

Authors' responses

We thank to the editor and the anonymous reviewers for comments motivating a revision of our paper which we are ready to perform based on their suggestions. We provide point-by-point responses to the reviewers' comments in the following, supplying the proposed updates and with few additional figures. Throughout this document, bold and italic fonts are used for the captions and reviewers' comments, respectively. Our responses are in normal fonts.

Authors' response to RC#1

Major comments

1. Temperature and salinity observations are assimilated and the evaluation of impacts is made only with temperature and salinity. The assimilation of temperature and salinity observations, however, impacts all other model fields. The evaluation of the observational impacts should be extended at least to currents.

In order to discuss this point, we provide a figure (Fig. 13) comparing salinity overlaid with current fields at 5 m. depth, obtained from experiments FB001 (left) and FB002 (right) for the exemplary case on 7 Jan 2009 at 00:00. The salinity differs significantly between the two experiments especially along the southeastern coast, while there is very little change in both qualitative and quantitative terms in the horizontal circulation, namely the current speed and direction, in the affected region that can be attributed to data assimilation. The effect of data assimilation is more pronounced in terms of the property fields, which alternatively indicates changes in stratification and vertical mixing along the southern coast.

Moreover, the same figure, but for the nature run is provided in the appendix as requested by the major comment A of the Reviewer#2. The circulation in the nature run (Fig. A2, introduced later) appear more intense compared to FB001 and FB002, however, without resulting in a significant change in the horizontal circulation patterns.

We can deduce that the impact of the assimilation on the circulation seems very small, if not negligible, compared to the differences of the experiments with the nature run.

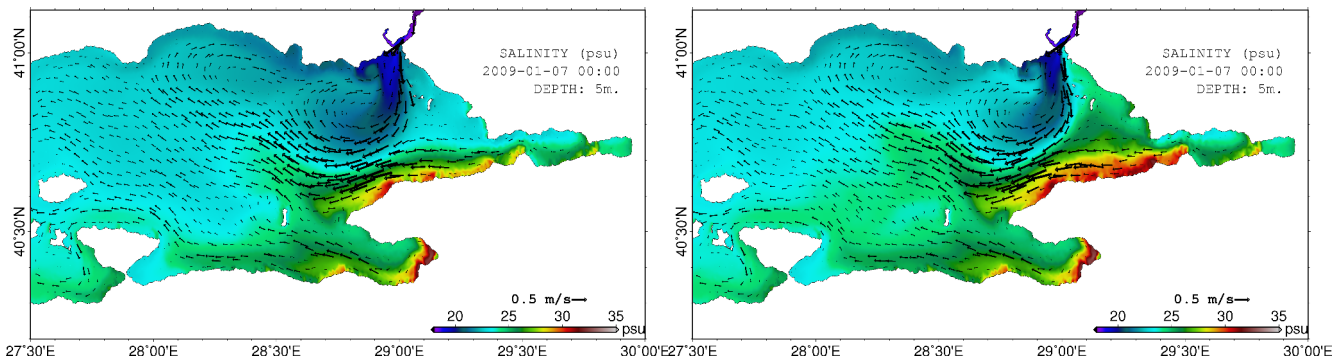


Fig. 13. Comparison of the 5 m. salinity fields on 7 January 2009 at 00:00 between FB001 (left) and FB002 (right). The corresponding circulation patterns are shown by current vectors.

2. It seems that using the Ferrybox system for observing temperature and salinity near the surface is more feasible than using floats or gliders. On the other hand, does the improved estimate of temperature and salinity fields near the surface significantly improve the support for the most important applications of oceanographic forecasts in the Marmara Sea?

We agree with this comment in general. We will introduce a small paragraph in the revised text, explaining the needs and our rationale as expressed in the following.

The Marmara Sea has unique dynamics of its circulation, generated by volume fluxes through the straits, interaction with the atmosphere and buoyancy effects in a strongly stratified environment. All these factors play crucial roles in the dynamical response of the system. Black Sea and Aegean Sea water masses transported through the Marmara Sea determine its vertical structure, which in turn impacts its internal dynamics. In principle, all of the above influences on the circulation dynamics have to be tested, by considering the individual and combined effects of the assimilation of different types of data.

To begin with the present study, we only considered elementary water properties observations that relatively easily could be obtained from available platforms. The present OSSE only attempts to initiate a first and essential step in the much needed extended studies of advanced modelling as well as data assimilation. We also note that there are not many near-real-time observations available in the Marmara Sea at present; building the necessary infrastructure to incorporate various other types of observations still needs further serious efforts.

The proposed initial observing system should result in better forecasts in terms of water properties of the upper layer, which promises to improve forecasts in the Aegean Sea. To further improve the forecasts, the assimilation of data on water properties, sea level and currents measured by floats and fixed stations (e.g. ADCPs, tide gauges and/or altimeter measurements), the use of these measurements to better estimate volume fluxes through the Bosphorus and Dardanelles Straits would be in order. Appropriate use of such extended measurements tested by continued OSSE's could also impact better estimates of the lower layer circulation driven by the density gradients.

3. There are many technical details about specific solutions implemented in the data assimilation scheme, but impacts those solutions are not tested in the study.

While implementing our solutions in this study, we benefited from the experience provided by similar studies in the literature, although we carefully adopted a version needed in our application. We agree with the reviewer that the impact of the present choices and other possible solutions should be tested, possibly by us and/or others in continuing studies in the region.

4. I think that the style of writing should be improved. It is very difficult to read and interpret many sentences providing important information.

We will do our best to improve the text in the revised version. The differences will be demonstrated by comparing the discussion paper and revised version.

Minor comments

1. Page 2, line 2: The “high resolution” of what?

The original phrase is “Until recently, the need for high resolution in the straits made it infeasible to model the complete TSS due to the computational cost.”

We clarify the phrase as “Until recently, building a model solving for the hydrodynamics of the complete TSS was not considered a feasible undertaking, implying high computational costs of the required horizontal and vertical resolution enabling to represent the sharp stratification and the extremely complex topography of the straits and shelf regions, largely differing from those of the larger neighboring basins.”

2. Page 2, line 3: What are “integral models”?

“.. integral models of the system ...”

We rephrase the phrase above as “...models of the whole system...”

3. *Page 2, line 12: The OSSE abbreviation is introduced, but it is not explained.*

We change it as “We follow the Observing System Simulation Experiments (OSSE) ...”

4. *Page 3, line 14: same as comment 2.*

“The high complexity of the system requires integral modeling approaches to represent the links between its different compartments.”

We replace the above sentence with “The high complexity of the system requires models that simultaneously solve for the whole system in its smallest resolved details and optimally representing multiple scales of interest.”

5. *Page 3, lines 19-20: This sentence is not related to the scope of the study.*

We omit the sentence “It was developed by the Alfred Wegener Institute as the first global ocean model using an unstructured mesh.”

6. *Page 4, line 11: What is “covariance information” in this context?*

We replace the term “information” with “distribution” to be more precise. The discussion addresses the issues related the prior distribution after resampling the ensemble to compute the covariance. The EAKF preserves the prior covariance distribution during sampling. The ensemble covariance is used to quantify the relation between pairs of state variables or an observation and a state variable in the linear Gaussian context of the ensemble Kalman filter.

7. *Page 4, line 12: What is “prior information” in this context?*

The same response as to comment #6. We replace the term “information” with “distribution”

8. *Page 4, line 13: Which covariances are updated?*

We replace the term with “analysis covariances”.

9. *Page 4, lines 20-24: Is vertical diffusion the most important for correctly simulating the depth of the interface? Vertical diffusion should be governed by slowly varying large-scale fields and perturbing it at the high frequency may add processes that may be unphysical.*

The depth of the interface in the TSS is mainly determined by the volume fluxes through the straits at seasonal time scales. On daily time scales, however, the wind forcing may alter the interface depth significantly, especially when there is a severe storm passage over the system (Book et al., 2014). Moreover, enhanced vertical mixing around the interface may contribute to the water exchange between strongly stratified upper and lower layers by entrainment processes (Özsoy et al. 2001).

We agree with the reasoning of the reviewer on vertical diffusion. However, what we perturb is the background vertical diffusivity K_v which is at least two orders of magnitudes smaller than the spatially varying vertical diffusivity K_v . We are aware that more care is needed for longer experiments. However, within the period of the experiments, we haven’t identified any instability developing due to the growth of the perturbation.

10. *Page 4, line 35: Localization may introduce strong dynamical imbalances. This contrasts the sentence on lines 11-13.*

This is correct. However, since we can only afford 30 ensemble members, localization is required for a system that produces forecasts with small RMSE. The localization half-width radius has been chosen

empirically to minimize RMSE while maintaining a prior RMSE to spread ratio that is approximately unity. A larger ensemble size would allow us to use less stringent localization, further reducing RMSE while also reducing dynamical imbalances. The fact that the assimilation cycle is stable over a number of days suggests that the dynamical imbalance is not large enough to dominate the balanced model dynamics.

11. *Page 6, lines 29-30: I do not understand this sentence.*

“This approach was chosen because FESOM in the Marmara Sea is sensitive to equally plausible salinity boundary conditions.”

We prefer to add an appendix to document the properties of the nature run and its difference from the forward model as requested by Reviewer#2. Very briefly, the nature run is supposed to be the best realization of the system whereas forward model is chosen as a different model or the same model with different resolution or different physics. In our study, different surface salinity boundary conditions have been implemented in the two configurations of the same model constituting the forward model and the nature runs. Both solutions appear realistic and equally plausible but different from each other, displaying sensitivity to the applied boundary conditions.

12. *Page 9, line 3: Argo floats and gliders are not instruments.*

Rephrased the following part “it is not easy to deploy instruments such as argo or glider since there is heavy ship traffic.” to read as follows: “it is not easy to deploy, for instance, argo floats or gliders close to the surface since ...”

13. *Page 9, lines 18-20: How is the sampling rate of 1 minute obtained output frequency of 1 hour.*

In the manuscript, it is explained as “... from hourly NR outputs at varying spatial location along the track of the ferries.” That means the synthetic observations falling into the same one hour interval are sampled from the same output but on different locations mimicking the motion of the ferries in time. We modify the above phrase as “...from hourly NR outputs at varying spatial locations so as to remain within one minute time intervals along the track of the ferries.”

14. *Page 9, line 24: What is the meaning of stochastic in this context?*

In stochastic EnKF, the observations are perturbed for each ensemble member before assimilation. In this work, the EAKF scheme used in the present work is deterministic since observations are identical for each ensemble member.

15. *Page 9, line 25: Do you want to say that errors of observations are uncorrelated?*

Obviously, repeated observations from the same instrument are expected to have some correlated error component. Here, we neglect the correlated error, as is done in many geophysical assimilation problems, because of the expense and difficulty of dealing with it explicitly. Employing an accurate estimate of the correlated error, or explicitly modeling the correlated error is difficult but would certainly lead to some improvement in our forecast fits to observations. Note that temperature and salinity errors are from unique instruments and might be expected to have less correlated error. Here, temperature and salinity errors are assumed to be uncorrelated and they are set different following an existing observing system presented by Grayek et al. (2011).

16. *Page 10, line 16: Updates are smaller than what? It looks like temperature and salinity are compared by magnitude, but they are two different physical parameters.*

The salinity updates at 20 m. are smaller than the salinity updates in the upper layers, although we refrain from presenting an extra figure in order to reduce clutter. Since the corrections around 20 m. are generally smaller than those at 5 m. for both temperature and salinity and we only choose to show the updates for temperature at 20 m.

17. Page 11: The bottom paragraph should be reformulated by using the correct terminology.

We rewrite the paragraph is as:

“The DART offers tools to control the data used without any difficulty. Each type of observation can either be assimilated or withheld to evaluate the resulting analyses. DART also provides a rudimentary quality control capability that can reject observations that are too different from the ensemble mean prior estimate. Synthetic observations are not used in the assimilation (rejected) if $x_b - y > TE(x_b - y)$. Here, y is the observation and x_b is the corresponding prior mean of the ensemble of forward operators. The expected value, E , is computed as:

$$E(x_b - y) = \sigma_{x_b}^2 + \sigma_y^2 \quad (1)$$

where σ stands for the standard deviation. T is chosen as 3 for both temperature and salinity in these experiments.”

18. Page 12, lines 1-4: I do not understand why there are outliers in an OSSE experiment? Are observations wrong, or some assumptions are not valid?

The outliers are completely an outcome of our choice considering the realistic errors in the Marmara Sea. In Aydoğdu et al. (2018), validation of the long-term simulations has been performed and presented. Comparison with the CTD observations shows that the errors may increase up to 3 psu or 3 °C, for salinity and temperature respectively, in the area of interest but not more. Therefore, we set a limit for the synthetic observations and do not assimilate them if they exceed these error ranges with respect to the prior ensemble mean. The difference between the nature run and ensemble mean can be high depending on the perturbations applied, therefore, there are some observations which we mark as outliers. The approach is chosen to stick with the realistic applications of the data assimilation where a quality check is generally required.

19. Page 15, line 4: Innovations are better than what?

We compute the innovation regardless if the observations are assimilated or not. Here, we compare the FB001 and FB002 cases, and conclude that the difference between the synthetic observations and the corresponding ensemble prior mean is smaller in FB002, implying improved forecast as a result of the assimilation.

20. Fig. 11: Salinity differences have very large gradients. I suspect that they form strong density gradients impacting currents. Currents should be included in the evaluation of the assimilation.

We added a new figure (Fig. 13) to compare the velocity fields in day 7. We couldn't identify any strong impact of the assimilation on the currents as discussed in our response to major comment #1.

Authors' response to RC#2

Major comments:

A. Unless this has been published elsewhere, it would be useful to have a brief analysis of the physical situation in the NR at the time of the assimilation experiments, if possible

Hopefully, these details will be published in another manuscript that is presently submitted to Ocean Science (Aydoğdu et al. 2018 in the references). Yet, to enable a review of the earlier results, we prefer to provide an appendix using the following figures along with further technical details of the nature run.

