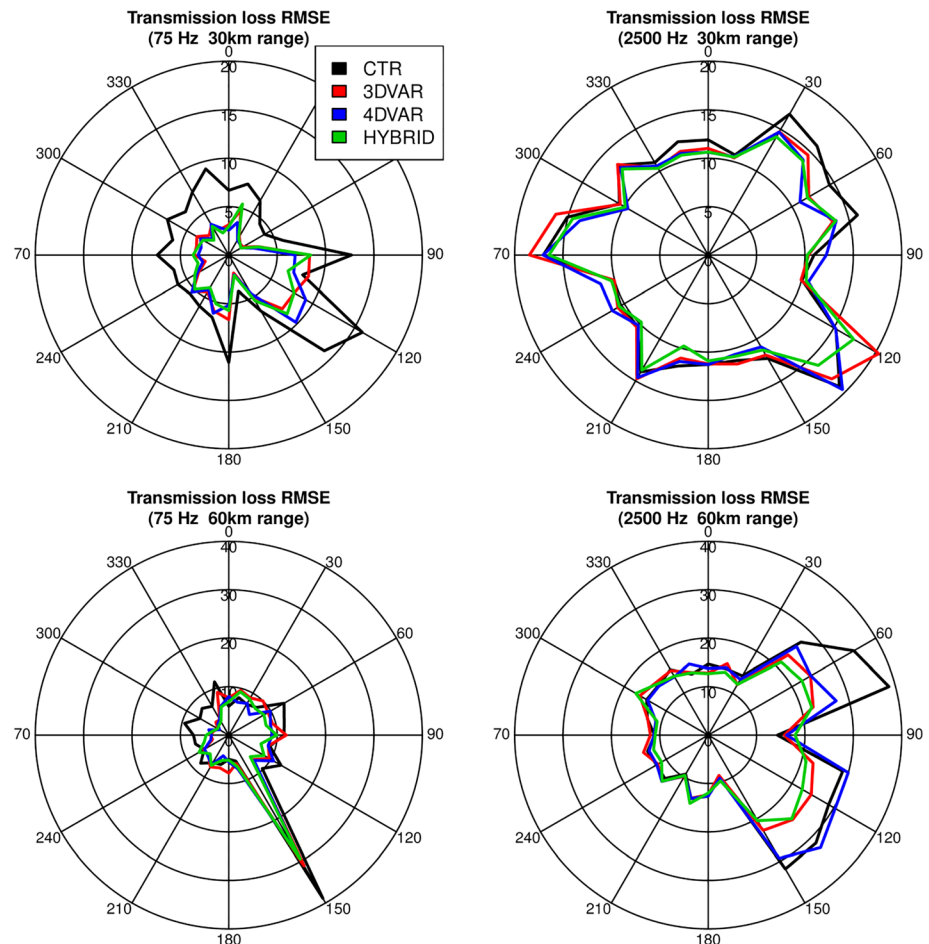

Figure 12. Time-averaged transmission loss RMSE (in dB) versus bearing angle, for the 75 and 2,500 Hz frequency OSSEs and the 30 and 60 km ranges.
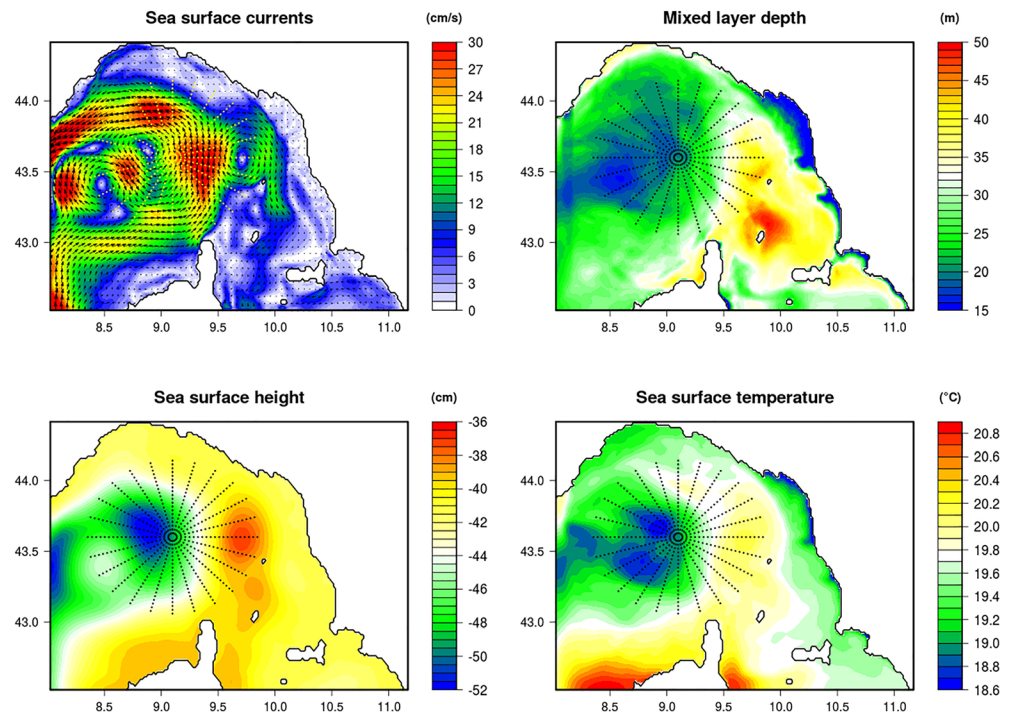
Figure 13. Mean state of sea surface currents, mixed layer depth, sea surface height, and sea surface temperature during the period from 16 October to 14 November 2017. Mean fields are calculated from the HYBRID experiment. Dots correspond to the acoustic propagation paths investigated in this work.

(HYBRID), compared to the 13 dB for the control experiment. In this acoustic scenario, the decrease of RMSE in 4DVAR compared to 3DVAR is not statistically significant, while those of HYBRID compared to 3DVAR or CTR are both significant at a 99% confidence level. For the 60 km range at 2,500 Hz, results are noisier than the other cases. There exists a clear improvement of 3DVAR with respect to CTR (14% RMSE decrease), while the improvements of HYBRID and 4DVAR are not significant, although the latter leads to the smallest TL RMSE.

