'Sampling strategies based on the Singular Value Decomposition for ocean analysis and forecast' by Maria Fattorini and Carlo Brandini

The paper is not suitable for publication. Rejection is suggested.

The paper investigates, for the ocean model, a well--known observation sampling strategy thoroughly assessed in the past for atmospheric model. The subject is not innovative and the investigation does not bring any new findings or useful practice for operational ocean forecast.
Moreover, the paper is not well organized and poorly written, which makes very difficult to follow and properly interpret. There are also many inaccuracies in the terminology used and in some explanations provided. The twin OSSE technique is used in a simplified model context and with perfect observations (unrealistic assumptions) and without any author's attempt to validate the realism of the NR and of the differences between the NR and the analysis fields. The small impact of the different observation sampling on the analysis and forecast (as expected) cannot justify any use or operational implementation of the methodology proposed.

In our opinion, it is not at all true that the sampling strategy used in our paper is well-known. Yes, other authors have used an approach based on SVD to determine the need for additional observations in atmospheric forecasting models. Some of the most significant articles (not all) in this regard have been mentioned in our introduction. It is important to point out that our work does not deal with the topic of targeted observations, i.e. to determine the need for additional observations to be used in DA. Our aim is to determine the best positions in which it is preferable to install a (fixed) in-situ observing network that gives the best benefit for DA. In oceanography, this kind of observation network normally consists in a small number of sensors, due to the considerable cost it has the installation and the maintenance of observation tools at sea, much more than in the atmosphere. We are talking about the optimal placement of a few fixed instruments, and not to locally increase the observations in targeted areas as in previous studies using SVD. In doing this, we did not limit ourselves to a SVD, like other authors, but we include also a criterion based on correlation and associated distance. As far as we know, this criterion is innovative. Using a large set of numerical experiments, we demonstrated that the distribution of observation points determined by this method has a considerable impact, not negligible, especially in relation to its relatively low cost.
Instead, we agree with the reviewer that we were maybe wrong to use the terminology typical of operational models, such as OSSE. Terms like "Nature Run"and "Free Run" are typical of such OSSE experiments, and we maintained the use of such terms in our work just because we think, in the future, that this will be the natural evolution of this work, towards more realistic cases. We intend to eliminate these ambiguities. In our study we simply use a well known idealized reference model (ocean Double Gyre) that has a deterministic behavior but it is

very sensitive to some initial settings (typically viscosity), for which the circulation shows a markedly non-stationary character (Fig 1). As in real cases, the (real) initial state of the system is not known, we initialize our model with climatology, and we estimate the state of the system through a small number of in-situ observations.
We recognize that there is a considerable distance between an idealized circulation model and an operational context. This does not mean that a sampling methodology tested for an idealized model can not be used for more realistic models, and we plan to perform more realistic studies as a future development of this work.
Regarding the writing of the paper, we will try to simplify many sentences and review the overall writing of this article, starting from the useful comments from all reviewers, and –if necessary- with the support of a mother tongue linguistic revision.

1) In the title as in the text, it should be clear than the strategies proposed are relative to observation sampling; 'Observations sampling strategies based on Singular Value Decomposition for ocean analysis and forecast.'

OK, we agree with this suggestion, and we will change the title, and the text

2) The introduction is not well organized and not well explained.
Once the ocean model and assimilation system are defined together with the purpose of the paper to find the optimal observing system network, which improves the analysis and forecast reliability at minimum observational cost; the authors should introduce:
   a) the methodology so far used to validate the assimilation of different observations samplings to forecast improvement
   b) the use of SV for observation sampling strategy
   c) the use of SV in the past for targeting technique and as forecast model fields perturbations for ensemble forecast (there is a clear difference in these 2 applications that the authors should clearly describe).
   d) The organization of the paper.

We apologize for not being able to clarify the difference between using SVs in different applications. We reiterate that this work does not strictly concern the targeting technique, but the design of a fixed measurement network, which is a somewhat different concept. In any case we will try to improve some unclear aspects of the introduction.