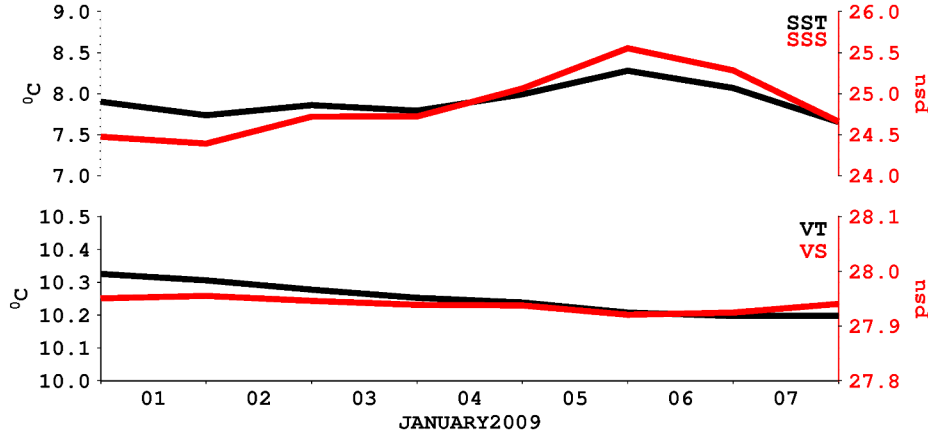


Fig. A1 Daily mean of sea surface temperature (SST) and salinity (SSS) and volume temperature (VT) and salinity (VS) in the Marmara Sea.

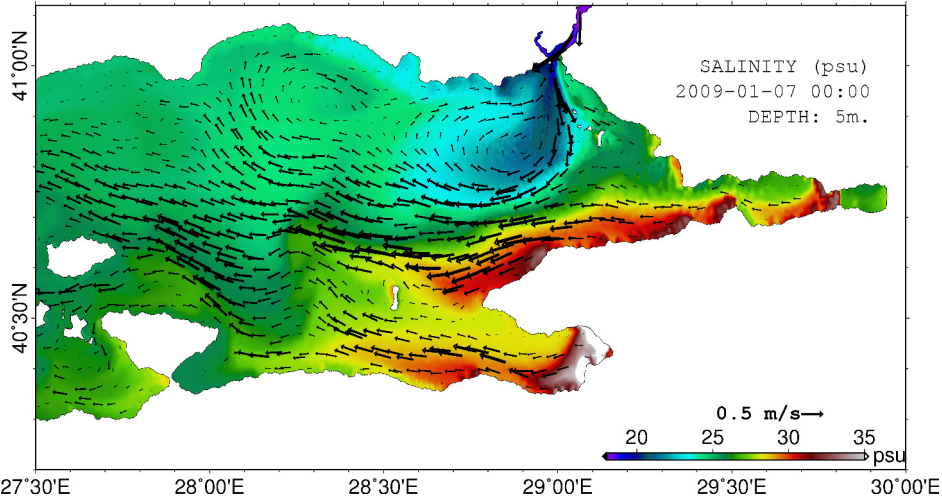


Fig. A2 5 m. salinity fields in 7 January 2009 at 00:00 simulated in the nature run. The corresponding circulation is overlaid by arrows. Corresponding fields for the FB001 and FB002 experiments are shown in Fig. 13 for comparison.

B. I had an overarching question in my mind throughout my reading of the manuscript: is the combined effect of i.c. perturbations (a mix of short-term time lag and interannual variability) and diffusivity perturbations able to explain at least part of the model-data differences, within observational error? This question can be posed for (B1) simulated data

The B1 question is a question of consistency of the innovations with Ensemble spread + obs error: it is partially covered in the ms. (e.g. Fig.8), but not exploited.

Coming back to question B1: I have been frustrated that Fig.8 shows RMSE (of the dimension of the variable) and spread (of the dimension of the variable, *squared*). It would have been better to show *MSE* (not RMSE) and spread (assuming that this is *prior* Ensemble spread): then you could have

tested whether the (prior) innovation variance was more or less of the same order as the (prior) "Total spread" (= your estimate of prior error + your estimate of obs error). I did the squaring visually, and the orders of both quantities do not match each other, especially for salinity. I believe that this should even briefly be discussed.

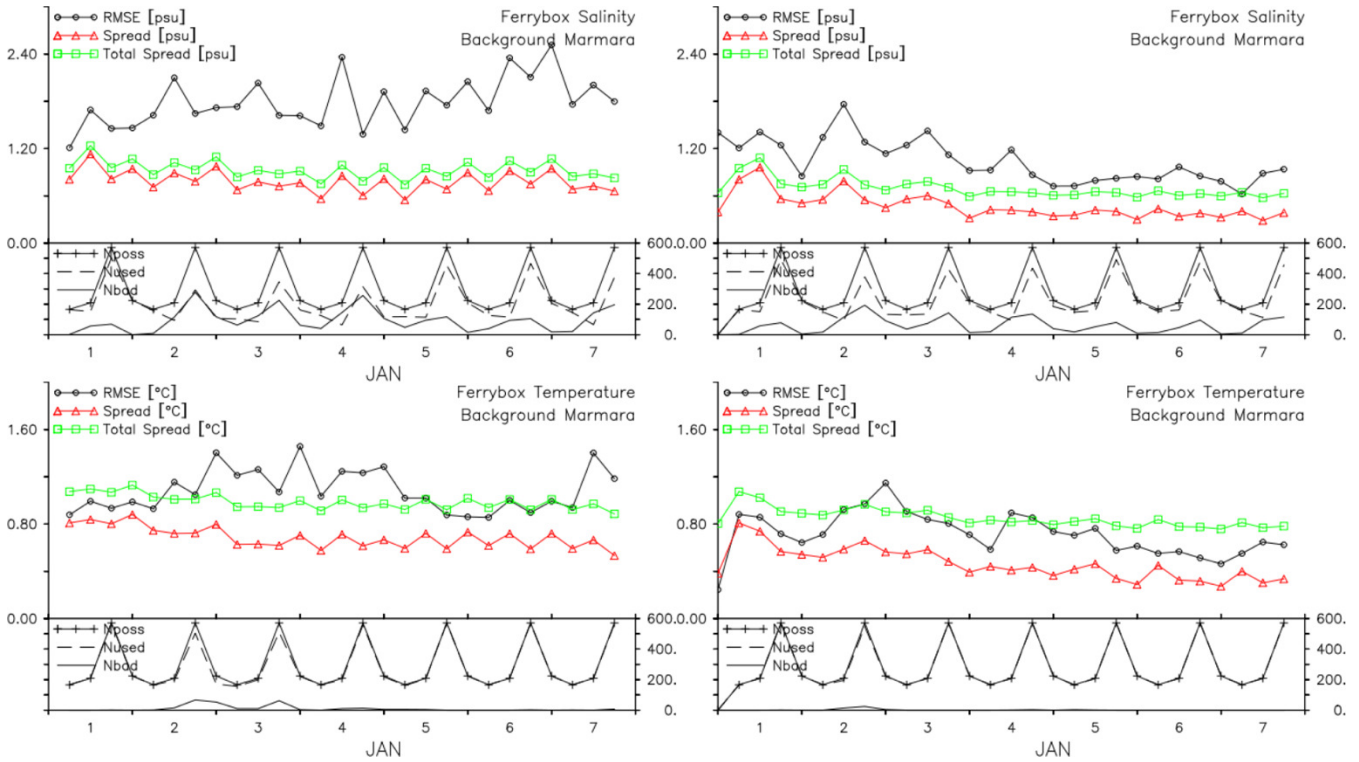


Fig. 8 Time series of mean squared innovations, spread and total spread of salinity (top) and temperature (bottom) for FB001 (left) and FB002 (right). The statistics are computed in the location of the observations used in the corresponding assimilation cycle. The spread is the square root of the variance. Total spread is the observational error added to the spread. Y-axis shows the range of each statistics indicated in the legend. Bottom panel of each figure shows the number of available observations (Nposs) in each assimilation cycle. For the assimilation experiment, FB002, Nused and Nout show the number of assimilated observations and outliers, respectively. For the experiment without assimilation, FB001, they are the number of observations which would be assimilated or rejected, respectively, in that specific assimilation cycle if assimilation was performed.

We agree and thank to the reviewer. We checked the calculation of the spread and found out that it is the standard deviation, not the variance. The statistics are accurate but units should be in psu and °C, for the salinity and temperature respectively. Therefore, they are comparable with the corresponding RMSE. We corrected the mistake on the legend as well as in the caption and provide the revised version of the Fig. 8. The resulting total spread is therefore, the square root of the sum of the variance and observation error. We will revise the manuscript accordingly.

The ensemble spread is initially provided by the perturbation of the initial temperature and salinity fields and maintained by the background vertical diffusivity perturbation. However, the growing RMSE points out different salinity boundary conditions chosen for the forward model and the nature run. In another word, the perturbations maintaining the spread are capable of sustaining it even the observations are assimilated. On the other hand, assimilation will significantly reduce the errors due to the physical processes simulated by chosen model schemes. We will revise the manuscript accordingly.

(B2) any existing real observations (e.g., SST).

The B2 question is about the realism of errors: it was not covered in the ms (only the conclusion mentions "lack of data").

The error assessment of the model using CTD observations is performed in the Marmara Sea and presented in another study (Aydoğdu et al 2018). We think the errors in the present work compare well with those and are therefore expected to be realistic. Lack of data is mentioned only for data assimilation purposes. We will summarize this aspect in the appendix that we shall provide according to what has been implied in major comment #A.

C. Why did you limit yourself to 7 days? Some of the error processes, especially those associated with mixing and stratification, could act on longer time scales.

The main challenge to perform longer experiments is commensurate with the computational cost of running ensembles using such a high-resolution model. However, given the present results, we are motivated to perform longer experiments which take into account also the suggestions addressed by the reviewers.

D. The localization cut-off scales are very short. Can't this trigger fast unphysical responses, for instance via temperature-sea level covariances?

Agree, that is possible. However, the cut-off radius has been chosen to minimize RMSE while maintaining a prior RMSE to spread ratio that is approximately unity. On the other hand, the spurious correlations related to larger cut-off radius may also trigger unphysical behaviors in energetic basin such as the Marmara Sea. But we haven't identified any unphysical behavior within the period of the experiments related to dynamical imbalances.

Minor comments:

1. *Page 4, lines 11-12: How is "prior covariance information" related to "dynamical balances"?*

We replace the term “information” with “distribution” to be more precise. The ensemble covariance is the statistical representation of the prior constraints (‘dynamical balances’) generated by the model. For a linear model, a sufficiently large ensemble would be able to represent all the balances. However, our model is nonlinear and cost forces us to use a small ensemble. This means that the ensemble sample covariances do not exactly represent the dynamical balances. The ensemble filter algorithm and inflation we use are designed to maintain the prior ensemble covariance as much as possible. The benefits of preserving the prior distribution are extensively discussed in Anderson (2001) compared to ensemble filters such as kernel or Gaussian resampling filters. However, given the errors in the sample covariances, localization is effective in reducing forecast RMSE compared to observations. Localization also disrupts the prior sample covariance, so can result in reduced dynamical balance.

2. *Page 4, lines 11 and 13: "the covariances are updated in every assimilation cycle": Isn't this contradictory with "it preserves the prior covariance information"?*

We replace the term information with distribution. Analysis covariances are updated in every assimilation cycle.

3. *Page 4, line 25: The generally adopted procedure to perturb diffusivity parameters does not use a centered Gaussian pdf*

We were not aware of it. We thank to the reviewer for the information and will further dig in the literature if we use the same perturbation methodology.

4. *Page 14, line 5: "is similar" -> "behaves similarly"*

Agreed.

5. *Page 15, lines 8-9: "...correct the subsurface fields": Fig. 9 is in data space (surface), so one cannot see a subsurface effect from that figure.*

We replace the term 'subsurface' with 'water column above the pycnocline, which is actually meant there.

Authors' response to editor' comments

1. *P. 2, ll. 26-28. The depth of the transition between the two layers is not mentioned (see also Fig. 4 and associated comments).*

We add the depth of transition to the end of the sentence “The exchange of contrasting water masses forms a highly stratified water column structure throughout the system ...” as “with a pycnocline around 20 m. depth.”

2. *Fig. 7. What are exactly the increments shown there (I understand there is one set of increments for each element of the assimilation ensemble)?*

The ensemble mean increment and innovation of the ensemble are used throughout this study.

3. *Eq. (1). I understand σ_{2y} is the assumed variance of the observational error. But what exactly is σ_{2xb} ? (I presume it is defined from the prior ensemble. But how exactly?). Anyway, you must also respond to minor comment 18 of referee 1.*

σ_{xb}^2 is the variance of the prior ensemble estimate to the observation y (the forward operator ensemble). We reformulated the paragraph as requested by reviewer #1 in minor comment 17.

4. *The spatial area over which the diagnostics shown in Fig. 8 have been computed does not seem to be mentioned.*

We add to the caption the sentence “The statistics are computed in the location of the observations used in the corresponding assimilation cycle.”

5. *I presume most of the readers (but maybe not all ...) will easily locate the Bosphorus and the Dardanelles on the various maps. But it may not be the same as concerns the Gulf of Izmit (p. 16, l. 11).*

Agreed. We can name the region as it as 'the northeastern Marmara Sea' and will mark locations on one of the Figures.

References

Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly weather review*, 129(12):2884–2903.

Aydoğdu, A., Pinardi, N., Özsoy, E., Danabasoglu, G., Gürses, Ö., and Karspeck, A. (2018) Circulation of the Turkish Straits System between 2008–2013 under complete atmospheric forcings, *Ocean Sci. Discuss.*, <https://doi.org/10.5194/os-2018-7>, (submitted, in review).

Book, J. W., Jarosz, E., Chiggiato, J., and Beşiktepe, Ş. (2014). The oceanic response of the Turkish Straits System to an extreme drop in atmospheric pressure. *Journal of Geophysical Research: Oceans*, 119(6):3629–3644.

Grayek, S., Staneva, J., Schulz-Stellenfleth, J., Petersen, W., and Stanev, E. V. (2011). Use of FerryBox surface temperature and salinity measurements to improve model based state estimates for the German Bight. *Journal of Marine Systems*, 88(1):45–59.