In Table 1, we summarize the validation results for both physical and acoustic parameters, in terms of Normalized RMSE (NRMSE; defined as the RMSE divided by the observed interquartile range). NRMSE is dimensionless and accounts for the variability of the validated variable. It is used here in order to compare results for oceanic and acoustic parameters. The largest sound speed errors are found at the thermocline (30–100 m, with values 3–5), where data assimilation provides its largest impact on physical quantities (reduction of NRMSE up to about 50%). Below and above, values of NRMSE are smaller (1.5–2.5), and the impact of data assimilation on sound speed accuracy never exceeds 30%. The predicted acoustic TL accuracy decreases with the range and frequency, as expected. For instance, values of the NRMSE for 60 km range at 2,500 Hz vary between 6 and 7.5, indicating that for this case, errors are significantly larger than those of sound speed, while for the low-frequency short-range case (30 km at 75 Hz), the errors are comparable with those of sound speed (1.7–3.2). The data assimilation impact is, in general, larger for the low-frequency case compared to 2,500 Hz, although the long-range 2,500 Hz case exhibits also a significant impact of data assimilation. The fact that HYBRID shows the best performances in all acoustic prediction cases matches its peak performance within the mixed layer sound speed validation, consistency with the location of the acoustic source depth.

TL RMSE statistics versus bearing (Figure 12) show that for moderate ranges such as 30 km, improvements are rather uniform across the different directions, while large improvements through data assimilation are found for the sections affected by the largest uncertainty at the longest range (60 km), likely related to the presence of complex bathymetry and the occurrence of small-scale oceanic features such as fronts and meanders. For the 60 km range, the presence of bathymetric gradients along the cross section corresponding to the 150° bearing angle leads to very large errors. Figure 13 provides the ocean mean state during the
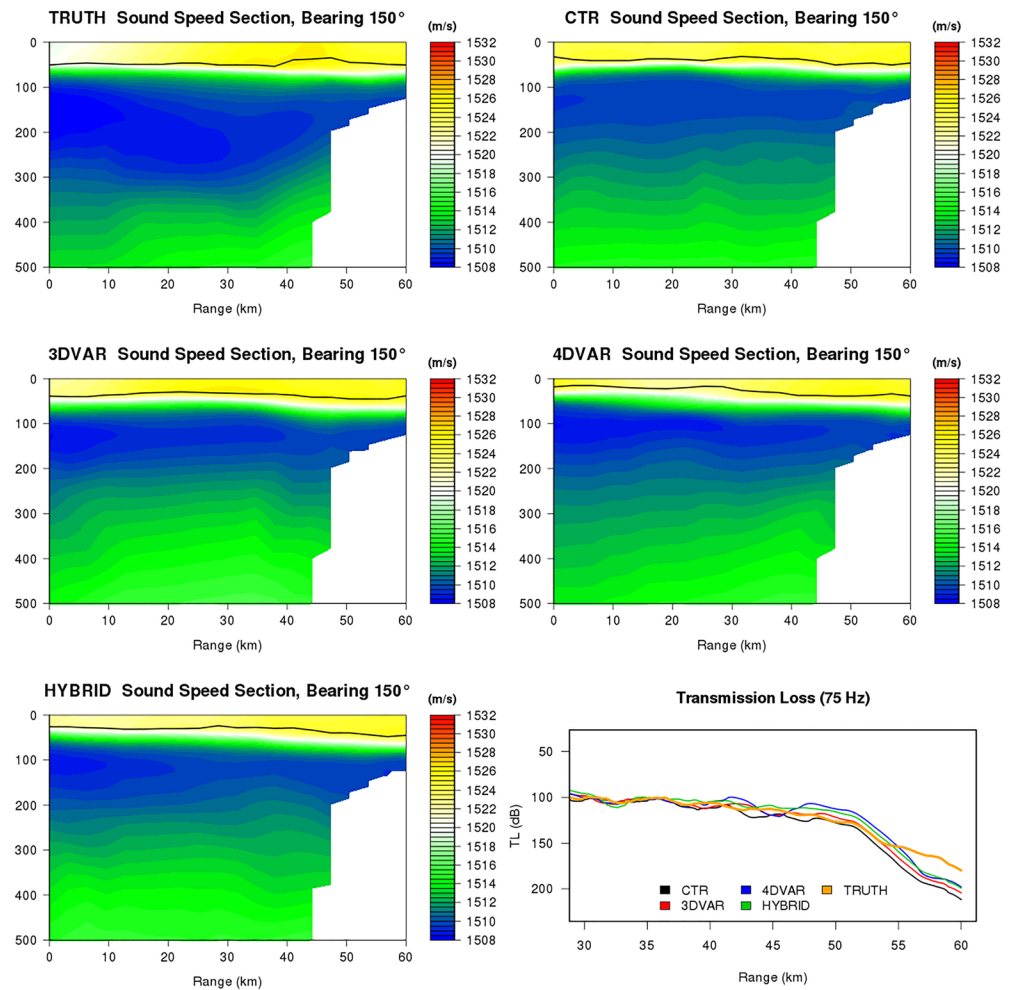
Figure 14. Cross sections of sound speed profiles along the propagation path of bearing equal to 150°, for the experiments explained in the text. The data refer to 5 November 2017. The black line corresponds to the mixed layer depth along the propagation path, for each experiment. The bottom-right panel shows the TL as a function of range from 30 to 60 km, for the signal frequency at 75 Hz and receiver depth at 60 m.

period 16 October to 14 November 2017 for selected parameters, suggesting indeed that for the long-range case (60 km), certain propagation paths vary widely and go across fronts and current systems, which may amplify the sensitivity of the predicted acoustic TL to the ocean state and in turn to the analysis schemes. This stresses the dependence of predicted TL on the ocean small-scale variability. For instance, at 150° (counting clockwise from north), the TL RMSE for the 60 km range at 75 Hz decreases from 40 to 33 dB with 3DVAR and to 31 and 28 dB with HYBRID and 4DVAR, respectively. Between 0° and 90°, where 3DVAR generally exhibits detrimental impact with respect to the CTR experiment, 4DVAR provides a positive impact of data assimilation (from 9.3 to 8.1 dB, averaged between 0° and 90° and compared to 9.8 dB of 3DVAR). Except a very few bearing angles, also at 2,500 Hz HYBRID and 4DVAR are able to improve the error statistics, although to a lesser extent than 75 Hz. HYBRID provides the best skill scores in particular between 30° and 150° at both ranges.

To summarize, we found significant improvements of 3DVAR against no data assimilation at both frequencies and ranges and improvements of 4DVAR and HYBRID in all cases. HYBRID performs better than any other experiment in all cases (except for 60 km range and 75 Hz, where it performs as well as 4DVAR).