Definitions and conclusions on methodologies from cited papers should be correctly explained. Very often the authors provide wrong or not clear explanations of other papers results. The authors also often confuse the analysis estimation with the model or background state.

The word data and observations is both used in the text: for clarity it is recommended to use only the term 'observations'.

OK, we will use only the term 'observations'.

3) Section 2: The authors first introduce the DG model then the NR and FR, finally the OSSE technique and then somehow the assimilation experiments and again the property of the NR and ultimately the observation sampling scenarios. Only in the following paragraph they poorly describe the DA system and the experiments performed to assimilate the simulated observations. Contrary to the actual organization, Section 2 should first introduce the OSSE technique properly referred (e.g Errico, R. M., R. Yang, N. Privé, K.--S. Tai, R. Todling, M. Sienkiewicz, and J. Guo, 2013. Development and validation of observing--system simulation experiments at NASA's Global Modeling and Assimilation Office. *Q. J. Roy. Meteor. Soc*, **139**, 1162--1178. doi: 10.1002/qj2027; Privé, N. C., R. M. Errico, and K.--S. Tai, 2013. The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *J. Geophys. Res. -- Atmos*, 118, 5332--5346. doi: 10.1002/jgrd.50452; Hoffman, R.N. and R. Atlas, 2016. Future Observing System
Simulation Experiments. *Bull. Amer. Met. Soc.*, **97**, 1601--1616. doi: 10.1175/BAMS--D--15--00200.1).

As we said earlier, we think that we should simplify the article by no longer referring to the OSSE technique. The reviewer is right as the OSSE methodology requires great effort in assessing how realistic are the synthetic data obtained from the model w.r.t. the observations, etc. BUT we did not make any OSSE, we only used some terms typical of the OSSE (Nature Run and Free Run), clarifying their meaning in this context. This confusion will be eliminated in the final version of this paper.

Then the description of the models used by the authors to represent all the OSSE components should be provided: NR, assimilation system and simulated observations. Finally, they should describe the assimilation experiments and the type of the observations simulated (instrument, variable observed, spatial distribution and frequency). It is not absolutely clear how the assimilation experiments were performed: length of the assimilation window (only in the
results it is said that the window is 5 days long);
in fact, it is 5 days
duration of the analysis experiments.
How many assimilation cycles were performed?
One assimilation cycle for each experiment. Each experiment characterized by a given number of observations extracted at fixed points.
Each cycle is made by up to 10 inner cycles and 2 outer cycles, unless a convergence criterion is satisfied before.

Were the observations repeatedly simulated and assimilated for weeks?
Experiments were repeated for different and independent time windows (10) without considering the results of DA for model re-initialisation.

How many analyses and forecast were performed?
We use 10 not consecutive assimilation windows lasting 5 days, each of them has one analysis cycle and one forecast.

Were the locations computed every analysis cycle?
No, locations were fixed for each different experiment.
Which observation type was simulated?
In this work, we only simulated velocities observations (e.g. ADCP-like observations). At the beginning, we also tried to assimilate temperature and salinity profiles, but in this particular model this doesn't give any benefit, because this system is predominantly a barotropic one.

Moreover, the authors should properly explain the purpose of introducing the FR, which starting from some averaged state, will be effectively filtered of small scales. Considering the idealized set up of the OSSE, validation of such an experiment is difficult to accomplish, however, it is necessary, to at least test how much the initial conditions of the experiment differ from the truth (NR), and if the differences/errors are at all realistic or at least useful for the study they are proposing.

We will simplify the article by no longer referring to the OSSE technique.

4) Results and conclusion

As the authors state in the conclusion, it is very difficult to get robust conclusion to be useful in an operational context from an idealized assimilation configuration were the ocean forecast model used is simplified and the observations assimilated perfect.
In reality, the impact of the observations in the analysis and forecast depend on the observations influence in the estimate, which is in turn depending on the observation and model error correlation and variance. The number and observation type assimilated can change the assimilation quality.
For example, a similar study would provide different results if on top of velocity observations in a random or according to SVD location, salinity and temperature observations were also routinely assimilated.
Also how feasible is a SVD technique in operational context: for example, how the observations are deployed in the computed locations and how often are deployed?