Özsoy, E., Di Iorio, D., Gregg, M. C., and Backhaus, J. O. (2001). Mixing in the Bosphorus Strait and the Black Sea continental shelf: observations and a model of the dense water outflow. *Journal of Marine Systems*, 31(1):99–135.

OSSE for a sustainable marine observing network in the Marmara Sea

Ali Aydoğdu^{1,2,3}, Timothy J. Hoar⁴, Tomislava Vukicevic⁵, Jeffrey L. Anderson⁴, Nadia Pinardi⁶, Alicia Karspeck⁴, Jonathan Hendricks⁴, Nancy Collins⁴, Francesca Macchia⁷, and Emin Özsoy⁸

¹Science and Management of Climate Change, Ca' Foscari University of Venice, Italy

²Centro Euro-Mediterraneo sui Cambiamenti Climatici, Bologna, Italy

³Nansen Environmental and Remote Sensing Center, Bergen, Norway

⁴National Center for Atmospheric Research, Boulder, Colorado, USA

⁵Office of Water Prediction, National Weather Service NOAA, USA

⁶Department of Physics and Astronomy, University of Bologna, Italy

⁷Barcelona Supercomputing Center (BSC), Barcelona, Spain

⁸Eurasia Institute of Earth Sciences, Istanbul Technical University, Turkey

Correspondence: Ali Aydoğdu (ali.aydogdu@nersc.no)

Abstract. An observing system simulation experiment (OSSE) is presented in the Marmara Sea. A high resolution ocean circulation model (FESOM) and an ensemble data assimilation tool (DART) are coupled. The OSSE methodology is used to assess the possible impact of a ferrybox network in the eastern Marmara Sea. A reference experiment without assimilation is performed. Then, synthetic temperature and salinity observations are assimilated along the track of the ferries in the second experiment. The results suggest that a ferrybox network in the Marmara Sea may improve the forecasts significantly. The salinity and temperature errors get smaller in the upper layer of the water column. The impact of the assimilation is negligible in the lower layer due to the strong stratification. The circulation in the Marmara Sea, particularly the Bosphorus outflow, helps to propagate the error reduction towards the western basin where no assimilation is performed. Overall, the proposed ferrybox network can be a good start to design an optimal sustained marine observing network in the Marmara Sea for assimilation purposes.

1 Introduction

The Marmara Sea is one of the compartments of a water passage known as the Turkish Straits System (hereafter TSS). The TSS connects the Black Sea and the Mediterranean by two narrow straits, namely the Bosphorus and the Dardanelles, along with the Marmara Sea.

Salty and dense Mediterranean waters and brackish Black Sea waters form a two-layer exchange flow through the TSS. In combination with the complex topography, flow structures in different scales generate a challenging environment for oceanographic studies.

Until recently, ~~the need for high resolution in the straits made it infeasible to model~~ building a model solving for the hydrodynamics of the complete TSS ~~due to the computational cost~~ was not considered a feasible undertaking due to high computational costs of the required horizontal and vertical resolution enabling to represent the sharp stratification and the extremely complex topography of the straits and shelf regions, largely differing from those of the larger neighboring basins.

5 However, ~~integral~~ models of the whole system have emerged with increasing computational capacity in recent years (Sannino et al., 2017; Gürses et al., 2016; Stanev et al., 2017).

In this study, to the best of our knowledge, we present the first data assimilation experiments performed in the Marmara Sea. In the region, in-situ observations are generally scarce and collected by dedicated projects (Ünlüata et al., 1990; Özsoy et al., 2001; Beşiktepe et al., 1994; Tuğrul et al., 2002; Chiggiato et al., 2012) for a limited area and time. Moreover, the spatial
10 resolution and frequency of satellite observations are still not sufficient to monitor the system continuously. For these reasons, a sustainable marine monitoring network is required in the Marmara Sea. We propose a ferrybox network mounted on existing public transportation infrastructure.

We follow the ~~OSSE~~ Observing System Simulation Experiments (OSSE) methodology to achieve our goal. OSSE has been used widely in the atmospheric community for four decades ~~for design of~~ to design new observation tools, for error assessment
15 in large models and parameter estimation (Arnold Jr and Dey, 1986; Masutani et al., 2010). Atlas (1997) summarizes the criteria established by the atmospheric community to perform credible OSSE. Halliwell Jr et al. (2014, 2015) gave an example of an ocean OSSE in the Gulf of Mexico by following those criteria. Aydoğdu et al. (2016) studied a fishery observing system in the Adriatic Sea taking the same criteria into account.

This paper is organized as follows: In the next section, the main characteristics of the TSS relevant to this study are summarized. In section 3, the model and data assimilation scheme that we used are documented. Section 4 is devoted to the OSSE
20 design. The nature run and forward model are introduced. Ferrybox network design is also detailed and the methodology for impact assessment is outlined. The results are discussed in section 5. The summary and conclusions are presented in the last section.

2 Overview of the Turkish Straits System

25 The Marmara Sea with the Bosphorus and the Dardanelles Straits constitute the Turkish Straits System (TSS) which connects the Black Sea and the Mediterranean. The exchange of contrasting water masses forms a highly stratified water column structure throughout the system with a pycnocline around 20 m. depth. The brackish surface water of the Black Sea flows towards the Mediterranean in the upper layer and the salty and dense Mediterranean Sea water occupies the lower layer of the water column (Ünlüata et al., 1990). In addition to the strong stratification, the complex topography of the two narrow straits and an internal
30 sea with shallow shelves and deep depressions presents a unique and challenging environment for oceanographic studies.

The Bosphorus is an elongated narrow strait connecting the Black Sea and the Marmara Sea. The upper layer flow is dominated by the Black Sea water with salinity about 18 psu. The lower layer water originates from the Mediterranean and reaches the Black Sea with a salinity of about 37 psu. The interface depth between the two layers is mainly determined by the

salinity gradient. In summer, a three layer temperature structure appears. A warm upper layer due to the atmospheric heat flux overlays a cold intermediate layer that propagates from the Black Sea. Warmer Mediterranean water constitutes the bottom layer (Altıok et al., 2012). The constriction in the mid-section and the sill in the southern section of the strait apply a hydraulic control on the flow and produce a surface jet at the Marmara Sea exit (Sözer and Özsoy, 2017). The velocity of the Bosphorus jet can exceed 2 m/s (Jarosz et al., 2011). Therefore, the jet is one of the key factors influencing the dynamics of the Marmara Sea. It enhances the mixing by entraining lower layer water (Ünlüata et al., 1990), energizes the Marmara Sea circulation in the eastern basin (Beşiktepe et al., 1994) and produces mesoscale eddies due to the potential vorticity balance (Sannino et al., 2017).

The wind is another important factor influencing the dynamics of the Marmara Sea. Many strong cyclones pass over the Marmara Sea especially in winter (Book et al., 2014). Northeasterly winds dominate the atmospheric circulation throughout the year.

The third important dynamical constituent is the density-driven baroclinic flow in the lower layer (Hüsrevoğlu, 1998). The velocity of the flow in the lower layer can reach 1 m/s (Jarosz et al., 2012).

The high complexity of the system requires ~~integral modelling approaches to represent the links between its different compartments~~ models that simultaneously solve the whole system in its smallest resolved details and optimally representing multiple scales of interest.

3 Ensemble modelling and data assimilation in the TSS

The ocean model used in this study is the Finite Element Sea-ice Ocean Model (FESOM). FESOM is an unstructured mesh ocean model using finite element methods to solve the hydrostatic primitive equations with the Boussinesq approximation (Danilov et al., 2004; Wang et al., 2008). ~~It was developed by the Alfred Wegener Institute as the first global ocean model using an unstructured mesh.~~ Gürses et al. (2016) applied FESOM to the TSS and performed realistic simulations of the complete system using atmospheric forcing. The TSS model domain extends zonally from 22.5°E to 33°E and meridionally from 38.7°N to 43°N. The mesh resolution is as fine as 65 m in the Bosphorus and 150 m in the Dardanelles. In the Marmara Sea, the resolution is always finer than 1.6 km and is not coarser than 5 km in the Black Sea and the Aegean Sea. The water column is discretized by 110 vertical z-levels. Vertical resolution is 1 m in the first 50 m depth and increases to 65 m at the bottom boundary layer in the deepest part of the model domain. The mesh has 70240 nodes at the surface layer and more than 3 million nodes for the 3D state variables i.e. temperature, salinity, zonal and meridional velocity.

~~Aydoğdu et al. (2018)~~ Aydoğdu (2017) performed inter-annual simulations using a similar model setup. Two six years simulations using different surface salinity boundary conditions have been completed and evaluated. It is shown that the model is able to simulate main characteristics of the system qualitatively. It simulates the two layer structure of the TSS, successfully. Moreover, dynamical properties such as sea level difference between different compartments and dynamics related to high frequency atmospheric events are well-captured. However, the error growth in temperature and salinity throughout the simulation is also demonstrated which motivates the need for data assimilation in the TSS.

In this study FESOM has been coupled with an ensemble-based data assimilation framework, Data Assimilation Research Testbed (DART). DART is an open-source community facility developed at NCAR (Anderson et al., 2009) that provides data assimilation tools to work with either low-order or high-order models for different research activities (Raeder et al., 2012; Karspeck et al., 2013; Schwartz et al., 2015; Hoteit et al., 2013).

5 DART includes several different stochastic and deterministic ensemble Kalman filtering algorithms. In this study, we use the Ensemble Adjustment Kalman Filter (EAKF) as described in Anderson (2001). The EAKF is a deterministic ensemble Kalman filter, where the observations are not perturbed randomly before they are assimilated. One of the main advantages of the EAKF for our application is that it preserves the prior covariance ~~information~~ distribution during resampling which is not the case for the ensemble filters such as kernel or Gaussian resampling filters. The ensemble covariance is used to quantify the relation
10 between pairs of state variables or an observation and a state variable in the linear Gaussian context of the ensemble Kalman filter. In other words, it is the statistical representation of the prior constraints (dynamical balances) generated by the model. For a linear model, a sufficiently large ensemble would be able to represent all the balances. However, our model is nonlinear and cost forces us to use a small ensemble. This means that the ensemble sample covariances do not exactly represent the dynamical balances. The ensemble filter algorithm and inflation we use are designed to maintain the prior ensemble covariance as much as possible. Moreover, localization is effective in reducing forecast RMSE compared to observations given the errors in the sample covariances. However, it disrupts the prior sample covariance, so can result in reduced dynamical balance. In the TSS, where different dynamics compete and generate circulation structures at various scales, the prior ~~information~~ distribution is crucial for maintaining the dynamical balances among different scales which justifies the choice of the EAKF. In addition, ~~the analysis~~ covariances are updated in every assimilation cycle which is important to sustain the high frequency variability of
15 the system.

One of the main issues to take into consideration in ensemble data assimilation is the filter divergence. It can result from insufficient ensemble variance which leads to assigning more weight to the prior information which in turn may cause rejection of information from observations. As a result, the analysis departs from the observations (Anderson, 2001). There are techniques to prevent filter divergence such as inflating the covariances (Anderson and Anderson, 1999). In the standard DART,
25 inflation of the prior ensemble leads to multiplying the prior covariances by a constant slightly larger than 1. Instead of DART's multiplicative covariance inflation, we have applied random perturbations to the background vertical diffusivity for each ensemble member since one of the main sources of error in the model is the vertical mixing and the position of the interface between the upper and lower layers. The background vertical diffusivity ~~κ_{v0}~~ κ_{v0} which is implicitly represented in the κ_v of the tracer equations is randomly perturbed all over the domain every timestep by fitting a gaussian with a mean of 0 and a
30 standard deviation of 5% of the ~~κ_{v0}~~ κ_{v0} . We note that κ_{v0} is approximately two orders of magnitude smaller than the spatially and temporally varying κ_v .

Houtekamer and Mitchell (1998) shows that spurious correlations in the prior information can be avoided by using large ensembles. Since it is computationally expensive to use large ensembles in large geophysical models, they proposed a spatial localization technique to overcome spurious small correlations associated with remote observations. The impact of an obser-
35 vation on state variables is reduced as a function of distance from the observation with impact going to zero at a finite cutoff