4.5. Examples of TL Events

Two case studies are presented in this section, as an example of specific events. First, underwater acoustic propagation at 75 Hz is investigated for the cross section with bearing equal to 150°, which is
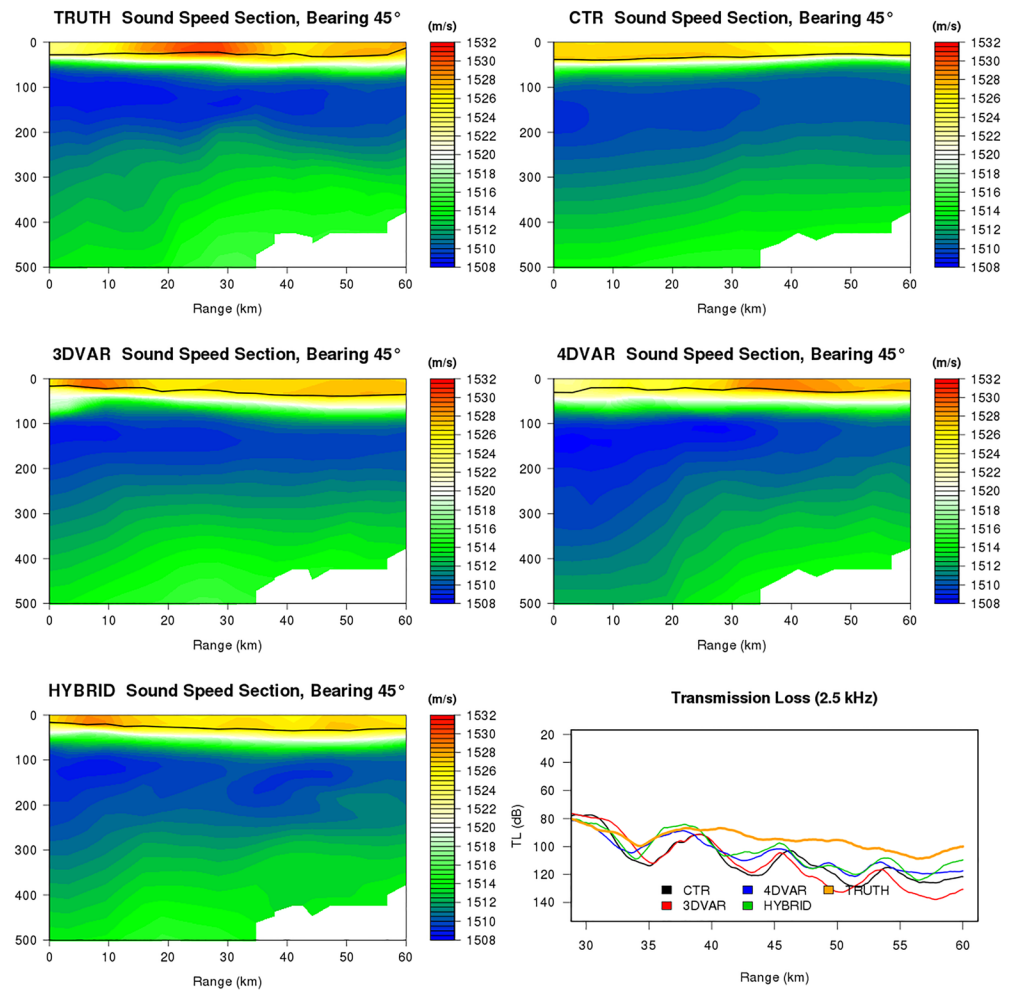
Figure 15. The same as Figure 14 but for the section with bearing equal to 45°. Transmission loss is at 2,500 Hz, for 23 October 2017.

characterized by the largest uncertainty in the CTR experiment (40 dB on average), for a specific day (5 November 2017). Figure 14 shows the sound speed cross section for the different experiments, along with the TL versus range diagram, at the receiver depth, for the experiments introduced earlier. This section is the one with the steepest decrease of bathymetry along the track, decreasing at the end of the section to about 150 m of depth off northeastern Corsica Island, due to the noticeable impact of geoacoustic uncertainty (see Figure 5).

While TL data are rather homogenous among the experiments until the 40 km range, all the experiments exhibit overestimation of the TL within the range 55–60 km, which causes the errors to be very large. However, 3DVAR, HYBRID, and 4DVAR all partly mitigate the TL overestimation compared to CTR. The associated structure of the sound speed sections appears different between the experiments in many aspects. The TRUTH (NR) exhibits a depth-averaged warming along the cross section (toward southeast), which in particular translates in colder and shallower mixed around the 45 km range, with a cold intrusion near the coast and a sharper thermocline. These features are not present in the CTR, which shows rather a homogenous stratification of temperature and sound speed. Data assimilation, and especially 4DVAR, is able to mitigate this effect and tilt the thermocline to some extent, although vertical gradients below the mixed layer are underestimated. The cold intrusion is partly reproduced with 4DVAR, although misplaced compared to TRUTH. The sound speed channel location and horizontal structure are consequently modified. In particular, data assimilation is able to deepen and narrow the location of the sound speed channel toward the coast, in better agreement with TRUTH than CTR. Below 300 m of depth, no experiment performs close to

TRUTH, as a consequence of the small impact of data assimilation in the deep ocean (see Figure 7). This is likely due to specification of observational errors at depth, although it seems not to affect acoustic propagation data for the sound-receiver scenario investigated in this work.

A second example is provided in Figure 15, together with the associated TL at 2,500 Hz, for a different date (23 October 2017) and bearing (45°). Here, the bathymetry decreases along the section but less sharply than the 150° case. There is, in this case, a slight deepening with the range of the mixed layer in TRUTH, while CTR shows, on the contrary, a shallower mixed layer toward the coast. In this case, the TL is overestimated by CTR and, to an even greater extent, 3DVAR, which fails in reproducing the TL over the 40–60 km range of about 10–20 dB. HYBRID is found to outperform all other experiments, significantly mitigating the TL overestimation. In general, these two examples illustrate how ocean data assimilation can impact the TL prediction as a function of the distance from the sound source.

5. Summary and Discussion

Assessing the impact of oceanic observations and oceanic data assimilation onto integrated ocean-acoustic predictions emerges as fundamental task for a large variety of purposes, ranging from ship noise prediction to active sonar-based applications. A well-known challenge in oceanic-acoustic coupling resides on the significant impact of the oceanic small scales (submesoscale), generally smaller than the resolution of the ocean model, on the acoustic fields. Such an impact of small-scale features increases dramatically with the acoustic frequency and makes it questionable whether the mesoscale or large-scale oriented analysis systems adequately ingest oceanic observations for various acoustic applications. Indeed, the main goal of the present study is to assess the feasibility of assimilating ocean observation for improving acoustic propagation prediction. Main novelties of the work reside on (i) model error-aware ensemble generation formulation for hybrid ensemble-variational data assimilation, (ii) assessment of the impact of different sources of uncertainty on oceanic and acoustic prediction accuracy, (iii) adaptation of the OSSEs procedure for the oceanic-acoustic modeling framework, and (iv) assessment and comparison of the impact that different ocean analysis schemes have on the accuracy of the predicted TL maps.

