SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Enhancing regional ocean data assimilation in high and mid latitude European seas
Computer Project Account:	Spitstor
Start Year - End Year :	2019 - 2021
Principal Investigator(s)	Andrea Storto
Affiliation/Address:	CMRE, La Spezia, Italy CNR-ISMAR, Rome, Italy
Other Researchers (Name/Affiliation):	Paolo Oddo Silvia Falchetti Craig Lewis

The following should cover the entire project duration.

Summary of project objectives

(10 lines max)

The project has several objectives, corresponding to the following tasks, in order to improve the data assimilation systems for the regional oceanographic configurations in use over mid and high latitudes.
Task A. Test the feasibility of assimilating small-scale current data collected by HF radars, drifters and ADCP profilers mounted on vessels and buoys;
Task B. Include uni (aka "weakly coupled") and multi-variate (aka "strongly coupled") data assimilation of sea-ice parameters in the Arctic analysis system exploiting the synergy of different observing networks.
Task C. Experiment multi-scale data assimilation in order to simultaneously ingest both the large and the small-scale information collected by gliders during the observational campaign;
Task D. Run ensemble variational experiments with stochastic physics in order to i) retune the background-error covariances for use in data assimilation and ii) provide an ensemble of realizations for forcing downstream acoustic models and characterize the uncertainty and the cross-covariances between physical and acoustic parameters;
Task E. Test optimal ways to assimilate SST observations (L2, also daytime) in the analysis systems, exploiting the synergy with in-situ profiles to verify the methodology, and in particular data from gliders piloted to follow the satellite tracks.

Summary of problems encountered

(If you encountered any problems of a more technical nature, please describe them here.)

Everything worked very well. Fetching old files from the ECFS archive was often extremely slow, thus slowing down the experiment performance. To avoid that this affected the jobs billing, we moved all ecp commands out of the job scripts.

Experience with the Special Project framework

(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

Very good experience. Perhaps a small suggestion could be to postpone the date of reporting, in order to make it correspondent to the yearly project duration (i.e. in December and not June).

Summary of results

(This section should comprise up to 10 pages, reflecting the complexity and duration of the project, and can be replaced by a short summary plus an existing scientific report on the project.)

The projects has allowed us to consolidate the optimal assimilation of in-situ and remotely sensed data, with particular focus on the coupled physical-acoustic environment and the adoption of techniques inherited from deep learning to substitute some pieces of the oceanic variational data assimilation scheme that we use.
In particular, the project consisted in several tasks, each with different scientific questions, that we briefly review in the following pages.

1) Optimal Assimilation of SST data.

Exploiting the potential of space-borne oceanic measurements to characterize the sub-surface structure of the ocean becomes critical in areas where deployment of in situ sensors might be difficult or expensive. Sea Surface Temperature (SST) observations potentially provide enormous amounts of information about the upper ocean variability. However, the assimilation of daytime SST retrievals, e.g., from infrared sensors into ocean prediction systems, requires a specific treatment of the diurnal cycle of skin SST, which is generally under-estimated in current ocean models due to poor vertical resolution at the air–sea interface and lack of proper parameterizations (see Figure 1a for a schematic representation of the problem).
To this end, a simple off-line bias correction scheme is proposed, where the bias predictors include, among others, the warm layer and cool skin warming/cooling deduced from a prognostic model. The selection of the bias predictors has been performed through Relevance Vector Machine (RVM). Furthermore, a localization procedure that limits the vertical penetration of the SST information in a hybrid variational-ensemble data assimilation system is formulated in Empirical Orthogonal Function (EOF) space, which is the space used by our data assimilation scheme to perform the variational minimization.
These two novelties are implemented and assessed within a regional ocean prediction system in the Ligurian Sea for the assimilation of daytime SST data retrieved with hourly frequency from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the geostationary satellite Meteosat-10. Experiments are validated against independent measurements collected by gliders, moorings, and drifters during the Long-term Glider Missions for Environmental Characterization (LOGCMEC17) sea trial. Results suggest that the simple bias correction scheme is effective in improving both the sea surface and mixed layer accuracy (see Figure 1b for an improvement of the forecasted diurnal cycle), correctly thinning the mixed layer compared to the control experiment (see Figure 1c to see how the SST assimilation impacts the mixed layer depth evolution), outperforming experiments with night-only data assimilation, and improving the forecast skill scores. Localization further improves the prediction of the mixed layer depth. It is therefore recommended that sophisticated bias correction and localization procedures are adopted for fruitfully assimilating daytime SST data in operational oceanographic analysis systems.
Details of this research have been published in Storto et al. 2019b (https://doi.org/10.3390/rs11232776)