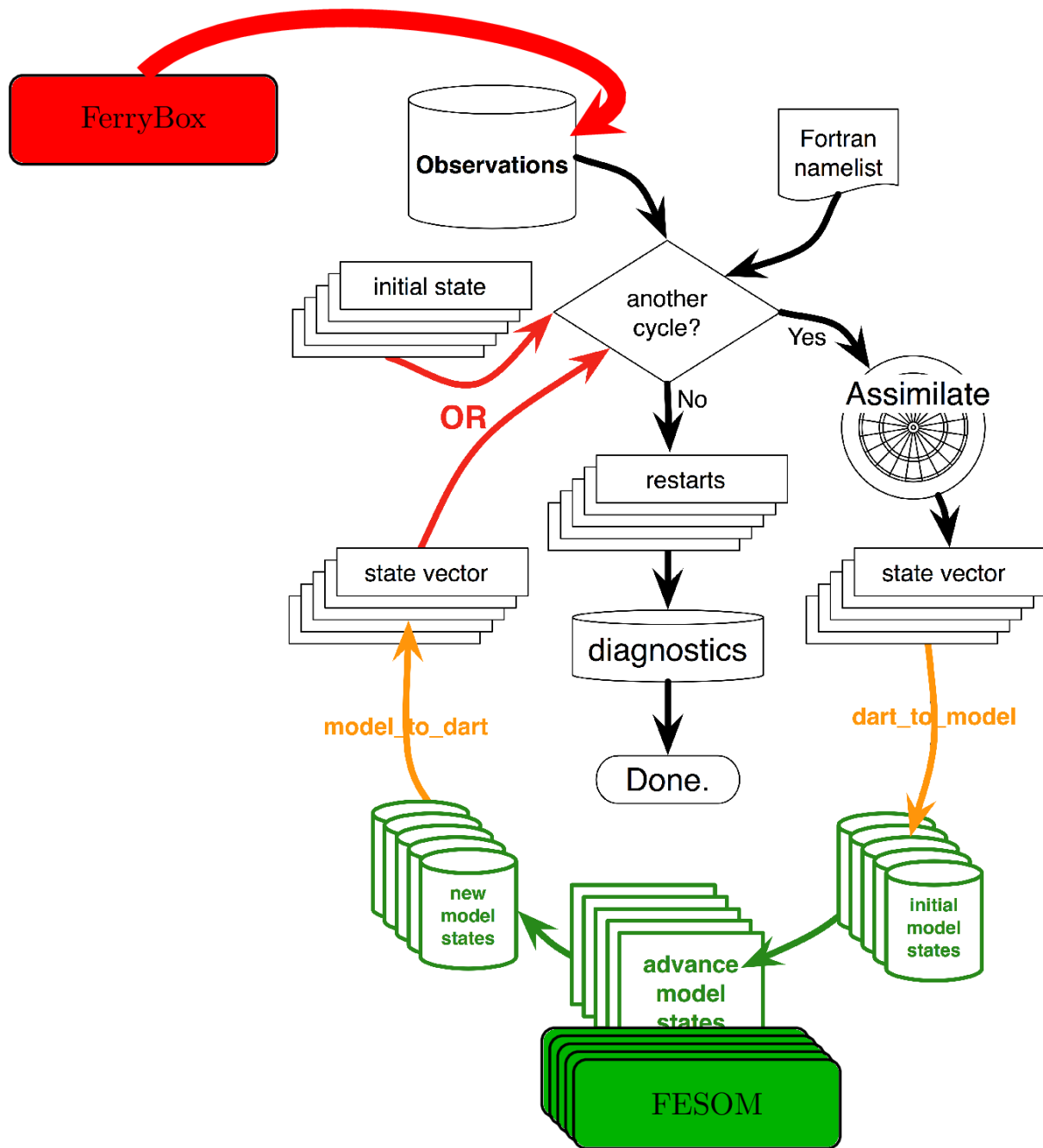


Figure 1. Flow of FESOM/DART system. Reproduced after Anderson et al. (2009) with modifications for the TSS application

value. In DART, the cutoff value can be set as a constant value in radians. The localization function is the fifth order piecewise continuous, compactly supported as presented in Gaspari and Cohn (1999). The horizontal cutoff radius, which is the half-

width of the Gaspari-Cohn kernel, is used to deal with the high spatial variability of the water masses along the TSS. Moreover, the water column in the TSS is strongly stratified in the vertical. Therefore, vertical localization is also applied to update of the state in the upper layer of the water column with less impact in the lower layer. As a result, the localization is an ellipsoidal volume centered on the observation being assimilated with a horizontal radius of two cutoff value and vertical radius of two cutoff value normalized by a factor. This approach is chosen not to face with problems, such as the breaking of the stratification during the integration.

For interfacing FESOM and DART, a forward operator that maps model state vector into observation space, characterized by a type of physical quantity, geo-referenced location and time was defined. For the ferrybox type of observations we used a simple, but theoretically justifiable nearest-neighbor mapping. Specifically, for each discrete observation value the forward operator first finds the closest horizontal grid point in the state vector and then the closest vertical location.

The details of the FESOM/DART coupling are shown schematically in Fig. 1. The interface to DART requires two model specific routines that convert the model restart files to DART state vectors and vice versa. For the model to DART step the state vector is provided from each ensemble member. The updated state vectors are then converted to the model restart files as an initial ensemble for the next assimilation cycle within which FESOM is applied to integrate the model state forward to the next analysis time. This process is repeated until the experiment finishes.

4 Design of the OSSE

The OSSE methodology used is shown schematically in Fig. 2. A proposed ferrybox network was assumed to be in the eastern Marmara Sea using the existing transportation infrastructure. Therefore, we were able to simulate the observations from the ferrybox network by tracking the ferrylines in the Marmara Sea. The impact assessment of this network was performed using the FESOM/DART ensemble data assimilation system.

4.1 The nature run and forward model

The OSSE methodology requires a representation of a reference true atmosphere or ocean which is called the nature run (NR). The NR is used for generating the synthetic observations to assimilate and for assessing the quality of the assimilation. For a fully objective OSSE the NR should be created using a simulation from a different model than the forward model (FM) used in the assimilation in an attempt to simulate the inevitable model errors.

In the current study the fraternal twin method was used, where the NR and FM are based on the two different configurations of the same model. The difference between the NR and the FM configurations is the surface salinity boundary condition, where the former employed the relaxation boundary condition for the sea surface salinity whereas the latter uses the mixed salinity boundary condition (Huang, 1993) (Huang, 1993, see also appendix A) .

This approach was chosen because FESOM in the Marmara Sea is sensitive to equally plausible salinity boundary conditions. This was demonstrated through long-term simulations using both configurations, which were validated with the actual in-situ observations (Aydoğdu et al., 2018) (Aydoğdu, 2017) . The simulations were performed for period 1 January 2008 to 31

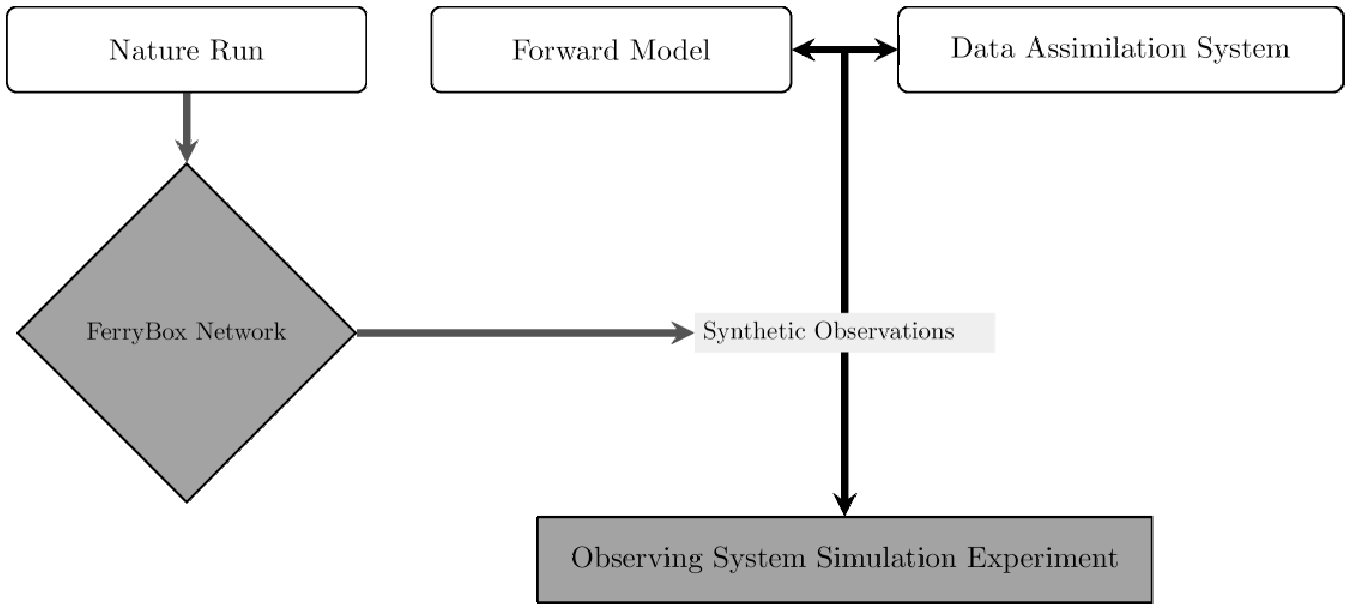


Figure 2. OSSE methodology applied in the Turkish Straits System. The forward model (FM) is a different configuration of the model setup used for the Nature Run (NR) as detailed in the text.

December 2013. It is demonstrated that they deviate from each other but both are still realistic during the OSSE period which is 1-8 January 2009. The output of the NR was saved hourly during the OSSE period to generate synthetic observations.

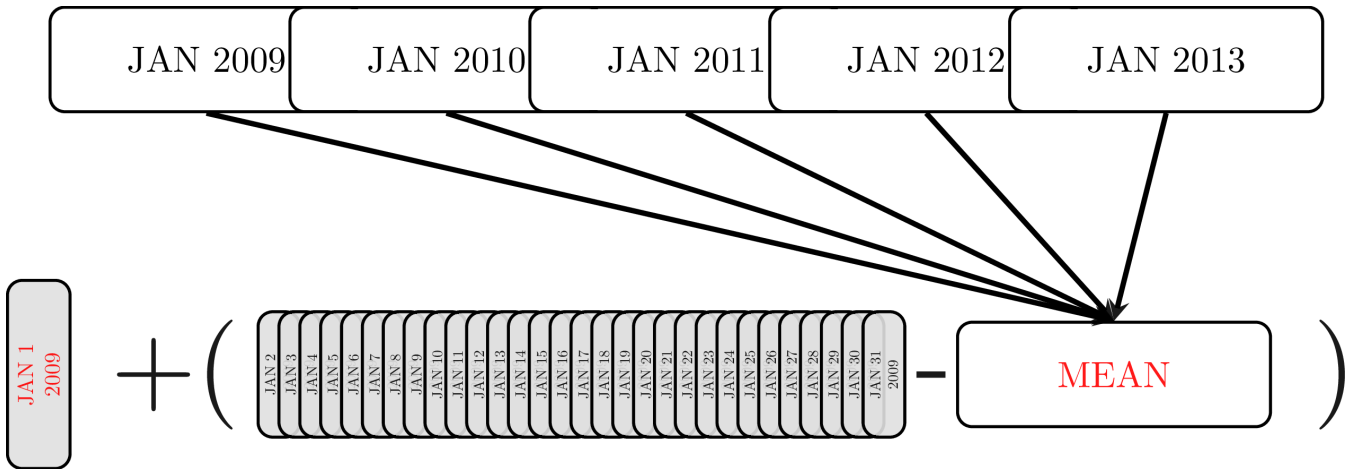


Figure 3. Schematic representation of the methodology used to generate the initial ensemble. MEAN is the average of January over five years between 2009-2013

The ensemble consists of 30 members. An initial ensemble is produced for 1 January 2009 as schematized in Fig. 3. The five-year means of the temperature and salinity for January are computed from the inter-annual-multi-year simulation. Then, the deviations of each day between 2-31 January 2009 from the inter-annual mean are calculated. Finally, these deviations are added to the temperature and salinity fields of 1 January 2009 to provide an initial perturbation for each ensemble member.

5 Fig. 4 shows the variance of the initial ensemble for salinity and temperature for the depths 5 m and 20 m. The horizontal distributions of the variance are similar for both variables. The variance is larger in the exits of the straits. In particular, the spread at the Bosphorus exit for the upper layer down to 15 m is larger due to the variability in the outflow. Correspondingly, the Dardanelles inflow to the Marmara Sea increases the variance in the lower layer. Such a distribution of initial variance in the Marmara Sea is appropriate to initialize the experiments since the assimilation of synthetic observations is performed in

10 the impact area of the Bosphorus outflow where the ensemble spread can be diminished by the assimilation.

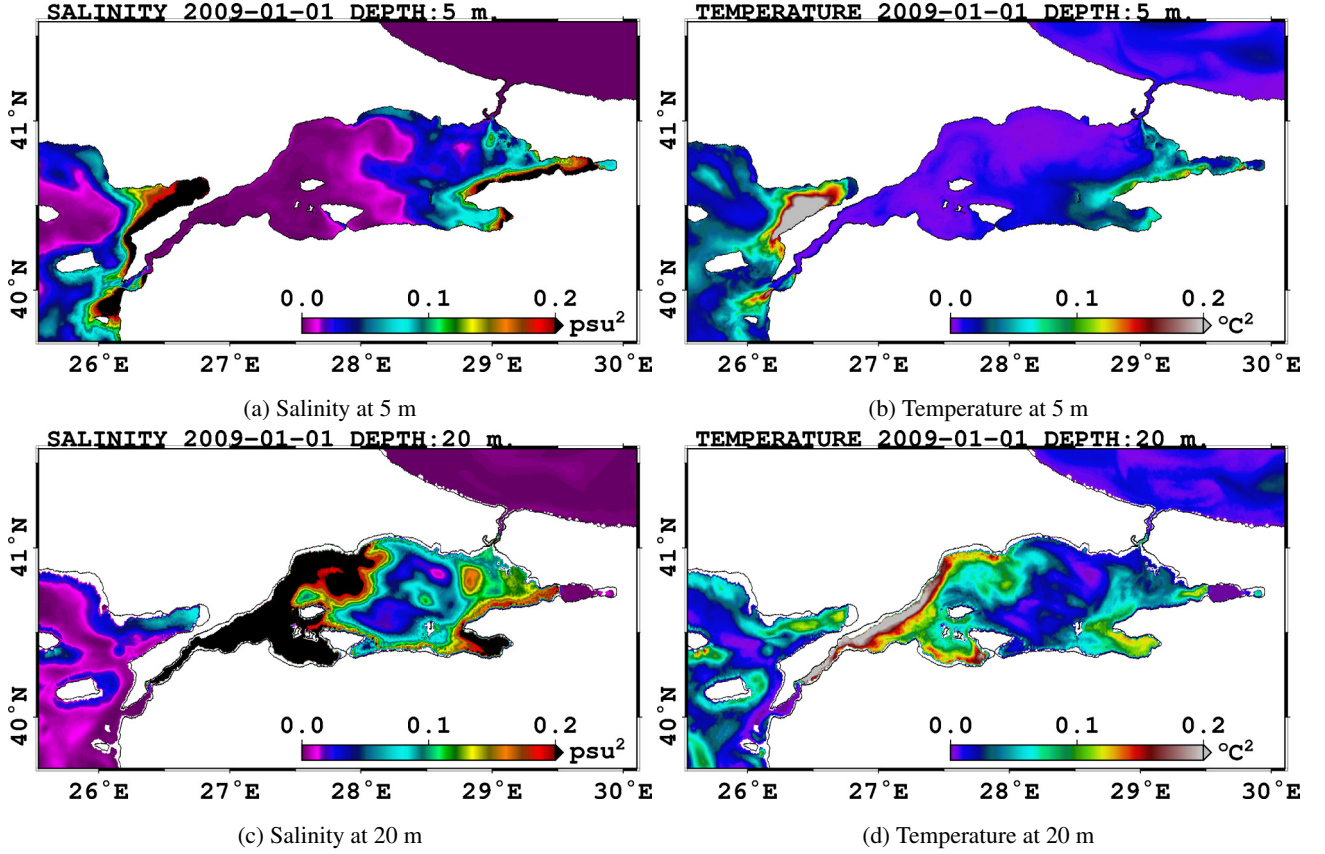


Figure 4. Salinity (left) and temperature (right) variance of the initial ensemble at 5 m (top) and 20 m (bottom) depth on 1 January 2009 at 00:00 UTC.