Due to the sparseness and costs of acoustic observation collection, in this work, we focus on extending the OSSEs procedures for oceanic-acoustic prediction. Oceanic and acoustic synthetic observations were initially extracted from an integrated ocean-acoustic NR with perturbed physics and boundary forcing. Oceanic observations are then assimilated into a regional ocean prediction system, after perturbing them with realistic errors, while synthetic acoustic observations are used for validating the experiments and assess the most impacting ocean analysis scheme on TL data.

The work was done with a regional ocean-acoustic modeling system that covers the Ligurian Sea. The acoustic scenario includes a sound source and a number of propagation paths radially arranged to cover a large part of the ocean model domain and consider two study frequencies, 75 and 2,500 Hz, representative of ship noise and active sonar propagation, respectively. The experimental period starts from August 2017, but validation with acoustic synthetic observations covers the 1 month period from 15 October to 14 November 2017. The modeling framework includes the primitive equation ocean model NEMO implemented at less than 2 km of horizontal resolution, the RAM range-dependent sound propagation model, and a variational data assimilation system. For the latter, a 4DVAR implementation that embeds simplified tangent-linear and adjoint models and a hybrid ensemble-variational scheme are implemented for comparison to the standard 3DVAR formulation. Specifically, care is given to the ensemble generation strategy that includes a novel stochastic physics scheme and perturbation of input data sets and boundary forcing, which in turn proved crucial in the hybrid experiments, although dedicated experiments should be performed to assess in details the role of perturbations in the hybrid scheme.

We first evaluate the uncertainty of the oceanic and acoustic fields, assessed through clustering of the ensemble members. The uncertainty in the LBCs dominates the total ensemble spread, although a nonnegligible impact is provided by stochastic physics around the mixed layer depth and SBCs in the proximity of the sea surface. TL uncertainty is also dominated by LBC uncertainty, although for selected geometries where the sea-bottom scattering plays an important role, geoacoustic parameter uncertainty becomes predominant. These findings are however intimately linked to the chosen geometry and study period and cannot be generalized to other prediction systems.

Through comparing the different analysis schemes, we found that the upper ocean (top 100 m) temperature, salinity, and sound speed are significantly improved by 4DVAR and HYBRID. In the comparison of TL values at 30 and 60 km range, the oceanic data assimilation is always found to improve significantly the TL accuracy over the control run. Moreover, HYBRID and 4DVAR improve the skill scores compared to 3DVAR. The HYBRID experiment is the experiment with the highest accuracy at all ranges and both frequencies, although differences for long ranges and high frequency (2,500 Hz at 60 km range) with respect to 3DVAR are not significant, due to pronounced noise in the TL signals. These findings are, in general, crucial to prove that the uncertainty in the geoacoustics parameters and acoustic geometry does not jeopardize the effectiveness of the data assimilation corrections.

The hybrid 4DVAR scheme (not implemented in this study) might further improve the verification skill scores, although this needs to be verified experimentally and may not necessarily occur, due to the strong nonlinearity of the coupled ocean-acoustic system. In any case, this study suggests that the optimal choice of the assimilation method depends upon (i) computational constraint; (ii) application, that is, frequency of interest of the underwater sound propagation; and (iii) target accuracy for acoustic predictions, considering all these factors together within a cost-benefit analysis.

The HYBRID scheme performs better than the other schemes, but it is the most computationally intensive scheme. The error decrease that HYBRID leads to appears linked to its best performances in the mixed layer sound speed validation. The choice of the optimal scheme shall also consider the specific study and computational demand. In most of the case studies, the errors in predicted TL are larger than those in the sound speed, indicating error amplification in the acoustic propagation fields due to their nonlinear sensitivity to the oceanic features. As expected, the error propagation amplifies for long ranges and high frequencies, suggesting that for such cases targeting the prediction problem with oceanic and acoustic observations might become necessary. Oceanic data assimilation is, however, able to significantly increase the TL prediction accuracy, although its actual impact depends upon the study case.

Although the sensitivity of oceanic-acoustic prediction systems on the oceanographic analysis scheme emerges clearly from this study, it will be important to confirm these findings in the future with real observation experiments. The synthetic observation framework is, however, an invaluable strategy for investigating optimal assimilation methods and/or observational sampling impact, for instance, for optimal planning of glider missions (e.g., Alvarez & Mourre, 2012). For instance, with this framework, (i) coupled ocean-acoustic data assimilation schemes can be evaluated (i.e., assimilating underwater acoustic measurements to improve the oceanic state); (ii) halving or modifying the oceanic acoustic sampling may be assessed prior to dedicated deployments and observational campaigns, eventually using also ensemble-based observation impact diagnostics (Ota et al., 2013; Storto et al., 2013). Indeed, coupled ocean-acoustic data assimilation—assessed with OSSEs first and with data from sea trial then—is the natural continuation of this work.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Data Availability Statement

The NEMO and RAM oceanic and acoustic models are freely available at https://www.nemo-ocean.eu/ and http://oalib.hlsresearch.com/PE/RAM/, respectively. Data from the OSSE experiments are available from the authors on demand.

Appendix A: The Oceanographic Data Assimilation Scheme

The oceanographic analysis scheme is based on a 3DVAR data assimilation system, which uses the standard incremental formulation (e.g., Courtier, 1997). The analysis increments δx are found for the ocean state x at the minimum of the following cost function:

$$J(\delta \boldsymbol{x}) = \frac{1}{2} \delta \boldsymbol{x}^T (\mathbf{B})^{-1} \delta \boldsymbol{x} + \frac{1}{2} (\mathbf{H}\delta \boldsymbol{x} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{H}\delta \boldsymbol{x} - \mathbf{d}), \tag{A1}$$

where $\mathbf{B}$ and $\mathbf{R}$ are the background- and observation-error covariance matrices, $\mathbf{d}$ is the vector of misfits, with $\mathbf{d} = \boldsymbol{y} - H(\boldsymbol{x^b})$, $\boldsymbol{y}$ the observation vector, $\boldsymbol{x^b}$ the background, $H()$ the observation operator mapping

the ocean state onto observation space, and $\mathbf{H}$ the tangent-linear version of the observation operator, linearized around the background state. This implementation considers the First Guess at Appropriate Time algorithm, namely, $\mathbf{d}$ is computed using the background fields at the exact time of the observations. In order to precondition the minimization problem and avoids the inversion of $\mathbf{B}$, a control vector ($\mathbf{v}$) transformation is applied, so that the cost function actually minimized is

$$J(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{H}V\mathbf{v}-\mathbf{d})^T\mathbf{R}^{-1}(\mathbf{H}V\mathbf{v} - \mathbf{d}), \tag{A2}$$

with $\delta x = V\mathbf{v}$ and $\mathbf{B} = VV^T$.