2) Assimilation of currents

Exploiting the potential of seawater velocity data from different observing networks has the potential of improving the small-scale oceanic variability, which is in turn essential for a number of applications. We have assessed the impact of assimilating such data in the Ligurian Sea model and in the nested Gulf of La Spezia model. The results indicate that the assimilation of current data improves the model validation skill scores and enhances the Liguro-Provencal current system, which in turn improves the vertical stratification of the ocean in the nested domain in comparison with independent observations from Scanfish. These promising results pave the way for real-time assimilation of velocity data and foster additional experiments, in particular in order to better understand the impact of the temporal frequency of data assimilation. In Figure 2, we show current speed RMSE (as a function of depth) in the experiment where all current data are assimilated (BASE+UV) against the experiment where only in-situ profiles and altimetry data have been assimilated (BASE), proving the effectiveness of the data assimilation procedure.

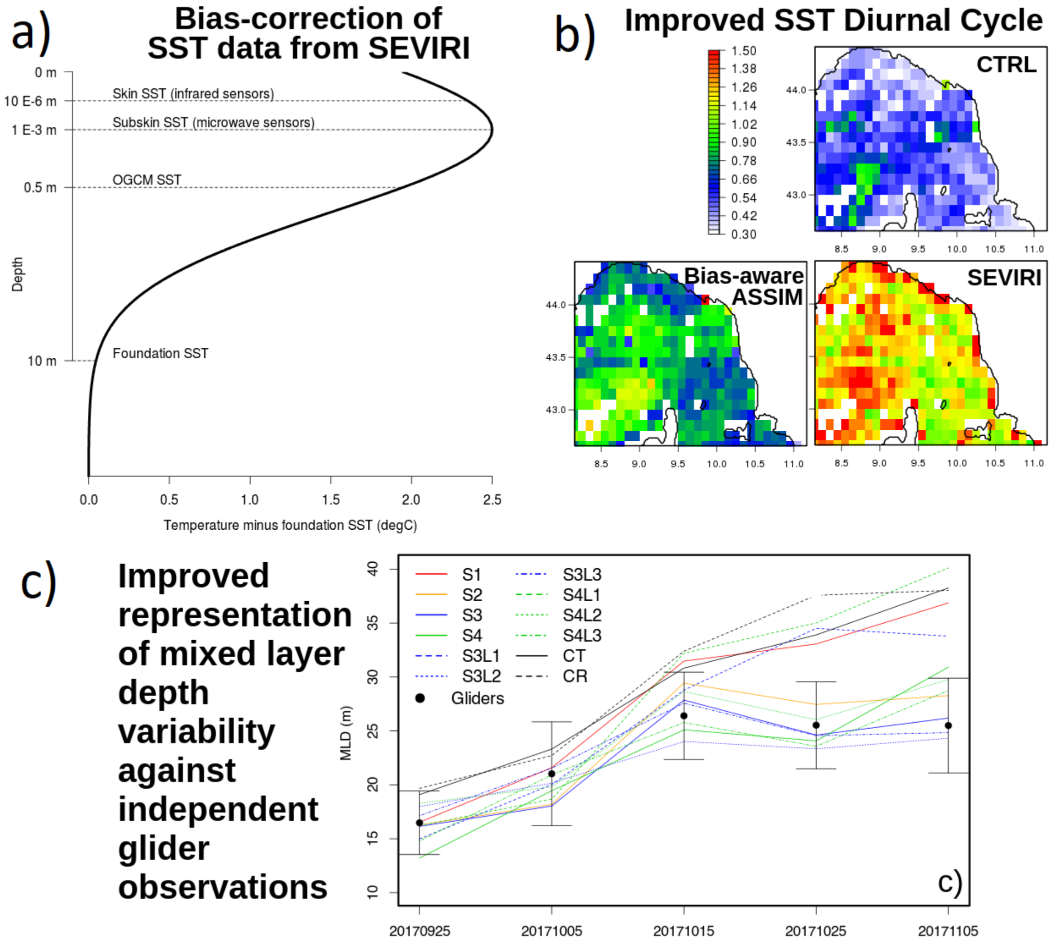

***Figure 1**. a) Typical profile of a temperature anomaly with respect to foundation SST for low-wind conditions. (Adapted from GHRSST). The depth axis is logarithmic. b) Averaged diurnal amplitude during the period 20171015–20171114, for two selected experiments in observation space compared to hourly SEVIRI data. The diurnal amplitude is calculated as the difference between maximum and minimum sea surface temperature over a day from hourly data, only for days with at least 12 hourly valid SEVIRI records. Model data are first interpolated to observation location/time to provide consistent assessment. c) Mixed layer depth time series from the experiments (in observation space, i.e., averaged over the glider profile locations) and from the two gliders. Data are 10 day averages, centered on the date shown on the x-axis, with standard deviation plotted as a vertical bar around the averaged observed values.*

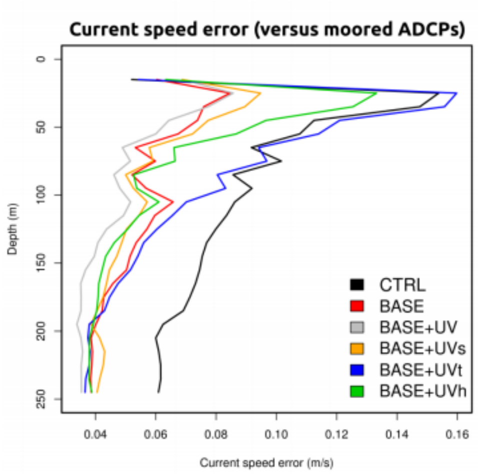

***Figure 2**. RMSE profiles of current speed against moored ADCP data in experiments assimilating current data from moored and ship-borne ADCPs, HF radars and drifters.*

3) Stochastic physics and ensemble-variational data assimilation schemes