4.2 The eastern Marmara Sea ferrybox network design

The Marmara region has the highest urban population in Turkey. It includes the metropolitan city of Istanbul which is divided into two by the Bosphorus and surrounded by the Black Sea in the north and the Marmara Sea in the south. As a consequence, a network of ferries is an essential means of transportation. The main hub for the ferries is Istanbul which is connected to the other cities around the Marmara Sea by several ferries (Fig. 5a).

In the Marmara Sea, it is not easy to deploy ~~instruments such as argo or glider~~, for instance, argo floats or gliders close to the surface since there is heavy ship traffic. However, the infrastructure for a ferrybox network is already available. The ferrylines in the Marmara Sea cover most of the eastern basin including the Marmara Sea exit of the Bosphorus. Using the ferry network as a monitoring system would be an efficient way to build a sustainable ocean observing network in the Marmara Sea.

Some existing state-of-the-art ferrybox networks are discussed by Petersen (2014). Usage has increased since the European network for ferrybox measurements project¹, especially in the northern European seas. In the Mediterranean, there is a ferrybox system between Piraeus and Heraklion operated by HCMR (Korres et al., 2014). A ferrybox mainly includes temperature, salinity, turbidity and chlorophyll-a fluorescence sensors, and a GPS receiver for measuring position. Oxygen, pH, pCO₂ or algal groups as well as air pressure, air temperature and wind sensors can also be installed (Petersen, 2014). The ferrybox observations can be used for analyzing the state of the ocean (Seppälä et al., 2007), comparison with other instruments (Sørensen et al., 2007) and can also be assimilated to improve the state of the ocean models (Grayek et al., 2011).

The sampling rate of the data can vary between systems. The data used in Grayek et al. (2011) is sampled at 10 s intervals. For our OSSE, we use a sampling rate of 1 min. following Korres et al. (2014). The synthetic observations are obtained from hourly NR outputs at varying spatial ~~location~~ locations so as to remain within one minute time intervals along the track of the ferries. The depth of the sampling is set to 5 m, an average depth given by Grayek et al. (2011). A random error, sampled by a gaussian around zero mean and standard deviation of 0.1°C for temperature and 0.04 psu for salinity, is added to each synthetic observation following Aydoğdu et al. (2016) in order to simulate realistic measurements taking the instrumental error into account. These synthetic observations are the same for each ensemble member i.e. the kalman filtering is not stochastic in which observations are perturbed. On the other hand, observational error matrices in the Kalman gain are chosen to be uncorrelated and constant diagonal with 0.5°C and 0.25 psu for temperature and salinity, respectively, as proposed by Grayek et al. (2011) considering other sources of error such as representativeness.

Three different ferry routes are chosen from the map in Fig. 5a, and their tracks are approximated by observing from the real-time Marine Traffic² application (Fig. 5b). The longest duration for a cruise in the eastern Marmara Sea is about 3.5 hours between Ambarlı-Topçular (Table 1). Another transect used here is YeniKapi-Yalova which takes about 75 min. and has cruises every two hours from each port. This route directly crosses the Bosphorus outflow and has the highest number of cruises a day. Therefore, we include six ferries in various periods of the day for YeniKapi-Yalova transect. The last transect chosen is Yenikapi-Bandırma crossing the Marmara Sea from north to south. The navigation on this line takes about 2.5 hours.

¹<http://www.ferrybox.org>

²<https://www.marinetraffic.com>

Route	Location	Distance	Speed	Duration	Departure Time
Yenikapi	28.956E-41.002N				09:45,15:45,21:45
		46.2km (28.7mi.)	23 kn	75 min.	
Yalova	29.274E-40.661N				07:15,13:45,19:45
Yenikapi	28.956E-41.002N				07:30
		110.0km (68.5mi.)	27 kn	150 min.	
Bandirma	27.967E-40.354N				18:30
Ambarli	28.676E-40.966N				20:00
		70.0km (43.2mi.)	12 kn	210 min.	
Topcular	29.434E-40.690N				16:00

Table 1. The unidirectional ferry tracks used in this study. The locations, distance between the ports, speed of the ferries, the duration of the cruise and time of departure from each port are listed.

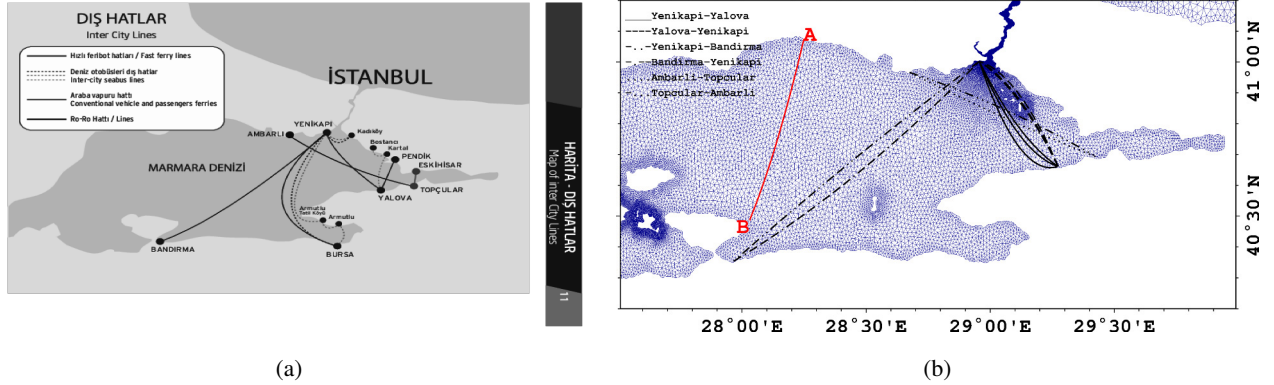


Figure 5. a) The routes of the intercity ferry lines from Istanbul and to Istanbul suggested by the operating company IDO (<http://www.ido.com.tr>). b) Approximate unidirectional ferry tracks. The legend shows the direction of the ferries. The section A-B is used only for impact assessment against the NR. Triangular mesh of the model is underlaid.

This transect is the only one that has direct impact in the southern basin. The resulting synthetic temperature and salinity observations sampled from the NR for the first day are shown in Fig. 6.

4.3 Experiments

Two experiments are performed as an initial evaluation for the data assimilation studies in the Marmara Sea (Table 2). The first experiment FB001 is a reference experiment without assimilation. It is used to evaluate the errors when the synthetic observa-

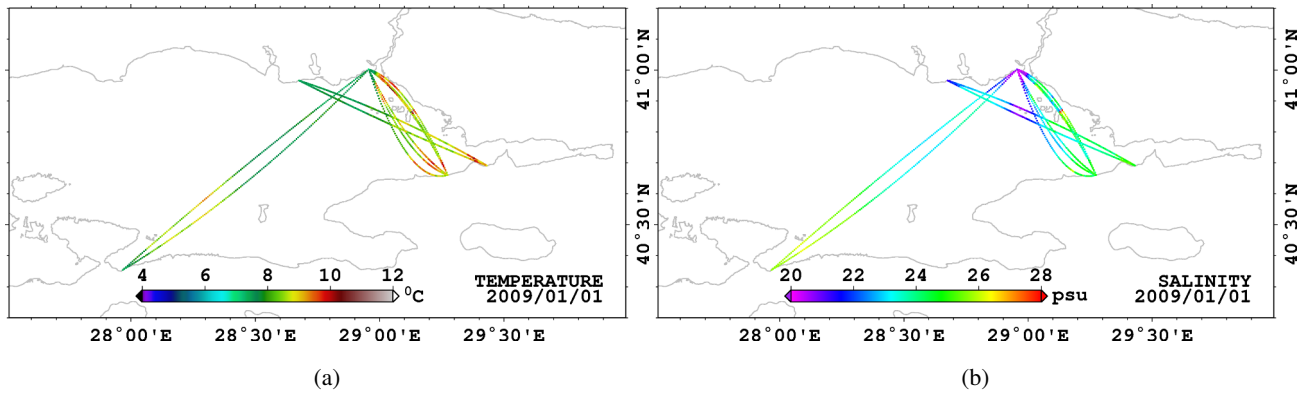


Figure 6. Synthetic a) temperature and b) salinity observations on 1 January 2009 sampled from the NR.

tions are not assimilated. In the second experiment, FB002, all the synthetic observations are assimilated in the corresponding assimilation window. A six hours width is chosen for each assimilation window, since the area is under the influence of the Bosphorus jet which may develop high frequency variability in the water mass structure and circulation at the upper layer, especially during severe storms and atmospheric cyclone passages.

	Start Date	End Date	Assimilation	A.Cycle	Evaluation
FB001	01-JAN-2009 00:00	8-JAN-2009 00:00	NONE	N/A	YES
FB002	01-JAN-2009 00:00	8-JAN-2009 00:00	ALL	6 hr	YES

Table 2. Summary of the OSSEs. Start and end date of both experiments are the same. There is no assimilation in FB001 whereas all the data are assimilated in FB002. Assimilation cycle is 6 hr for FB002.

- 5 The horizontal cutoff radius is set to 6.36 km and the vertical cutoff radius is considered as 15 m centered on the observation location. The mean temperature and salinity increments after the first and third assimilation cycle are shown in Fig. 7a and Fig. 7b, respectively. As can be seen from the temperature corrections after the first cycle, the whole track of the Yenikapi-Bandırma route is not assimilated at once since the southern section of the data is not in the current assimilation window. Moreover, the updates on the salinity fields are smaller around 20 m (Fig. 7b), compared to the upper layers (not shown) closer
- 10 to the observation locations.

4.4 Methodology for impact assessment

- The DART offers tools to have the control on the data without any difficulty. ~~It allows to decide to use a set of data for assimilation or evaluation. It also checks if any operator fails during the mapping of the state vector to the observations~~ Each type of observation can either be assimilated or withheld to evaluate the resulting analyses. DART also provides a rudimentary
- 15 quality control capability that can reject observations that are too different from the ensemble mean prior estimate. Synthetic

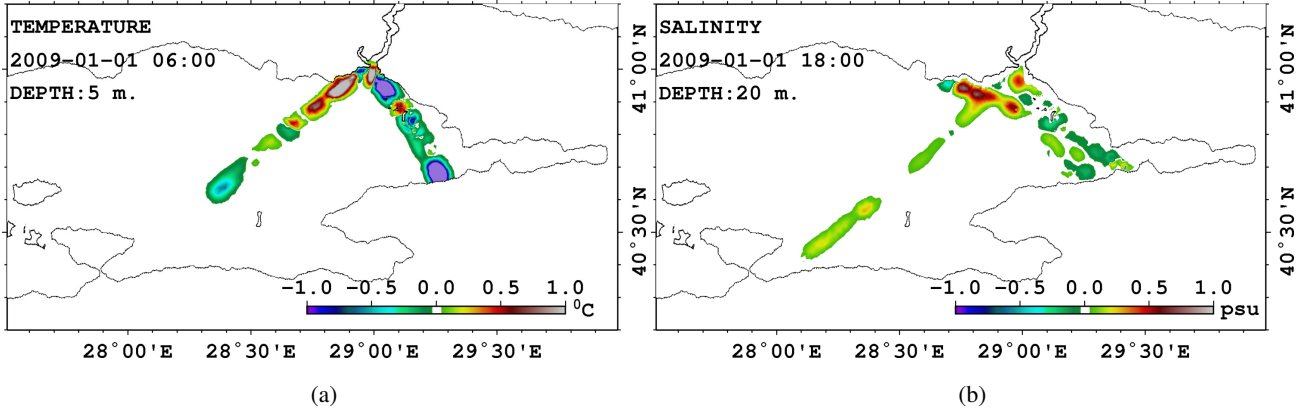


Figure 7. [Increments](#)-[Mean increments](#) after assimilation for a) temperature at 01-01-2009 06:00 at 5 m depth and b) salinity at 01-01-2009 18:00 at 20 m depth.

observations are not used in the assimilation (rejected) if $x_b - y > TE(x_b - y)$. Here, y is the observation and x_b is the corresponding prior mean of the ensemble of forward operators. The expected value, E , is computed as:

$$E(x_b - y) = \sqrt{\sigma_{x_b}^2 + \sigma_y^2} \quad (1)$$

where σ stands for the standart deviation. T is chosen as 3 for both temperature and salinity in these experiments [considering the real model errors reported by Aydoğdu \(2017\)](#).

The first diagnostic we use to assess the impact of the observations is the RMS of innovations which are the root mean square of the difference between the prior and observation. We also use horizontal maps of the innovations to determine the spatial distribution of error.

The OSSE methodology allows various ways to evaluate the analysis since the NR is assumed to be the true state of the system. We exploit this assumption to compare the experiments with the NR to understand the impact of assimilation better.

One diagnostic we use in this sense is to compare the NR and ensemble mean in the vertical. Although we don't assimilate any data below 5 m and we limit the radius for vertical updates it is important to assess the vertical distribution of the errors given the strongly stratified water column in the Marmara Sea. For this purpose, we compute the difference between prior ensemble mean and NR along the transect A-B (see Fig. 5) in the first 50 m depth.

Second diagnostic is the RMS of difference between the prior ensemble mean and the NR computed in the first 10 m of the water column for the whole basin. The propagation of the error reduction can be traced from the spatial maps of this diagnostic.