The operator V is thus the left square root of $\mathbf{B}$, and it is defined as a sequence of operators:

$$V = V_bV_hV_v, \tag{A3}$$

where V_v and V_h account for the vertical covariances and horizontal correlations, modeled through multivariate EOFs and the application of a recursive filter operator, respectively. The operator V_b accounts for cross-covariances, that is, balances between different ocean parameters. In particular, while temperature and salinity are independent, sea level is obtained as a combination of balanced and unbalanced component, which formally reads

$$\delta\eta = \delta\eta_b + \delta\eta_u = K_dK_\rho V_hV_v\, \mathbf{v} + V_hV_v\, \mathbf{v}, \tag{A4}$$

with K_d the dynamic height operator applied to density fields (obtained with the K_ρ operator applied to temperature and salinity), such that

$$K_d : \delta\eta = -\frac{1}{\rho_0}\int_{z_l}^{0}\delta\rho\, \mathrm{d}z, \tag{A5}$$

with ρ_0 the reference density (1,026 kg/m^3), z_l the bottom depth for integration (also called level-of-no-motion), and $\delta\rho$ the density increment ($\delta\rho = K_\rho(\delta T, \delta S)$), calculated from the tangent-linear version of the UNESCO 1983 (Fofonoff & Millard, 1983) polynomial equation of state, linearized around the background state (T^b, S^b). In practice, the characterization of (A3) requires the definition of vertical EOFs and horizontal correlation length scales (in V_v and V_h, respectively), which are computed off-line, prior to the variational minimization, from data sets of anomaly of temperature, salinity, and sea surface height.

Fast minimization is achieved through the L-BFGS minimizer (Byrd et al., 1995), which requires the explicit computation of the gradient of the cost function, as

$$\nabla J(\mathbf{v}) = \mathbf{v} + (\mathbf{H}V)^T\mathbf{R}^{-1}(\mathbf{H}V\mathbf{v} - \mathbf{d}). \tag{A6}$$

Preliminary sensitivity experiments were conducted in order to retune the background- and observation-error covariance matrices. In particular, background-error vertical covariances (EOFs) were estimated from the data set of temperature, salinity, and sea level anomalies with respect to the long-term mean (from September to November 2017), taken as daily means every 3 days to ensure ergodicity of the statistics. For each grid point, a set of EOFs was calculated including also the neighbor grid points to avoid possible rank deficiency.

The hybrid ensemble-variational capability of the data assimilation scheme is introduced essentially by augmenting the control vector to include a flow-dependent term ($\mathbf{v}_f$), to which a flow-dependent background error covariance matrix is associated ($B_f = V_fV_f^T$), further to the usual stationary term ($\mathbf{v}$). This is done by redefining the control variable transformation such that

$$\delta x = \alpha V\mathbf{v} + \beta\, V_f\mathbf{v}_f, \tag{A7}$$

which is equivalent to consider the background-error covariance matrix as a linear combination of the flow-dependent and stationary components (Oddo et al., 2016)

$$B = a\, V\, V^T + (1 - a)\, V_f V_f^T. \tag{A8}$$

Note that the previous relationships hold with $\alpha = \sqrt{a}$ and $\beta = \sqrt{(1 - a)}$ (see Wang et al., 2007) and when the two components can be approximated as independent (orthogonal, see Oddo et al., 2016), namely, no cross-covariances between stationary and flow-dependent components are formally allowed. The form of V_f is as in Equation A3, namely, it is a parametric definition for which EOFs and correlation length scales are computed from the ensemble system; V is exactly the same as in standard 3DVAR, that is, with $a = \alpha = 1$ the hybrid scheme simplifies to the 3DVAR in (A2).

Finally, the 4DVAR extension of the data assimilation system implies redefining the cost function as

$$J(v) = \frac{1}{2} v_0^{\,T} v_0 + \frac{1}{2}(\mathbf{HMV}v_0 - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{HMV}v_0 - \mathbf{d}), \tag{A9}$$

where $\mathbf{M}$ is the tangent-linear version of the model, and v_0 indicates the control vector valid at the beginning of the assimilation time window, with $\delta x = \mathbf{M}V v_0$ the new control variable transformation that consider δx as the four-dimensional ocean state. Including for simplicity, the tangent-linear model in the sequence of the chain operator V yields

$$V = \mathbf{M}\, V_b V_h V_v. \tag{A10}$$

A simplification embedded in our 4DVAR implementation resides on switching the balance and tangent-linear model operators, such that Equation A10 is redefined as

$$V = V_b \mathbf{M}\, V_h V_v. \tag{A11}$$

This change of operator sequence implies that that tangent-linear model is responsible for the temporal propagation of only balanced state variables, namely, only temperature and salinity. Sea level is obtained as a combination of balanced and unbalanced components such that

$$\delta \eta = \delta \eta_b + \delta \eta_u = K_d K_\rho \mathbf{M}\, V_h V_v\, v_0 + V_h V_v\, v_0. \tag{A12}$$

Equation A12 indicates that the propagation of the sea level increment within the assimilation time window is achieved not directly with the tangent-linear model but with the balance operator, which is applied to the temperature and salinity increments propagated in turn by the tangent-linear and adjoint models.

The tangent-linear model $\mathbf{M}$ (and consequently the adjoint model) includes only a number of physical processes, that is, horizontal and vertical advection and mixing, and air-sea exchanges. Detailed information on the tangent-linear model $\mathbf{M}$ is contained in Storto et al. (2018). The 4DVAR implementation considered in this study includes one single outer loop. Inner loops are performed at the same resolution as the NEMO ocean model implementation, thus inheriting mixing, time step, and other setup from the ocean model configuration. The frequency of nonlinear trajectories for the tangent-linear approximation is 3 hourly. In the future, a simplification operator for running the inner loops at a lower resolution than the outer loops will be implemented for sake of computational saving. The hybrid 4DVAR extension is straightforward and could be achieved by merging the previous extensions, that is, using the 4DVAR cost function (A9) with the hybrid covariance definitions of (A7) and (A8), although here it is not considered because of its very high computational costs.

Appendix B: SPPT Scheme

SPPT is the perturbation scheme where state variable tendencies are perturbed collinearly to the subgrid parameterization tendencies, whose errors are assumed to upscale from unresolved scales. The scheme was originally developed at ECMWF for introducing stochastic physics in the atmospheric model for medium- to long-range ensemble predictions (Buizza et al., 1999). Recently, it has been adopted for ocean simulations, providing promising results in long-range prediction applications (Andrejczuk et al., 2016).