Generating optimal perturbations is a key requirement of several data assimilation schemes. We have developed a new stochastic physics package for ocean models, implemented in the NEMO ocean general circulation model. The package includes three schemes applied simultaneously: stochastically perturbed parametrization tendencies (SPPT), stochastically perturbed parameters (SPP) and stochastic kinetic energy backscatter (SKEB) schemes. The three schemes allow for different temporal and spatial perturbation scales. Within a limited-area ocean model configuration, ensemble free-running simulations were performed to assess the impact and reliability of the schemes. They prove complementary in increasing the ensemble spread at different scales and for different diagnostics. The ensemble spread appears reliable; for instance, it proves consistent with the root-mean-square differences with respect to higher-resolution (sub-mesoscale) simulations that here represent the “truth” (in the sense that it includes “unresolved physics”). Interestingly, both the SPPT and the SKEB schemes lead to an increase of eddy kinetic energy at small spatial scales (2–10 km), and contribute to modify the ensemble mean state, mitigating warm biases near the thermocline due to the enhancement of the upper ocean vertical mixing (see Figure 3). The schemes have been later implemented and tuned in the global NEMO configurations in use at ECMWF, within the C3S_321b project.
As an application of the stochastic packages, the ensemble anomaly covariances coming from the ensemble free-running simulations were used to feed large-scale anisotropic covariances that complement smaller-scale ones in a hybrid-covariance regional analysis and forecast system in the Mediterranean Sea. Ensemble-derived covariances are formulated as slowly varying three-dimensional low-resolution empirical orthogonal functions (EOFs), which is a relatively new formulation of background-error covariances (opposed to standard horizontal versus vertical separation). The improvements due to the addition of such covariances to the stationary ones are found significant in real-data experiments, within verification skill scores against glider profile data, remotely sensed observations and current speed measurements from drifters, radar and moorings.

4) Coupled ocean-acoustic forecasting

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, we have conducted integrated ocean-acoustic Observing System Simulation Experiments (OSSE). 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 were also 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. Figure 4 shows an example of verification skill scores (TL at 75 Hz and at 30km range), indicating the benefits of more complex data assimilation systems. These findings foster developments of coupled data assimilation for operational underwater acoustic propagation.