We are talking about a limited number of observations located at fixed observation points, to give design criteria for in-situ network. This kind of study, in our opinion, it is absolutely feasible at a relatively low cost, as it only needs a climatology or other kind of solution taken from operational oceanography

products in use. Obviously for real ocean system, we agree that a rigorous OSSE must be adopted.

From previous studies, It has been found that the quality of the analysis strictly depend on the background quality, which has been proved to improve with high density, homogeneous distribution and different varieties of satellite observations assimilated. Variational assimilation systems using million of observations every cycle have been proved being unaffected by the assimilation of targeting observations (Hamill, T. M., F. Yang, C. Cardinali, and S. J. Majumdar, 2013: Impact of targeted Winter Storm Reconnaissance dropwindsonde data on midlatitude numerical weather predictions. Mon. Wea. Rev., 141, 2058–2065, doi:10.1175/MWR--D--12--00309.1.; Majumdar, 2016: A review of targeted observations, *https://doi.org/10.1175/BAMS--D--14--00259.1*).
Our problem is not the same and the methods adopted in the mentioned paper have nothing to do with our work.

From previous studies, It has also been understood that effort should be spent on DA and model development to better describe not only the model dynamics but also background and model error, observation error correlation and observation biases. All these aspects are more important to taken into consideration than targeting observation strategies, which have been proved inefficient for operational purposes.

In conclusion, considering all the past experience acquired on the use of observations in numerical atmospheric models, I am quite skeptical on the use of targeted observations or targeted observation sampling strategies (measuring only specific areas and levels) for model forecast.

We are not working on targeting observation strategies, but we think that the operational oceanography context, in which we have much less observations than in atmosphere, is somewhat different.

Detailed comments:
1)Pag.1 line 25 to 27: the sentence should be rephrase as 'improve the forecast reliability when the model forecast is properly initialized with fields obtained through a data assimilation procedure. Data assimilation (DA) scheme, in fact, combines observations and model first guess (typically few hours forecast) weighted by their respective accuracies to obtain the best unbiased estimation of the ocean state. In operational practices, '

We can rephrase this sentence, although we would like to express a more general concept. Data Assimilation, and in particular variational assimilation, is not only suitable for improving initial conditions, but even boundary conditions, forcings, model parameters.

Pag 1 line 27: 'In the operational practice, a common procedure of initializing a simulation starting by external data (e.g. climatology, objective

analysis, model analysis, etc.) requires a spin--up interval during which the solution is not useable: assimilation of suitable data can strongly reduce model errors and hopefully produce more reliable solutions.'
This sentence is very confusing: spin up is a different problem than model error. In general, DA schemes provide a solution that should avoid any spin up but would not eliminate the model error.

You are right, this sentence is confusing. In our opinion, it is very important to use DA to reduce spin-up time (that can be very long, much longer than in atmospheric model). Variational assimilation is particularly helpful since the solution is consistent with physics (many older assimilation methods, still used in oceanography, like OI or nudging, are not). This is very important to better initialize models. Obviously DA is aimed at improving the quality of analysis and then forecasts. "Model error" itself is not reduced by DA (although, by the "weak constraint" formulation of 4D-VAR even "model error" can be reduced). In any case, "model uncertainties" are reduced by DA.

Page 1 line 30 to 33: this sentence is very confusing. The authors should clearly say which observations can be assimilated in an ocean model, satellite-- and ground--based and shortly explain advantages and limitation of both observing systems. Please rephrase.

In operational oceanography, data used for DA are usually: SST and SSH (or SLA). This is a huge amount of data, and - depending on the specific application – they need a proper treatment of data. This information is limited to the surface. It is very important to assimilate in-situ data (mainly CTD profiles and ARGO buoys) and velocities (ADCP, lagrangian drifters). Coastal radars (HF radars) are becoming more and more important, but they only measure surface velocities. In this paper we concentrated only on ocean in-situ data, these have only a local representativeness and present some disadvantages (installation, maintenance): their cost is very relevant.
Is it so necessary to better specify all that?