The SPPT implementation presented here has been coded from scratch in NEMO 3.6, and it is inspired by the revisited SPPT scheme of Palmer et al. (2009).

Consider the general equation for the tendency of an ocean state variable $\boldsymbol{X}$ (e.g., temperature, salinity, or currents):

$$\frac{\partial \boldsymbol{X}}{\partial t} = D(\boldsymbol{X}) + P(\boldsymbol{X}), \tag{B1}$$

where $\frac{\partial \boldsymbol{X}}{\partial t}$ is the total tendency, $D(\boldsymbol{X})$ is the tendency from the dynamics, namely, equal to $-\nabla \cdot (\boldsymbol{XU})$ with $\boldsymbol{U}$ the Eulerian velocity, and $P(\boldsymbol{X})$ represents parameterized processes, including diffusion and eventually boundary conditions. In the SPPT scheme, the tendencies from parameterized processes are perturbed collinearly to the model tendencies:

$$\frac{\partial \boldsymbol{X}}{\partial t} = D(\boldsymbol{X}) + (1 + \boldsymbol{\xi}(t, z)) \, P(\boldsymbol{X}), \tag{B2}$$

where the perturbation field is given by

$$\boldsymbol{\xi}(t, z) = \boldsymbol{\gamma}(z)[c \, \boldsymbol{\xi}(t-1, z) + d \, \boldsymbol{\phi}(t)] \tag{B3}$$

is an AR(1) process, with $c = \exp(-1/\tau)$ and $d = \sqrt{(1 - c^2)}$. τ is the decorrelation time scale; $\boldsymbol{\phi}(t)$ is the white noise field, obtained in turn through applying independent random realizations drawn from the $N(0, \sigma^2)$ distribution to each grid point and correlating them in space by means of a Laplacian filter. Finally, $\boldsymbol{\gamma}(z)$ is a vertical weight function, conceived to possibly zero the perturbation fields at the sea surface and in correspondence of the vertical level at the sea bottom. The perturbation field $\boldsymbol{\xi}(t,z)$ is bounded to values between -1 and 1, in order not to reverse the sign of the perturbed tendency with respect to the original tendency. The Laplacian filter (or first-order Shapiro filter; Shapiro, 1970) is a low-pass filter which permits spatial correlations of the spatially independent random realizations, which reads

$$P_{(i,j)}^k = \frac{1}{8}\Big(4\, P_{(i,j)}^{k-1} + P_{(i-1,j)}^{k-1} + P_{(i+1,j)}^{k-1} + P_{(i,j-1)}^{k-1} + P_{(i,j+1)}^{k-1}\Big), \tag{B4}$$

where k indicates the kth iteration of the filter. For such first-order filter, the response function R at a given length scale L for a mesh resolution equal to Δx and N iterations is given by

$$R = \left(1 - \sin^2\left(\frac{\pi L}{\Delta x}\right)\right)^k. \tag{B5}$$

The SPPT configuration used for the present study considers as perturbed tendencies horizontal and vertical mixing and solar radiation penetration scheme, with a decorrelation time scale τ equal to 2 days, the dimensionless standard deviation σ equal to 0.5, and the Laplacian filter configured with 50 passes, which roughly correspond to 50% attenuation at the spatial scales of 50 km. Additionally, perturbations closer than 8 km to the coast are linearly smoothed to 0 at the shoreline. Other parameterized tendencies (e.g., bulk formulas and boundary conditions) remain unperturbed.

References

Alvarez, A., & Mourre, B. (2012). Optimum sampling designs for a glider–mooring observing network. *Journal of Atmospheric and Oceanic Technology*, *29*, 601–612. https://doi.org/10.1175/JTECH-D-11-00105.1

Andrejczuk, M., Cooper, F. C., Juricke, S., Palmer, T. N., Weisheimer, A., & Zanna, L. (2016). Oceanic stochastic parameterizations in a seasonal forecast system. *Monthly Weather Review*, *144*, 1867–1875. https://doi.org/10.1175/MWR-D-15-0245.1

Arnold, C. P. Jr., & Dey, C. H. (1986). Observing-systems simulation experiments: Past, present, and future. *Bulletin of the American Meteorological Society*, *67*, 687–695.

Atlas, R. (1997). Atmospheric observations and experiments to assess their usefulness in data assimilation. *Journal of the Meteorological Society of Japan*, *75*, 111–130.

Bershad, S., & Weiss, M. (1976). Deck41 surficial seafloor sediment description database. National Geophysical Data Center, NOAA. https://doi.org/10.7289/V5VD6WCZ

Bessières, L., Leroux, S., Brankart, J.-M., Molines, J.-M., Moine, M.-P., Bouttier, P.-A., et al. (2017). Development of a probabilistic ocean modelling system based on NEMO 3.5: Application at eddying resolution. *Geoscientific Model Development*, *10*, 1091–1106. https://doi.org/10.5194/gmd-10-1091-2017

Bishop, C. H., & Satterfield, E. A. (2013). Hidden error variance theory. Part I: Exposition and analytic model. *Monthly Weather Review*, *141*, 1454–1468.

Brankart, J.-M. (2013). Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling. *Ocean Modelling*, *66*, 64–76.

Brankart, J.-M., Candille, G., Garnier, F., Calone, C., Melet, A., Bouttier, P.-A., et al. (2015). A generic approach to explicit simulation of uncertainty in the NEMO ocean model. *Geoscientific Model Development*, *8*, 1285–1297. https://doi.org/10.5194/gmd-8-1285-2015

Buehner, M. (2005). Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quarterly Journal of the Royal Meteorological Society*, *131*, 1013–1043. https://doi.org/10.1256/qj.04.15

Buizza, R., Miller, M., & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *125*, 2887–2908. https://doi.org/10.1002/qj.49712556006

Buongiorno Nardelli, B., Tronconi, C., Pisano, A., & Santoleri, R. (2013). High and ultra-high resolution processing of satellite sea surface temperature data over Southern European seas in the framework of MyOcean project. *Remote Sensing of Environment*, *129*, 1–16. https://doi.org/10.1016/j.rse.2012.10.012

Burgers, G., van Leeuwen, P. J., & Evensen, G. (1998). Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, *126*, 1719–1724. https://doi.org/10.1175/1520-0493(1998)126%3C;1719:ASITEK%3E2.0.CO;2

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, *16*(5), 1190–1208. https://doi.org/10.1137/0916069

Castor, K., Gerstoft, P., Roux, P., Kuperman, W. A., & McDonald, B. E. (2004). Long-range propagation of finite-amplitude acoustic waves in an ocean waveguide. *The Journal of the Acoustical Society of America*, *116*(4), 2004–2010.