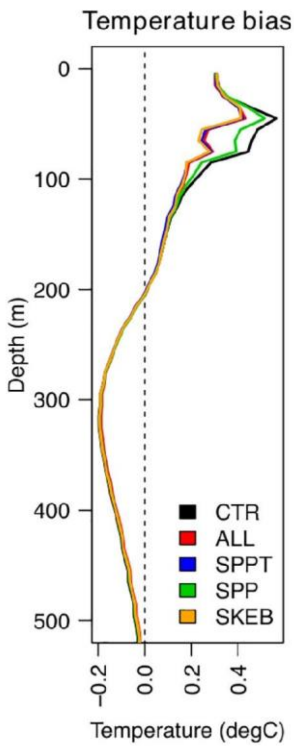

Figure 3. *Stochastic physics schemes are found to mitigate warm biases around the thermocline w.r.t. unperturbed runs, due to enhanced vertical mixing during. The bias is the averaged difference of glider observations minus model-equivalents. CTR: Unperturbed; SPPT, SPP, SKEB implement the corresponding stochastic scheme; ALL: the three perturbations schemes are applied together. Results are from a 3-month free-running NEMO simulations in the Ligurian Sea.*

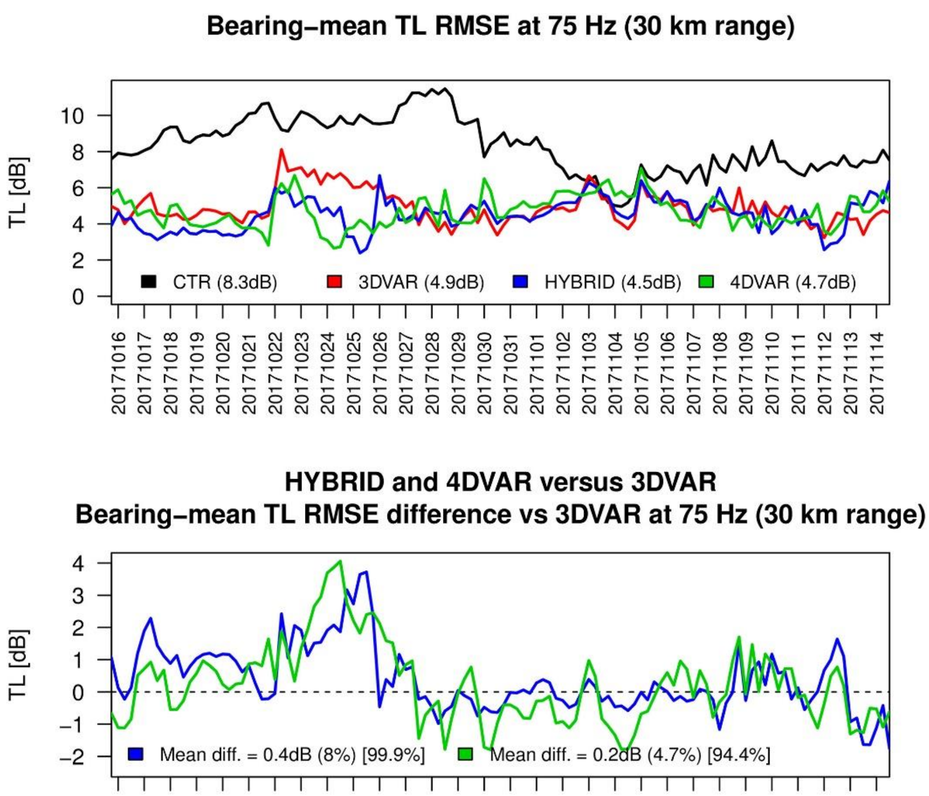

Figure 4*. Bearing mean 30 km transmission loss RMSE against the synthetic acoustic observations for the 75 Hz configuration, for the four experiments CTR, 3DVAR, HYBRID, and 4DVAR (top panel).*

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.