5 Results

Figure 8 shows the time evolution of [prior](#) RMS of innovations, ensemble spread and total spread for temperature and salinity. The RMS of salinity innovations continuously grows in FB001. It fluctuates around 2 psu and reaches to 2.4 psu at the end of

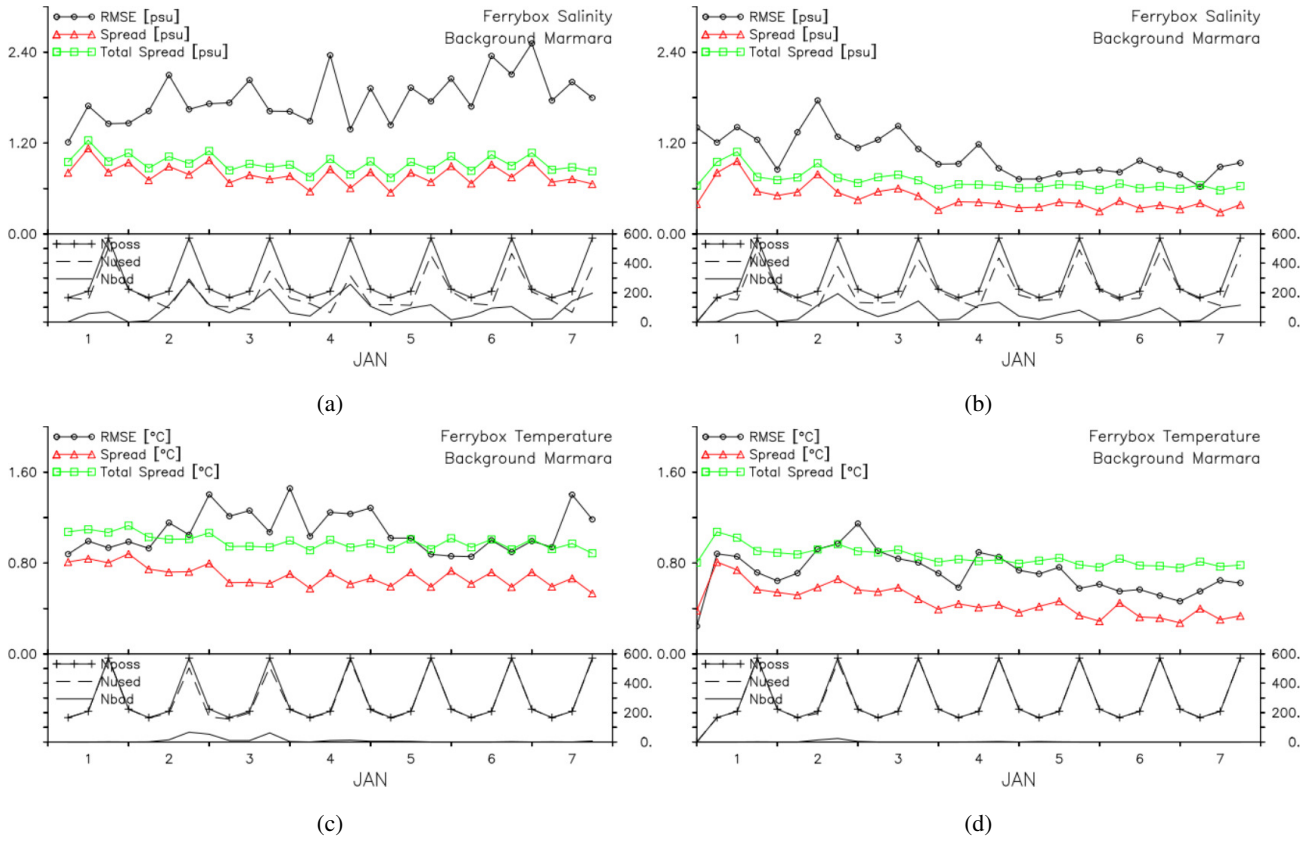


Figure 8. Timeseries of prior RMS of innovations, spread and total spread of salinity (top) and temperature (bottom) for FB001 (left) and FB002 (right). The spread is statistics are computed in the variance, therefore it location of the observations used in the corresponding assimilation cycle. The spread is the square root of the unit of the corresponding state variable variance. Total spread is the square root of the observational error added to the spread variance. Y-axis shows the range of each statistics indicated in the legend. Bottom panel of each figure shows the number of available observations (Nposs) in each assimilation cycle. For the assimilation experiment, FB002, Nused and Nout show the number of assimilated observations and outliers, respectively. For the experiment without assimilation, FB001, they are the number of observations which would be assimilated or rejected, respectively, in that specific assimilation cycle if assimilation was performed.

sixth day. In FB002, assimilation of the observations decrease the RMS of innovations, significantly. The RMS of innovations is generally below 1.2 psu. Although there is an increase of error in the first two days a gradual reduction takes place in the following days. The RMS of temperature innovations is similar behaves similarly to that of salinity after the second day. The error grows in FB001 even though the trend is not as obvious as in salinity errors of the same experiment. The analysis is improved at the end of experiment FB002 compared to FB001. The range of errors are comparable to the real model errors with respect to in-situ observations as demonstrated by Aydoğdu (2017). Another important result is that the ensemble still has spread comparable to the RMS errors at the end of the experiments (larger in FB001). In other words, the ensemble didn't

collapse after a week of assimilation. We recall that the ensemble spread is maintained by perturbing the background vertical diffusivity. The method seems promising at least for a week long period.

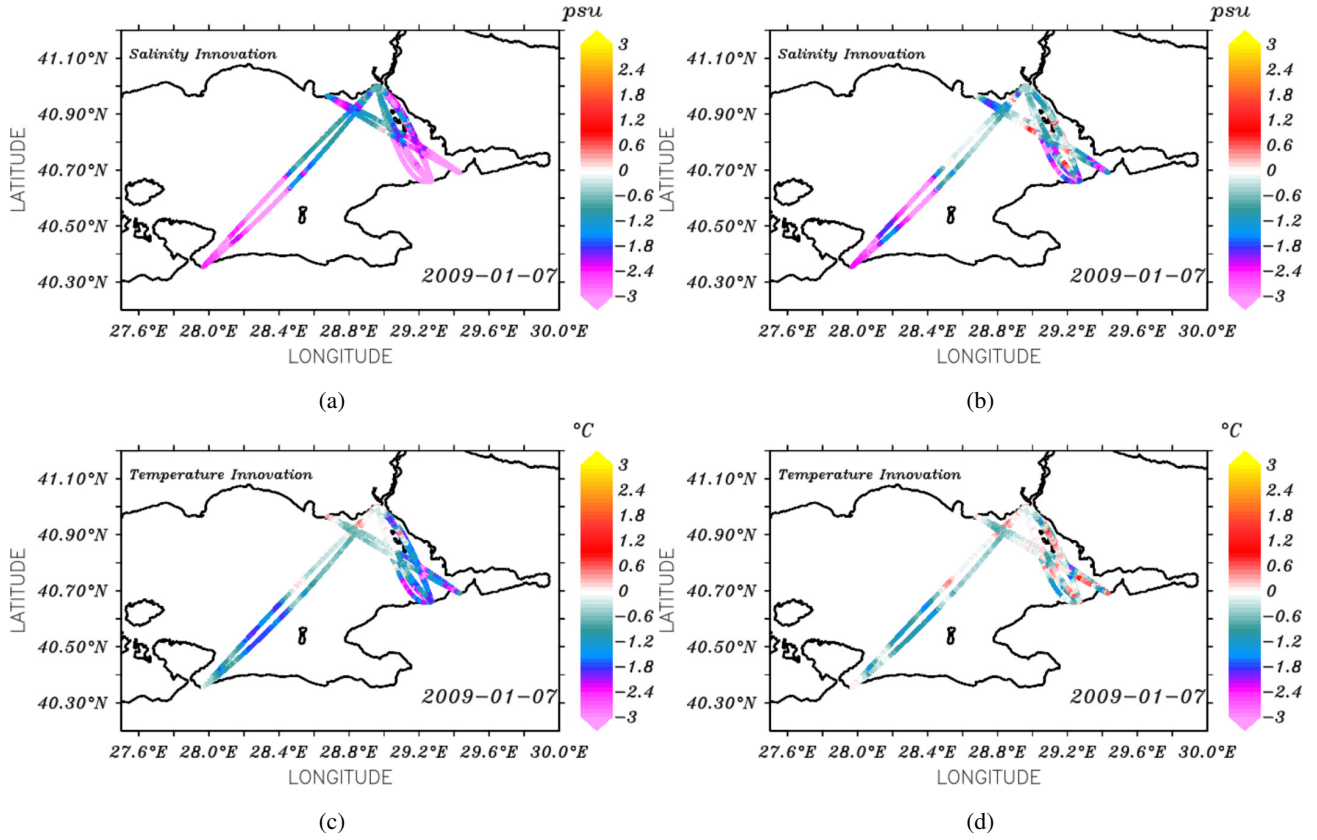


Figure 9. Horizontal distribution of salinity (top) and temperature (bottom) innovations along the ferry tracks in 7 January 2009 for FB001 (left) and FB002 (right). Observations with a innovation out of the range ± 3 are considered as outliers.

As discussed in section 4.4, the data are subjected to a quality control before assimilation. The bottom panels of Fig. 8 shows the number of available observations (N_{poss}), number of used observations (N_{used}) and number of outlier observations (N_{out}). The decrease in the number of outlier observations in FB002 points out an improvement also in the regions in which the innovations are larger as can be deduced from Fig. 9.

In the last day of experiments (Fig. 9), the salinity innovations are better almost everywhere in FB002 compared to FB001. There is a significant reduction in errors in the northern basin. Assimilation decreases the number of outlier salinity observations especially on the route between Yenikapi and Yalova. Moreover, innovations are also improved to a lesser extent in the southern basin where fewer observations are assimilated. The number of outlier temperature observations is very small in both experiments. Overall, the assimilation of temperature and salinity observations in the selected transects notably helps to correct the subsurface fields water column above the pycnocline.

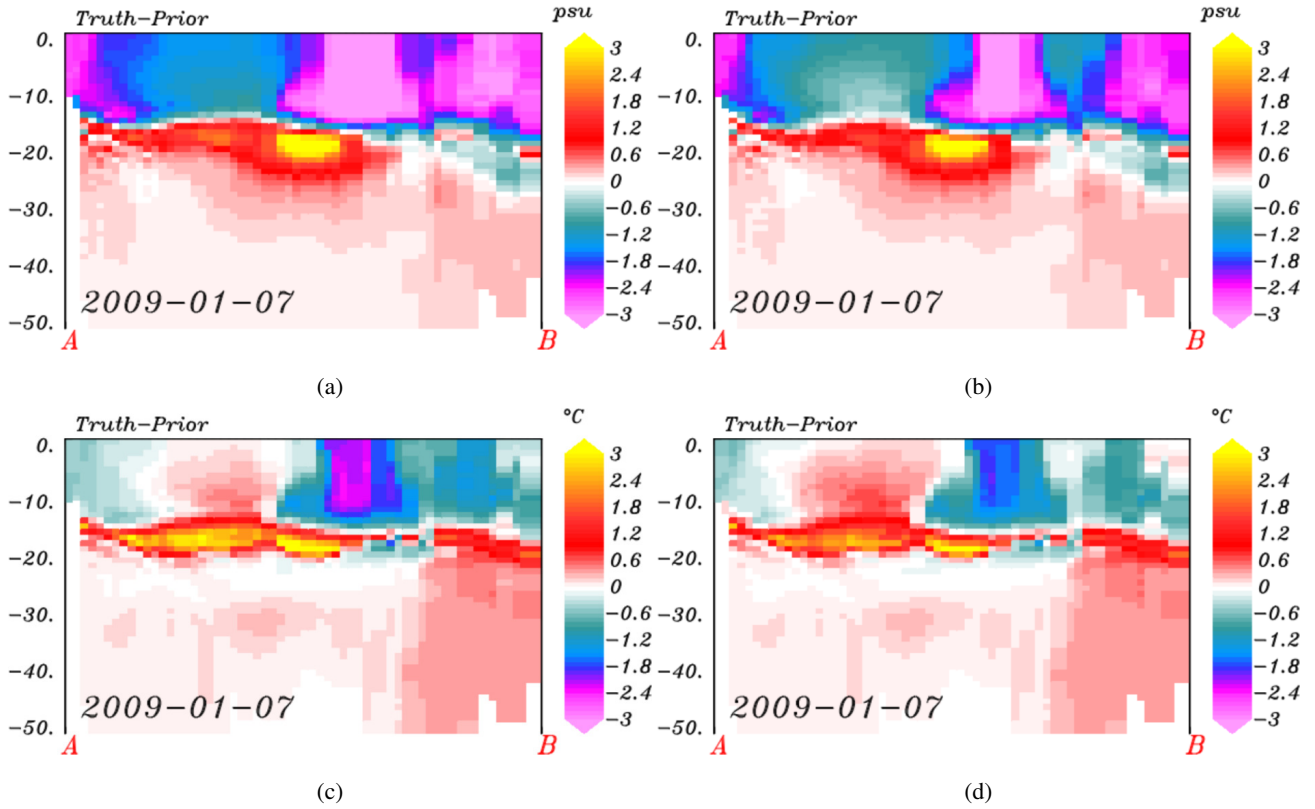


Figure 10. Vertical distribution of salinity (top) and temperature (bottom) difference between the NR and the prior ensemble mean along the cross-section A-B (see Fig. 5) in 7 January 2009 for FB001 (left) and FB002 (right).

The improvement of the analysis is also noticed in remote areas such as the A-B transect in the central basin (see Fig. 5). Figure 10 shows the difference between the truth and prior state down to 50 m depth along the A-B transect after the last assimilation cycle. Comparison of salinity differences in FB001 and FB002 reveals the improvement in the northern section down to 15 m depth. The southern part away from the coast also gets better. The middle of the transect still has large discrepancies in both experiments at the end of seven days. The correction in the remote area suggests a mechanism related to the outflow of the Bosphorus and the surface circulation of the Marmara Sea. The water masses which are corrected by assimilation in the eastern basin are pushed towards the west and reduce the error.