Clayton, A. M., Lorenc, A. C., & Barker, D. M. (2013). Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, *139*, 1445–1461. https://doi.org/10.1002/qj.2054

Collins, M. D. (1993). A Split-step Padé solution for the parabolic equation method. *The Journal of the Acoustical Society of America*, *93*, 1736–1742.

Collins, M. D. (1994). Generalization of the split-step Padé solution. *The Journal of the Acoustical Society of America*, *96*, 382–385.

Collins, M. D., Cederberg, R. J., King, D. B., & Chin-Bing, S. A. (1996). Comparison of algorithms for solving parabolic wave equations. *The Journal of the Acoustical Society of America*, *100*, 178–182.

Courtier, P. (1997). Variational methods. *Quarterly Journal of the Royal Meteorological Society*, *75*, 211–218.

Donlon, C. J., Martin, M., Stark, J. D., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2011). The Operational Sea Surface Temperature and Sea Ice analysis (OSTIA). *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2010.10.017

Dosso, S. E., Giles, P. M., Brooke, G. H., McCammon, D. F., Pecknold, S., & Hines, P. C. (2007). Linear and nonlinear measures of ocean acoustic environmental sensitivity. *The Journal of the Acoustical Society of America*, *121*, 42–45. https://doi.org/10.1121/1.2382719

Errico, R. M., Yang, R., Privé, N. C., Tai, K.-S., Todling, R., Sienkiewicz, M. E., & Guo, J. (2013). Development and validation of observing-system simulation experiments at NASA's Global Modeling and Assimilation Office. *Quarterly Journal of the Royal Meteorological Society*, *139*, 1162–1178. https://doi.org/10.1002/qj.2027

Fofonoff, N. P., & Millard, R. C. (1983). Algorithms for computation of fundamental properties of seawater, UNESCO Technical Papers in marine science, 44, Paris. available at. https://unesdoc.unesco.org/ark:/48223/pf0000059832

Fortin, V., Abaza, M., Anctil, F., & Turcotte, R. (2014). Why should ensemble spread match the RMSE of the ensemble mean? *Journal of Hydrometeorology*, *15*, 1708–1713.

Halliwell, G. R. Jr., Kourafalou, V., Le Henaff, M., Atlas, R., & Shay, L. K. (2015). OSSE impact analysis of airborne ocean surveys for improving upper-ocean dynamical and thermodynamical forecasts in the Gulf of Mexico. *Progress in Oceanography*, *130*, 32–46. https://doi.org/10.1016/j.pocean.2014.09.004

Halliwell, G. R., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Hénaff, M. L., & Atlas, R. (2014). Rigorous evaluation of a fraternal twin ocean OSSE system for the open Gulf of Mexico. *Journal of Atmospheric and Oceanic Technology*, *31*, 105–130. https://doi.org/10.1175/JTECH-D-13-00011.1

Hamill, T. M., & Snyder, C. (2000). A hybrid ensemble Kalman filter-3D variational analysis scheme. *Monthly Weather Review*, *128*, 2905–2919.

Hamilton, E. L., & Bachman, R. T. (1982). Sound velocity and related properties of marine sediments. *The Journal of the Acoustical Society of America*, *72*, 1891–1904. https://doi.org/10.1121/1.388539

Helber, R. W., Barron, C. N., Carnes, M. R., & Zingarelli, R. A. (2008). Evaluating the sonic layer depth relative to the mixed layer depth. *Journal of Geophysical Research*, *113*, C07033. https://doi.org/10.1029/2007JC004595

Hoffman, R. N., & Atlas, R. (2016). Future observing system simulation experiments. *Bulletin of the American Meteorological Society*, *97*, 1601–1616. https://doi.org/10.1175/BAMS-D-15-00200.1

Hopson, T. M. (2014). Assessing the ensemble spread–error relationship. *Monthly Weather Review*, *142*, 1125–1142. https://doi.org/10.1175/MWR-D-12-00111.1

Houtekamer, P., Buehner, M., & De La Chevrotière, M. (2018). Using the hybrid gain algorithm to sample data assimilation uncertainty. *Quarterly Journal of the Royal Meteorological Society*, 1–22. https://doi.org/10.1002/qj.3426

Hursky, P., Porter, M. B., Cornuelle, B. D., Hodgkiss, W. S., & Kuperman, W. A. (2004). Adjoint modeling for acoustic inversion. *The Journal of the Acoustical Society of America*, *115*(2), 607–619. https://doi.org/10.1121/1.1636760

Jensen, F. B., Kuperman, W. A., Porter, M. B., & Schmidt, H. (2011). *Computational ocean acoustics*, (2nd ed.p. 794). New York, U.S: Springer-Verlag. https://doi.org/10.1007/978-1-4419-8678-8

Juricke, S., Palmer, T. N., & Zanna, L. (2017). Stochastic subgrid-scale ocean mixing: Impacts on low-frequency variability. *Journal of Climate*, *30*, 4997–5019. https://doi.org/10.1175/JCLI-D-16-0539.1

Lam, F. P., Haley, P. J., Janmaat, J., Lermusiaux, P. F. J., Leslie, W. G., Schouten, M. W., et al. (2009). At-sea real-time coupled four-dimensional oceanographic and acoustic forecasts during battlespace preparation 2007. *Journal of Marine Systems*, *78*(1), S306–S320.

Large, W.G., & Yeager S.G. (2004). Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies, Technical Report no 460, NCAR Technical Note (2004)

Lei, L., & Hacker, J. P. (2015). Nudging, ensemble, and nudging ensembles for data assimilation in the presence of model error. *Monthly Weather Review*, *143*, 2600–2610. https://doi.org/10.1175/MWR-D-14-00295.1

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., et al. (2018). Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1∕12° high-resolution system. *Ocean Science*, *14*, 1093–1126. https://doi.org/10.5194/os-14-1093-2018

Lermusiaux, P. F. J., Xu, J., Chen, C., Jan, S., Chiu, L. Y., & Yang, Y. (2010). Coupled ocean-acoustic prediction of transmission loss in a continental shelfbreak region: Predictive skill, uncertainty quantification, and dynamical sensitivities. *IEEE Journal of Oceanic Engineering*, *35*, 895–916. https://doi.org/10.1109/JOE.2010.2068611

Lorenc, A. C. (2003). Modelling of error covariances by 4D-Var data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *129*, 3167–3182. https://doi.org/10.1256/qj.02.131

Lurton, X. (2010). *An introduction to underwater acoustics: Principles and applications*, (2nd ed.p. 680). Berlin Heidelberg, Germany: Springer-Verlag.

Madec G, the NEMO team (2012). NEMO ocean engine. Note du Pole de modélisation de l'Institut Pierre-Simon Laplace, France, No 27 ISSN No 1288–1619

Ménétrier, B., & Auligné, T. (2015). Optimized localization and hybridization to filter ensemble-based covariances. *Monthly Weather Review*, *143*, 3931–3947.