5) Use of neural networks in data assimilation to replace observation operators

Variational data assimilation requires implementing the tangent-linear and adjoint (TA/AD) version of any operator. This intrinsically hampers the use of complicated observations. We have assessed a new data-driven approach to assimilate acoustic underwater propagation measurements [transmission loss (TL)] into a regional ocean forecasting system. TL measurements depend on the underlying sound speed fields, mostly temperature, and their inversion would require heavy coding of the TA/AD of an acoustic underwater propagation model. The nonlinear version of the acoustic model is applied to an ensemble of perturbed oceanic conditions. TL outputs are used to formulate both a statistical linear operator based on canonical correlation analysis (CCA), and a neural network–based (NN) operator. For the latter, two linearization strategies are compared, the best-performing one relying on reverse-mode automatic differentiation, natively provided by the Tensorflow library (compared to numerical differentiation through Richardson’s extrapolation). The new observation operator is applied in data assimilation experiments over the Ligurian Sea (Mediterranean Sea), using the observing system simulation experiments (OSSE) methodology to assess the impact of TL observations onto oceanic fields. TL observations are extracted from a nature run with perturbed surface boundary conditions and stochastic ocean physics. Sensitivity analyses indicate that the NN reconstruction of TL is significantly better than CCA. Both CCA and NN are able to improve the upper-ocean skill scores in forecast experiments, with NN outperforming CCA on the average (see Figure 5 for the skill score profiles). The use of the NN observation operator is computationally affordable, and its general formulation appears promising for the adjoint-free assimilation of any remote sensing observing network.

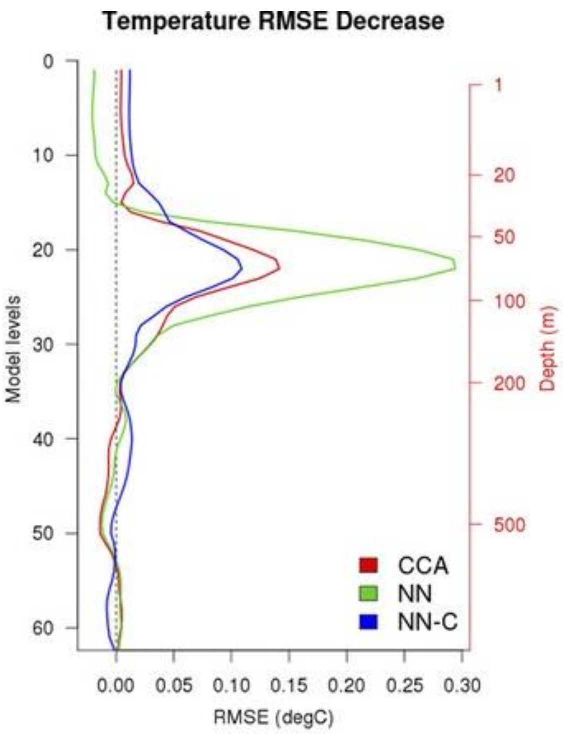

***Figure 5**. temperature RMSE difference profiles with respect to the Ctrl experiment*

List of publications/reports from the project with complete references

1. Storto, A., P. Oddo, E. Cozzani, and E. F. Coelho, 2019: Introducing Along-Track Error Correlations for Altimetry Data in a Regional Ocean Prediction System. J. Atmos. Oceanic Technol., 36, 1657–1674, https://doi.org/10.1175/JTECH-D-18-0213.1
2. Storto, A.; Oddo, P. Optimal Assimilation of Daytime SST Retrievals from SEVIRI in a Regional Ocean Prediction System. Remote Sens. 2019, 11, 2776, doi:10.3390/rs11232776
3. Storto, A., Falchetti, S., Oddo, P., Jiang, Y.- M., & Tesei, A. (2020). Assessing the Impact of Different Ocean Analysis Schemes On Oceanic and Underwater Acoustic Predictions. Journal of Geophysical Research: Oceans, 125, e2019JC015636. https://doi.org/10.1029/2019JC015636
4. Storto, A, Andriopoulos, P. A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. Q J R Meteorol Soc. 2021; 1691– 1725. https://doi.org/10.1002/qj.3990
5. Storto, A., De Magistris, G., Falchetti, S., & Oddo, P. (2021). A Neural Network–Based Observation Operator for Coupled Ocean–Acoustic Variational Data Assimilation, Monthly Weather Review, 149(6), 1967-1985. Retrieved Jun 11, 2021, from https://journals.ametsoc.org/view/journals/mwre/149/6/MWR-D-20-0320.1.xml

Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new Special Project.)

We are not involved in other special projects, but we plan to submit an application to continue the project, with several other tasks and different scientific questions.