Figure 11 depicts the RMS of the difference between NR and prior [ensemble mean](#) salinity at the first 10 m depth. It clearly supports the mechanism suggested above. The distribution of RMS of differences in the non-assimilation case shows that the Bosphorus outflow also has some capability to reduce the difference since the Black Sea water masses govern the upper layer of the Marmara Sea (Fig. 11a and 11c). These two snapshots from the fifth and the last day of the FB001 show the error reduction in some regions due to the dynamics. Therefore, the conclusion that the improvement is due to the assimilation is not straightforward. However, comparison of the FB001 and FB002 reveals the role of the assimilation of ferry tracks on the

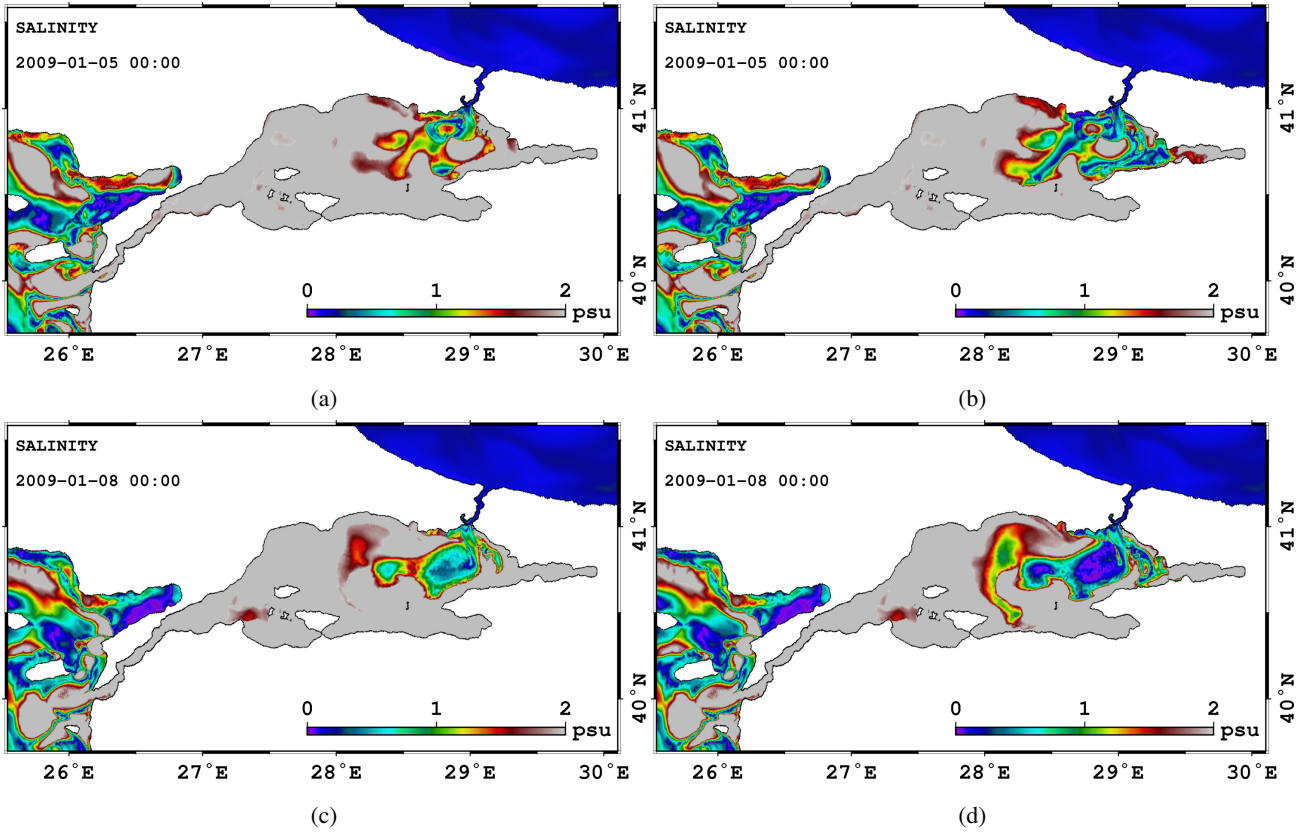


Figure 11. RMS of the difference between NR and prior [ensemble mean](#) salinity at the first 10 m. Comparison of FB001 (left) and FB002 (right) are shown for 5 January 2009 (top) and 8 January 2009 (bottom). RMS of difference is higher than 2 psu in the gray areas.

error reduction, clearly. In the fifth day, lower RMS of differences in FB002 (Fig. 11b) extend towards the [Gulf of Izmit in the east-eastern-most part of the basin](#) which is absent in FB001 (Fig. 11a). The westwards propagation is more pronounced in the last day in FB002 (Fig. 11d) compared to FB001 (Fig. 11c). The central basin has lower errors almost everywhere in the north-south orientation including the A-B transect shown before.

- 5 [Finally, temperature](#) [Temperature](#) fields are already closer to the [truth-NR](#) in the fifth day even in the western basin (Fig.12). After seven days of assimilation, the temperature error in the Bosphorus plume is significantly reduced. The southern and central basin has improvements as much as 0.5°C locally.

- 10 [Finally, Fig.13 compares FB001 and FB002 in terms of salinity fields overlaid with corresponding circulation at 5 m. depth for the exemplary case on 7 January 2009 at 00:00. The salinity differs significantly between the two experiments especially along the southeastern coast, while there is very little change in both qualitative and quantitative terms in the horizontal circulation, namely the current speed and direction, in the affected region that can be attributed to assimilation. The effect of assimilation is more pronounced in terms of the property fields, which alternatively indicates changes in stratification and](#)

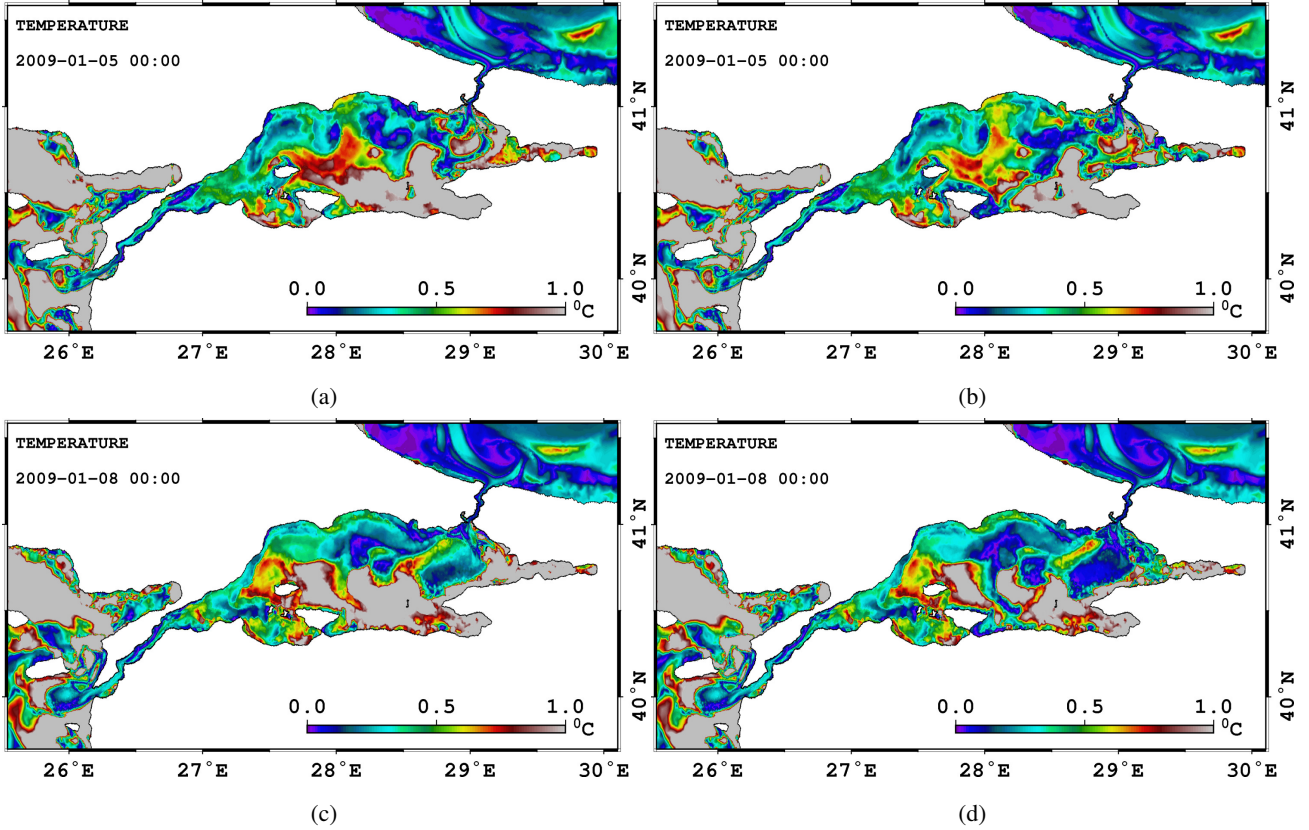


Figure 12. RMS of the difference between NR and prior temperature at the first 10 m. Comparison of FB001 (left) and FB002 (right) are shown for 5 January 2009 (top) and 8 January 2009 (bottom). RMS of difference is higher than 1°C in the gray areas.

vertical mixing along the southern coast. This can be seen also in the comparison of the experiments (Fig.13) with the nature run (Fig.A1b) which shows that salinity in the FB002 is closer to the nature run than FB001 where the synthetic observations are assimilated. However, circulation is always more intense in the nature run compared to both experiments but with similar circulation pattern.

5 6 Summary and Discussion

We have described data assimilation experiments performed in the Marmara Sea. The main characteristics of the TSS have been summarized. For the study, a general ocean circulation model, FESOM and an ensemble data assimilation framework, DART have been coupled. The implementation of the data assimilation scheme has been reported.

The TSS is an important water passage for the oceanography of the neighboring Black and Aegean Seas. It also has important
10 impacts on their ecosystem by maintaining the exchange of water masses and nutrients. The high population in the cities

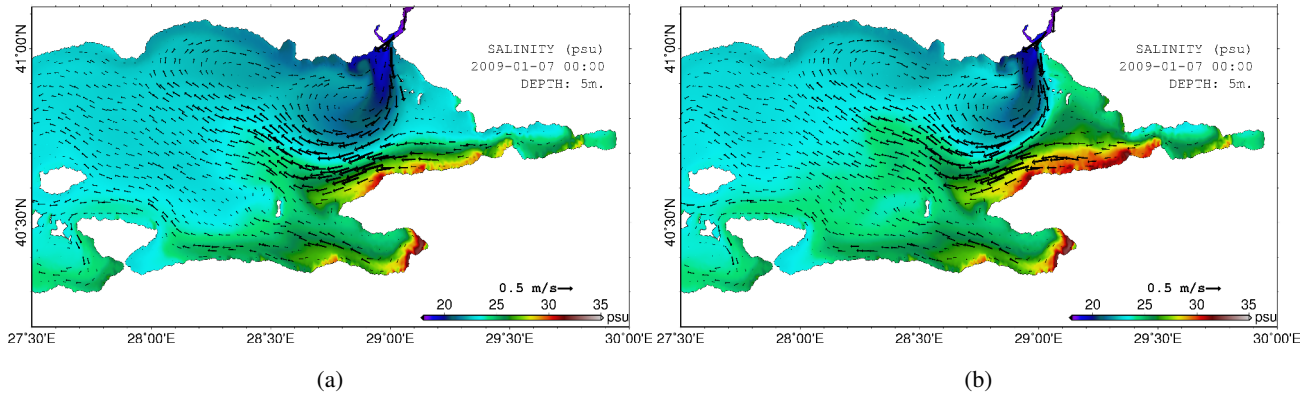


Figure 13. Salinity overlaid with current fields at 5 m. depth, obtained from experiments a) FB001 and b) FB002 on 7 January 2009 at 00:00.

surrounding the TSS and intense marine traffic through the passages add social and economic reasons to monitor the TSS in a sustainable way.

Real observations in the TSS have been obtained by dedicated projects for short time periods or have limited spatial coverage. Moreover, the satellite measurements are still low-resolution for monitoring and assimilation purposes. In this study, we proposed a sustainable marine monitoring network using the ferrylines in the eastern Marmara Sea. We think that equipping the ferries which operate daily everyday from Istanbul to various cities around the Marmara Sea with temperature and salinity sensors can provide immense amounts of data both in time and space. Given this motivation, we tested a ferrybox network including some of the ferry transects in the basin.

The OSSE methodology has been used to assess the impact of ferrybox measurements. We tried to satisfy the main criteria determined by approximately forty years experience of the atmosphere and ocean communities. However, it was still not possible to perform an OSE using real observations to compare with OSSE due to the lack of data during the experiment period.

The results of the two experiments presented here are promising. We showed that the assimilation of the salinity and temperature observations significantly improve the analysis in the Marmara Sea. The Bosphorus jet has an important role in the propagation of the error reduction towards the western basin where no data is assimilated. Moreover, the Marmara Sea circulation helps to improve the southern basin even for short timescales. The lower layer doesn't show any response to assimilation since a vertical localization around the observations around 5 m is applied to keep the impact in the upper layer as much as possible. Moreover, the stratification between the upper and lower layers is too strong so that it prevents the interaction between the two layers. Assimilation of temperature and salinity observations doesn't alter the circulation in the Marmara Sea significantly within the experiment period.

In conclusion, the results encourage further data assimilation studies in the Marmara Sea. The investigations can be extended to different observing systems, different areas of the sea or different dynamical focuses. Moreover, we believe the unique dynamics of the system demonstrated its ability to be a good natural laboratory for future data assimilation studies.

Appendix A

A1

The OSSE is performed in a fraternal twin setup as discussed in section 4.1. The nature run (NR) and forward model (FM) are integrated using similar model configurations, but different surface salinity boundary conditions. The NR and FM use the boundary condition A1a and A1b, respectively.

$$K_v \partial_z S \Big|_{z=\eta} = \gamma(S^* - S_0) - S_{corr} \quad (\text{A1a})$$

$$K_v \partial_z S \Big|_{z=\eta} = S_0(E - P - R) + \gamma(S^* - S_0) - S_{corr}^* \quad (\text{A1b})$$

In the boundary conditions (A1a) and (A1b), S_0 and S^* are the surface salinity and the reference salinity, respectively, γ is the relaxation coefficient. S_{corr} and S_{corr}^* are correction terms applied to the salinity corresponding to boundary conditions (A1a) and (A1b), respectively. The correction term is required for mass conservation in a closed lateral boundary model and discussed extensively in Aydoğdu (2017).