Page 1 line 33 to 36: Please rephrase as 'The main limitation of in--situ observation networks is the high cost for installation and maintenance over time; it is very important, therefore, to design an observing system, which maximize the impact of the observations in the forecast and minimize the cost.
OK

Pag 2 line 1 to 5: the sentence is not clear, please rephrase as 'The impact or benefit of the observations can be measured as an improvement of the analysis and forecast reliability. It is important that the criteria used to measure the observations benefit are objective and easily implementable in a operational context.
So the major problem is to identify an observation network configuration that provides the best impact once the observations have been assimilated.'

OK

Pag 2 line 7: delete 'assimilated'. The sentence line 5 to 7 is again not clear and should be rephrased.
Ok

Pag 2 line 7 to 10: the sentence is wrong. In a DA scheme, the observations correct the trajectory (first guess) according to their influence that mainly depends on the observation and model error covariance matrices. Clearly, if the observations are located in areas where error in the initial condition is fast growing, they can better control this growth.

You are right. We will try to improve the clarity of this sentence

Pag 2 line 11 to 16: the sentence is very confusing, what is the perturbation theory? Why do the author say ' in fact' and then introduce the main sources of model errors? Please rephrase.
We will try to improve the clarity of this sentence

Pag 2 line 23: please modify the sentence 'lied in the elements of the predicting system' with 'embedded in the predicting system'
OK

Pag 2 line 29 to 31: What the sentence means? Again very confusing.
OK, We will revise this sentence.

Pag 3 the sentence 'Observing tools can be deployed inside a verification area to minimize the forecast error covariance, estimated by ensemble, with respect to a set of possible observation deployments (i.e. a set of observation operators) by computing how the ensemble members are transformed through DA (Buehner and Zadra 2006).' Is very unclear. Which observing tool can be deployed in a verification area? Do you mean observations deployed in an area? Minimize the forecast error covariance? Do you mean minimize the forecast error or the spread of the forecast error? Observations or observations operator? And what does it mean that the ensemble members are transformed by DA?

You are right, we attempted to simplify this sentence.

Pag 3 the sentence 'In literature the behavior of perturbations in dynamical system was mainly addressed by stability analysis, in that instability is linked to the existence of exponentially growing normal modes in the linearized dynamical equations'
What does it mean? Which perturbations? Do the authors want to say that the instability is linked to the normal modes? How?
What does mean 'the perturbations dynamics' Do the authors want to say 'the perturbations of a dynamic system?'

You are right, we attempted to simplify this sentence.

Pag 3 line 18: 'reset' . Do the authors mean 'further developed'? Pag 3 line 33: 'the state estimate' should be the analysis estimate or the initial condition estimate OK

Pag 3 line 34: 'Observation networks could be designed to catch, as frequently as needed, real data in specific areas characterized by strong perturbation growth' should be rephrased as 'Observing systems networks should be optimized to particularly observe areas where model inaccuracies can fast growing'
OK

Pag 3 line 36: delete 'criteria'
OK

Pag 3 line 39: 'sampling' should be 'assimilating'
OK

Pag 3: the sentence 'SVD was used for others different applications such as perturbing the initial state in ensemble forecasting' is wrong.
It is not the initial state that is perturbed but the forecast. In the ensemble analysis is instead the initial state that is perturbed.
You are right.

Pag 4 line 3: the sentence 'A review of experiments of sampling' should be 'A review of observation sampling'
OK

Pag 4 line 4: should be 'in selecting the observations'
OK

Pag 4 line 13: should be 'additional observations from aircraft'
OK

Pag 4 line 19: the sentence 'and the number of additional observing tools to be deployed'
is it 'observations to be deployed'?
We don't understand, sorry. Observing tools in oceanography are usually said to be "deployed", such as lagrangian drifters and profilers.