Ngodock, H., & Carrier, M. (2014). A 4DVAR system for the Navy Coastal Ocean Model. Part I: System description and assimilation of synthetic observations in Monterey Bay. *Monthly Weather Review*, *142*, 2085–2107. https://doi.org/10.1175/MWR-D-13-00221.1

Ngodock, H., Carrier, M., Fabre, J., Zingarelli, R., & Souopgui, I. (2017). A variational data assimilation system for the range dependent acoustic model using the representer method: Theoretical derivations. *The Journal of the Acoustical Society of America*, *142*(1), 186–194. https://doi.org/10.1121/1.4989541

Oddo, P., Bonaduce, A., Pinardi, N., & Guarnieri, A. (2014). Sensitivity of the Mediterranean Sea level to atmospheric pressure and free surface elevation numerical formulation in NEMO. *Geoscientific Model Development*, *7*, 3001–3015. https://doi.org/10.5194/gmd-7-3001-2014

Oddo, P., Storto, A., Dobricic, S., Russo, A., Lewis, C., Onken, R., & Coelho, E. (2016). A hybrid variational-ensemble data assimilation scheme with systematic error correction for limited-area ocean models. *Ocean Science*, *12*, 1137–1153. https://doi.org/10.5194/os-12-1137-2016

Ota, Y., Derber, J. C., Kalnay, E., & Miyoshi, T. (2013). Ensemble-based observation impact estimates using the NCEP GFS. *Tellus A*, *65*, 20,038. https://doi.org/10.3402/tellusa.v65i0.20038

Palmer, T. N., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G., et al. (2009). Stochastic parametrization and model uncertainty. ECMWF Tech. Memo. 598, 42 pp.

Pecknold, S., & Osler, J. C. (2012). Sensitivity of acoustic propagation to uncertainties in the marine environment as characterized by various rapid environmental assessment methods. *Ocean Dynamics*, *62*(2), 265–281. https://doi.org/10.1007/s10236-011-0497-1

Penny, S. G., Behringer, D. W., Carton, J. A., & Kalnay, E. (2015). A hybrid Global Ocean data assimilation system at NCEP. *Monthly Weather Review*, *143*, 4660–4677. https://doi.org/10.1175/MWR-D-14-00376.1

Pettenuzzo, D., Large, W. G., & Pinardi, N. (2010). On the corrections of ERA-40 surface flux products consistent with the Mediterranean heat and water budgets and the connection between basin surface total heat flux and NAO. *Journal of Geophysical Research*, *115*, C06022. https://doi.org/10.1029/2009JC005631

Pinardi, N., Cavaleri, L., Coppini, G., De Mey, P., Fratianni, C., Huthnance, J., et al. (2017). From weather to ocean predictions: An historical viewpoint. *Journal of Marine Research*, *75*(3), 103–159. https://doi.org/10.1357/002224017821836789

Robinson, A., & Lee, D. (1994). *Oceanography and acoustics: Prediction and propagation models*, (p. 257). Melville, NY: American Institute of Physics, AIP-Press.

Rodwell, M. J., Lang, S. T., Ingleby, N. B., Bormann, N., Hólm, E., Rabier, F., et al. (2016). Reliability in ensemble data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *142*, 443–454. https://doi.org/10.1002/qj.2663

Rouseff, D., & Ewart, T. E. (1995). Effect of random sea surface and bottom roughness on propagation in shallow water. *The Journal of the Acoustical Society of America*, *98*, 3397–3404.

Shapiro, R. (1970). Smoothing, filtering, and boundary effects. *Reviews of Geophysics and Space Physics*, *8*, 359–387.

Storto, A. (2016). Variational quality control of hydrographic profile data with non-Gaussian errors for global ocean variational data assimilation systems. *Ocean Modelling*, *104*, 226–241. https://doi.org/10.1016/j.ocemod.2016.06.011

Storto, A., Masina, S., & Dobricic, S. (2013). Ensemble spread-based assessment of observation impact: Application to a global ocean analysis system. *Quarterly Journal of the Royal Meteorological Society*, *139*, 1842–1862. https://doi.org/10.1002/qj.2071

Storto, A., & Oddo, P. (2019). Optimal assimilation of daytime SST retrievals from SEVIRI in a regional ocean prediction system. *Remote Sensing*, *11*, 2776. https://doi.org/10.3390/rs11232776

Storto, A., Oddo, P., Cipollone, A., Mirouze, I., & Lemieux, B. (2018). Extending an oceanographic variational scheme to allow for affordable hybrid and four-dimensional data assimilation. *Ocean Modelling*, *128*, 67–86. https://doi.org/10.1016/j.ocemod.2018.06.005

Storto, A., Oddo, P., Cozzani, E., & Ferreira Coelho, E. (2019). Introducing along-track error correlations for altimetry data in a regional ocean prediction system. *Journal of Atmospheric and Oceanic Technology*, *36*, 1657–1674. https://doi.org/10.1175/JTECH-D-18-0213.1

Storto, A., & Randriamampianina, R. (2010). Ensemble variational assimilation for the representation of background error covariances in a high-latitude regional model. *Journal of Geophysical Research*, *115*, D17204. https://doi.org/10.1029/2009JD013111

Sutton, P. J., Worcester, P. F., Masters, G., Cornuelle, B. D., & Lynch, J. F. (1993). Ocean mixed layers and acoustic pulse propagation in the Greenland Sea. *The Journal of the Acoustical Society of America*, *94*, 1517–1526.

Trémolet, Y., & Le Dimet, F. X. (1996). Parallel algorithms for variational data assimilation and coupling models. *Parallel Computing*, *22*, 657–674. https://doi.org/10.1016/0167-8191(96)00018-X

Wang, D., Lermusiaux, P. F. J., Haley, P. J., Eickstedt, D., Leslie, W. G., & Schmidt, H. (2009). Acoustically focused adaptive sampling and on-board routing for marine rapid environmental assessment. *Journal of Marine Systems*, *78*, S393–S407.

Wang, X., Snyder, C., & Hamill, T. (2007). On the theoretical equivalence of differently proposed ensemble-3DVAR hybrid analysis schemes. *Monthly Weather Review*, *135*, 222–227.

Weatherall, P., Marks, K. M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J. E., et al. (2015). A new digital bathymetric model of the world's oceans. *Earth and Space Science*, *2*, 331–345. https://doi.org/10.1002/2015EA000107

Weaver, A. T., Deltel, C., Machu, E., Ricci, S., & Daget, N. (2005). A multivariate balance operator for variational ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *131*, 3605–3625. https://doi.org/10.1256/qj.05.119