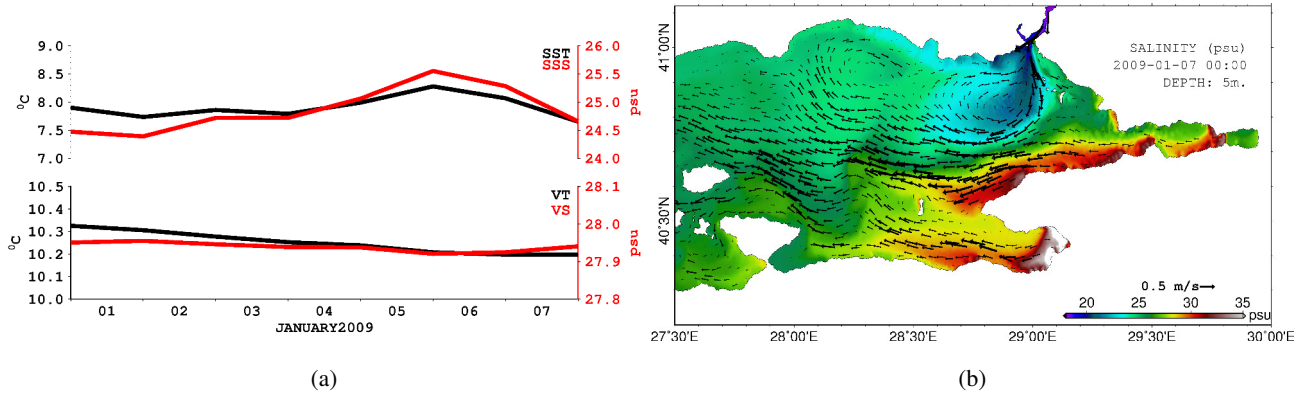


Figure A1. a) Daily timeseries of sea surface temperature (SST), sea surface salinity (SSS), volume temperature (VT) and volume salinity (VS) obtained from the nature run throughout the experiment period between 1-7 January 2009 b) Salinity overlaid with current fields at 5 m. depth, obtained from nature run on 7 January 2009 at 00:00

In Fig.A1a, the daily timeseries of the temperature and salinity are presented. The surface mean temperature and salinity show slight variability between 1-7 January 2009. The salinity field overlaid with circulation at 5 m. depth on 7 January 2009 at 00:00 is shown in Fig.A1b. Low salinity water coming from the Black Sea occupies the area close to the Bosphorus Strait. High salinity near the southern and eastern coasts implies an upwelling of the Mediterranean origin water in the lower layer. The circulation in NR actually favours upwelling and is dominated mainly by westward currents with two anti-cyclones, one in the northern basin and the other generated by the Bosphorus jet. The NR is more saline compared to the experiments FB001 and

FB002 (see Fig.13) which are performed using the FM. This is because of the water flux term in the surface salinity boundary condition used in FM. However, both model configurations are realistic as discussed in Aydoğdu (2017).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study is a part of the PhD thesis of Ali Aydoğdu. He is funded by Ca' Foscari University of Venice and CMCC during his PhD and REDDA project of the Norwegian Research Council during his Post-doc research period. The National Center for Atmospheric Research (NCAR) is Federally Funded Research and Development Center, sponsored by NSF. We are very grateful to two anonymous referees for helping us to improve the paper substantially.

References

- Altıok, H., Sur, H. İ., and Yüce, H.: Variation of the cold intermediate water in the Black Sea exit of the Strait of Istanbul (Bosphorus) and its transfer through the strait, *Oceanologia*, 54, 233–254, 2012.
- Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Avellano, A.: The data assimilation research testbed: A community facility, *Bulletin of the American Meteorological Society*, 90, 1283–1296, 2009.
- Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, *Monthly weather review*, 129, 2884–2903, 2001.
- Anderson, J. L. and Anderson, S. L.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Monthly Weather Review*, 127, 2741–2758, 1999.
- Arnold Jr, C. P. and Dey, C. H.: Observing-systems simulation experiments: Past, present, and future, *Bulletin of the American Meteorological Society*, 67, 687–695, 1986.
- Atlas, R.: Atmospheric observation and experiments to assess their usefulness in data assimilation, *J. Meteor. Soc. Jpn*, 75, 111–130, 1997.
- Aydoğdu, A.: Advanced modeling and data assimilation methods for the design of sustained marine monitoring networks, Ph.D. thesis, Università Ca'Foscari Venezia, 2017.
- Aydoğdu, A., Pinardi, N., Pistoia, J., Martinelli, M., Belardinelli, A., and Sparnocchia, S.: Assimilation experiments for the Fishery Observing System in the Adriatic Sea, *Journal of Marine Systems*, 162, 126 – 136, <https://doi.org/http://dx.doi.org/10.1016/j.jmarsys.2016.03.002>, <http://www.sciencedirect.com/science/article/pii/S0924796316300045>, progress in marine science supported by European joint coastal observation systems: The JERICO-RI research infrastructure, 2016.
- Aydoğdu, A., Pinardi, N., Özsoy, E., Danabasoglu, G., Gürses, O., and Karspeck, A.: Numerical simulations of the Turkish Straits System for the 2008–2013 period, in preparation, 2018.
- Beşiktepe, Ş. T., Sur, H. İ., Özsoy, E., Latif, M. A., Oğuz, T., and Ünlüata, Ü.: The circulation and hydrography of the Marmara Sea, *Progress in Oceanography*, 34, 285–334, 1994.
- Book, J. W., Jarosz, E., Chiggiato, J., and Beşiktepe, Ş.: The oceanic response of the Turkish Straits System to an extreme drop in atmospheric pressure, *Journal of Geophysical Research: Oceans*, 119, 3629–3644, <https://doi.org/10.1002/2013JC009480>, 2014.
- Chiggiato, J., Jarosz, E., Book, J. W., Dykes, J., Torrisi, L., Poulain, P. M., Gerin, R., Horstmann, J., and Beşiktepe, Ş.: Dynamics of the circulation in the Sea of Marmara: Numerical modeling experiments and observations from the Turkish straits system experiment, *Ocean Dynamics*, 62, 139–159, <https://doi.org/10.1007/s10236-011-0485-5>, 2012.
- Danilov, S., Kivman, G., and Schröter, J.: A finite-element ocean model: principles and evaluation, *Ocean Modelling*, 6, 125–150, 2004.
- Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, *Quarterly Journal of the Royal Meteorological Society*, 125, 723–757, 1999.
- Grayek, S., Staneva, J., Schulz-Stellenfleth, J., Petersen, W., and Stanev, E. V.: Use of FerryBox surface temperature and salinity measurements to improve model based state estimates for the German Bight, *Journal of Marine Systems*, 88, 45–59, <https://doi.org/10.1016/j.jmarsys.2011.02.020>, <http://www.sciencedirect.com/science/article/pii/S0924796311000492>, 2011.
- Gürses, Ö., Aydoğdu, A., Pinardi, N., and Özsoy, E.: A finite element modeling study of the Turkish Straits System, in: *The Sea of Marmara - Marine Biodiversity, Fisheries, Conservations and Governance*, edited by Öztürk, B. et al., pp. 169 – 184, TUDAV Publication, 2016.
- Halliwell Jr, G., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Le Hénaff, M., and Atlas, R.: Rigorous evaluation of a fraternal twin ocean OSSE system for the Open Gulf of Mexico, *Journal of Atmospheric and Oceanic Technology*, 31, 105–130, 2014.

- Halliwell Jr, G., Kourafalou, V., Le Hénaff, M., Shay, L., and Atlas, R.: 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/http://dx.doi.org/10.1016/j.pocean.2014.09.004>, <http://www.sciencedirect.com/science/article/pii/S0079661114001517>, 2015.
- 5 Hoteit, I., Hoar, T., Gopalakrishnan, G., Collins, N., Anderson, J., Cornuelle, B., Köhl, A., and Heimbach, P.: A MITgcm/DART ensemble analysis and prediction system with application to the Gulf of Mexico, *Dynamics of Atmospheres and Oceans*, 63, 1–23, 2013.
- Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an ensemble Kalman filter technique, *Monthly Weather Review*, 126, 796–811, 1998.
- Huang, R. X.: Real Freshwater Flux as a Natural Boundary Condition for the Salinity Balance and Thermohaline Circulation Forced by Evaporation and Precipitation, [https://doi.org/10.1175/1520-0485\(1993\)](https://doi.org/10.1175/1520-0485(1993)), 1993.
- 10 Hüsrevoğlu, S.: Modeling Of The Dardanelles Strait Lower-layer Flow Into The Marmara Sea., Master's thesis, Institute of Marine Sciences, METU, 1998.
- Jarosz, E., Teague, W. J., Book, J. W., and Beşiktepe, Ş.: On flow variability in the Bosphorus Strait, *Journal of Geophysical Research: Oceans* (1978–2012), 116, 2011.
- 15 Jarosz, E., Teague, W. J., Book, J. W., and Beşiktepe, Ş.: Observations on the characteristics of the exchange flow in the Dardanelles Strait, *Journal of Geophysical Research: Oceans*, 117, <https://doi.org/10.1029/2012JC008348>, c11012, 2012.
- Karspeck, A. R., Yeager, S., Danabasoglu, G., Hoar, T., Collins, N., Raeder, K., Anderson, J., and Tribbia, J.: An Ensemble Adjustment Kalman Filter for the CCSM4 Ocean Component, *Journal of Climate*, 26, 7392–7413, <https://doi.org/10.1175/JCLI-D-12-00402.1>, <http://dx.doi.org/10.1175/JCLI-D-12-00402.1>, 2013.
- 20 Korres, G., Ntoumas, M., Potiris, M., and Petihakis, G.: Assimilating Ferry Box data into the Aegean Sea model, *Journal of Marine Systems*, 140, 59–72, <https://doi.org/10.1016/j.jmarsys.2014.03.013>, <http://www.sciencedirect.com/science/article/pii/S0924796314000736>, 2014.
- Masutani, M., Schlatter, T. W., Errico, R. M., Stoffelen, A., Andersson, E., Lahoz, W., Woollen, J. S., Emmitt, G. D., Riishøjgaard, L.-P., and Lord, S. J.: Observing system simulation experiments, in: *Data Assimilation*, pp. 647–679, Springer, 2010.
- Özsoy, E., Di Iorio, D., Gregg, M. C., and Backhaus, J. O.: Mixing in the Bosphorus Strait and the Black Sea continental shelf: observations and a model of the dense water outflow, *Journal of Marine Systems*, 31, 99–135, 2001.
- 25 Petersen, W.: FerryBox systems: State-of-the-art in Europe and future development, *Journal of Marine Systems*, 140, Part A, 4–12, <https://doi.org/http://dx.doi.org/10.1016/j.jmarsys.2014.07.003>, <http://www.sciencedirect.com/science/article/pii/S0924796314001699>, 2014.
- Raeder, K., Anderson, J. L., Collins, N., Hoar, T. J., Kay, J. E., Lauritzen, P. H., and Pincus, R.: DART/CAM: An Ensemble Data Assimilation System for CESM Atmospheric Models, *Journal of Climate*, 25, 6304–6317, <https://doi.org/10.1175/JCLI-D-11-00395.1>, <http://dx.doi.org/10.1175/JCLI-D-11-00395.1>, 2012.
- 30 Sannino, G., Sözer, A., and Özsoy, E.: A high-resolution modelling study of the Turkish Straits System, *Ocean Dynamics*, 67, 397–432, <https://doi.org/10.1007/s10236-017-1039-2>, 2017.
- Schwartz, C. S., Romine, G. S., Sobash, R. A., Fossell, K. R., and Weisman, M. L.: NCAR's Experimental Real-Time Convection-Allowing Ensemble Prediction System, *Weather and Forecasting*, 30, 1645–1654, <https://doi.org/10.1175/WAF-D-15-0103.1>, <http://dx.doi.org/10.1175/WAF-D-15-0103.1>, 2015.
- 35 Seppälä, J., Ylöstalo, P., Kaitala, S., Hällfors, S., Raateoja, M., and Maunula, P.: Ship-of-opportunity based phycocyanin fluorescence monitoring of the filamentous cyanobacteria bloom dynamics in the Baltic Sea, *Estuarine, Coastal and Shelf Science*, 73, 489 –

- 500, <https://doi.org/http://dx.doi.org/10.1016/j.ecss.2007.02.015>, <http://www.sciencedirect.com/science/article/pii/S0272771407000571>, 2007.
- Sørensen, K., Grung, M., and Röttgers, R.: An intercomparison of in vitro chlorophyll a determinations for MERIS level 2 data validation, *International Journal of Remote Sensing*, 28, 537–554, <https://doi.org/10.1080/01431160600815533>, <http://dx.doi.org/10.1080/01431160600815533>, 2007.
- 5 Sözer, A. and Özsoy, E.: Modeling of the Bosphorus exchange flow dynamics, *Ocean Dynamics*, pp. 1–23, <https://doi.org/10.1007/s10236-016-1026-z>, 2017.
- Stanev, E. V., Grashorn, S., and Zhang, Y. J.: Cascading ocean basins: numerical simulations of the circulation and interbasin exchange in the Azov-Black-Marmara-Mediterranean Seas system, *Ocean Dynamics*, pp. 1–23, 2017.
- 10 Tuğrul, S., Beşiktepe, T., and Salihoğlu, I.: Nutrient exchange fluxes between the Aegean and Black Seas through the Marmara Sea, *Mediterranean Marine Science*, 3, 33–42, 2002.
- Ünlüata, Ü., Oğuz, T., Latif, M., and Özsoy, E.: On the physical oceanography of the Turkish Straits, in: *The physical oceanography of sea straits*, pp. 25–60, Springer, 1990.
- Wang, Q., Danilov, S., and Schröter, J.: Finite element ocean circulation model based on triangular prismatic elements, with application in
15 studying the effect of topography representation, *Journal of Geophysical Research: Oceans* (1978–2012), 113, 2008.