Pag 4 line 20: should be 'Studies conducted at ECMWF'
OK

Pag 4 line 21: what does mean 'The question of predictability is strongly related to both observations and assimilation scheme used to synthesize initial conditions' ?

Atmosphere predictability is not related to the model and observations used.
You are absolutely right.

Pag 4 line 24: change the sentence to 'any observation sampling strategy'
OK

Pag 4 line 34: change the sentence to 'an optimal observation sampling strategy'
OK

Pag 4 line 35: the sentence 'By optimal sampling, we mean a strategy based on the use' should be 'In particular, an optimal observations sampling should use'
OK

Pag 4--5: what does it mean 'We have tested a strategy based on the SVD of the linearized dynamical operator whose validity requires sufficiently small error to ensure the model linearization around the background trajectory relevant respect to the dynamics of linearized system around the real system trajectory.'?
OK, we need to simplify this sentence.

Pag 5 line 5: correct to 'the ocean model ROMS'. Please refers ROMS as model and not code (there are many other 'code' to correct in the paper)
OK
Pag 5 line 10: 'possible strategy for model sampling'. Do you mean observation sampling?
Yes, you are right.

Pag 5 line 20: 'In fact, the DG dynamics has been largely studied as a theoretical scheme of typical seasonal and inter--annual oscillations in the large--scale circulation observed in mid--latitude oceans.'
The sentence is not clear, is it: 'DG simplified dynamics has been used to understand the ocean seasonal and inter--annual oscillations of the large scale circulation.'?
We change the text in: 'DG simplified dynamics has been used as an idealized model to reproduce the ocean seasonal and inter-annual oscillations of the large scale circulation.'

Pag 5 line 31: 'This configuration has been also reported by ROMS developers (Moore et al. 2004) to describe the functionalities of specific modeling tools such as the tangent linear model and the adjoint model of nonlinear code, as well as other tools like the SVD of the tangent linear propagator.' Which configuration? Reported? Specific modelling tools? Also the adjoint can only be applied on linear operator. I do not know what they authors mean.

Simply that: we consider the tangent linear model and the adjoint of tangent linear model, derived from the original nonlinear code, as modeling tools developed and documented in the work done by other scientists.

2) Pag 7 line 3: 'dirtied' is not the correct word. Observations are weighted by their accuracy or penalized by their error variance.
Also a perfect observation has 0 error variance, which technically cannot be strictly applied in a variational DA scheme. Can the authors explain how they did it?
You are right, we modified the text.

Pag 7: 'The control vector corresponds just to the free--run (background) initial conditions since the reality has been constructed by using the same model and parameters. Hence we assume that the model is able to reproduce all the processes of our virtual reality, as also forcing and boundary conditions are the same.'. What does it mean? Is a background field or an initial condition field. Control vectors are the model parameters minimized during the assimilation process, what is this control vector been introduced?
The authors assume that the NR is close to reality; this cannot be assumed, it must be proved. It is one of the very crucial aspects of the OSSE to prove that the NR is able to resolve the particular phenomenon, situation and pattern is going to be studied (see above suggested papers on OSSEs).

You are right, but as we said before, we understand that the use of some terms typical of the OSSE can be confusing. It is not the goal of this work to demonstrate how realistic is a mathematical model (Double-Gyre), used by so many authors in the past for studies concerning the dynamics of geophysical systems.

Pag 7 line 12: 'velocity observation'? do the author mean 'wind observation'? u and v component? Which instrument type and where is based? How many ocean layer are observed?
Velocity observations are in the form of ocean current profiles, extracted from what we call the NR (synthetic data). Typically such profiles are obtained by Acoustic Doppler Current Profiles (ADCPs). We extracted this data at all ocean layers of interest, at each observation position.

Pag 7 section 2.2: 'the incremental formulation of the 4DVar implemented in ROMS (ROMS--IS4DVar) and based on the Lanczos algorithm: it identified iteratively the incremental vector to correct the control vector which minimizes the cost function, as explained in (Moore et al. 2011). The minimum of the cost function, which corresponds to the maximum likelihood between model and data, is obtained by searching for the zero of the gradient of the cost function. The control vector is represented by the initial state vectors, so just the initial conditions are adjusted by data.' ;
this sentence is very confusing, the terms definition is wrong and the explanation at least inappropriate.

The Lanczos algorithm is a minimisation algorithm, the control vectors are the model parameters to be minimised and the minimisation is performed in terms of 'incremental approximation', that is, the minimisation problem is written as a function of the departure from the background. The approximate minimisation problem thus defined is solved using an iterative algorithm: this is the inner loop of 4D--Var. Usually, a preconditioned Lanczos--conjugate gradient algorithm is used to solve the inner loop minimisation problem. After this minimisation, the departures and trajectory can be recomputed using the nonlinear model and a new linearised problem is defined. The process can be repeated: this is the outer loop of incremental 4D--Var. If the linearised problem is reasonably close to the nonlinear problem its solution should be an approximation of the solution of the nonlinear problem. At the next outer loop iteration, the starting point is closer to the solution. The algorithm should converge to the solution of the nonlinear problem, although there is no general theoretical proof of convergence.
If necessary we can work on a more accurate and precise description of what the minimization algorithm does, however this algorithm is not part of our work.

In the following lines (26 to 38) the authors provide a very detailed description of the B matrix but do not provide any information on the length of the assimilation system, the number of assimilations performed and the description of the observation assimilated. It is only later on, in the results section that by chance they say the assimilation window was 5 days. Is it really 5 days?

Yes, 5 days.

Pag 9 line 5 to 8: 'As described in the previous sections, our experiment consists in sampling, through different possible strategies, an idealized ocean system (DG), then to assimilate such data in a twin model, and finally to identify which extracted dataset gives the maximum benefit both in the most correct identification of the final state, both in the subsequent forecast. A first possibility is to sample the system randomly, using a limited number of observation points.'
The authors study consists of 1) find an optimal observations distribution to best represent the initial ocean state. 2) Such observations set, once assimilated by using the OSSE identical twin experiment, would provide 3) a set of forecasts that once compared with the truth (NR) will identify the best forecast and therefore 4) the best observations sample assimilated.

Thanks for suggestion, we are attempting to improve this part of manuscript.

Pag 9 line 13: 'We started with 20 observation points and this test was repeated several times with different datasets. In fact, randomness can produce datasets more or less impactful for DA, hence in order to remove the impact of such chance factor, the test has been repeated considering

different positions of observations, creating in this way an ensemble of analyses (Fig. 2).'
I think the authors simply wanted to say that 'different random observation sampling set were provided'.
How many? Pag 9 line 20: 'Figure 2 shows, by mean of a Taylor diagram of the results, how analysis depends strictly by the position of observations, as it shows a wide spread between the analyses produced by assimilating different datasets corresponding to different networks having the same number of observing tools' Which observations tools?
Current velocity profilers

Page 9 line 25: 'statistical skills'. RMSE, STD and correlation are not skills but statistical indices used to quantify the forecast skill.
You are right. We will modify the text accordingly.

Figure 1 should well explain the domain of the basin, axes and contours. Are the authors showing only the surface currents? What about the depth? And the ocean circulation?
Yes, in the maps we only showed surface currents. This (simplified) ocean system is strongly barotropic and no relevant differences were found in different vertical layers.

Figure 4 should be eliminated and explained without showing; there are too many figures in the paper.
We don't agree. Fig. 4 is very useful to better understand the influence of the number of observations on the analyses. Eventually we can eliminate other figures if there are too many, but not this one.

The authors wrote: 'However, information obtained from points too close to each other is likely to be too correlated (i.e. redundant).
Correlation does not mean redundant: observation error correlation if properly treated in the assimilation system provides a good information, which increases the influence of the observation in the fit.

Our results show clearly that –assimilating a limited number of observations – it is better to place them not only in areas characterized by strong perturbation growth (identified by SVs) but also at positions sufficiently far from each other, in order to avoid redundancy that could result from concentrating too many observations in the most active regions of the flow.