Interactive comment on "Using Canonical Correlation Analysis to produce dynamically-based highly-efficient statistical observation operators" _by_ Eric Jansen et al.

Anonymous Referee #1

Received and published: 19 February 2019

General comments

The paper describes a method for producing efficient estimates of skin and subskin SST using the Canonical Correlation Analysis (CCA) statistical technique. This is described in the context of observation operators which are used in data assimilation schemes to provide information about the model-observation misfits in order to correct the model. The aim of the paper is to make better use of satellite SST observations in data assimilation systems compared to most existing operational systems by improving the way the diurnal cycle of SST is represented. The work is therefore valuable and the paper is well written.

Near the end of the last section, there is a reference to another paper submitted by Korres et al. which appears to address a very similar topic by the same list of authors. This other paper should be referred to in the introduction and the aims of the two papers differentiated clearly. It might not be possible to address some of the comments below without using results presented in the Korres et al. paper. If that is the case perhaps the authors should consider merging the papers or having a two part paper.

The SOSSTA project is presented in three separate papers. The first (Pimentel et al. 2019) describes the modelling of the diurnal cycle in the Mediterranean Sea. The second (this paper) describes the method of building an observation operator by parameterising the results of an external model. It uses the dataset from the first paper as an example to demonstrate the method. The third paper (Korres et al., 2019) uses the GOTM datasets and the method presented in this paper and applies it in the POSEIDON data assimilation system. We believe that the topics are sufficiently different and self-contained to warrant publish them as separate papers. Moreover, we are hoping to publish additional papers such as Korres et al., 2019 that document the application of the SOSSTA operator in other models/systems.

Specific comments

1. The paper describes the method in a general sense with SST as a "use case". It might help the flow of the paper if the use of the method for SST was more central to the paper so that, for instance, the introduction would have more information about the literature on SST analysis and data assimilation. The way it is currently structured, more review is needed on the various other types of meteorological and oceanographic observation operators used, for instance radiative transfer models.

The technique for building a parameterisation using canonical correlation analysis on an existing dataset is very general and can be applied in many situations. To emphasise the general nature of the method we decided to first describe the method and then focus on one specific application.

2. The scheme is not tested in a data assimilation system (although that appears to be done in the Korres et al. paper). Would the aim of a scheme using the CCA method be to correct for errors in the diurnal cycle of SST at 1 m depth using all available satellite SST measurements, or to correct the model's foundation SST?

The aim of the observation operator is to provide the best possible estimate of SST from the model background, based on the actual atmospheric conditions. The correction to the model temperature depends somewhat on the modelling of the model covariance in the data assimilation system. In general the correction will affect all the layers that in turn are affected by diurnal variability

3. There are shortcomings in the design of the experiments to test the performance of the CCA method in section 4.3. The validation is performed over the same period as the model data used to generate the CCA OO, and then compared to the same model data. There is no comparison of the CCA results to real observations.

The purpose of Sect. 4.3 is to show how the simplified parameterisation of the CCA method is able to approximate the full GOTM model. As the training dataset contains O(1000) profiles and the parameterisation has only O(10) parameters, there is no risk of overfitting.

Nevertheless, we agree that a demonstration of the method on a separate dataset would make the message of Sect. 4.3 stronger and remove doubts that people may have about the method. We have decided to revise Sect. 4.3 and to use an independent dataset for the comparisons. This independent dataset is obtained by withholding every other profile (along the zonal direction) in the input dataset from the calculation of the CCA OO.

4. The comparisons with other SST assimilation methods described in section 5 assume that the GOTM model and the atmospheric forcing driving it are correct. If the skin model in GOTM had a bias for instance, could it be worse to use the CCA OO based on that model than the other methods tested? A more independent way of assessing the CCA OO method based on GOTM is needed, e.g. by comparing to real observations.

This is a valid point and we fully agree. We have revised Sect. 5 of the paper and the comparison with other methods is now performed using the SEVIRI L3C dataset of subskin temperature. The performance metrics are calculated using only the withheld profiles, as described above at point 3.

5. There is not any discussion of the need for the adjoint of the observation operator in data assimilation (which is obviously very efficient when using the

CCA method).

We have rephrased the final comment of Sect. 3 to be more inclusive of 4D variational schemes: "The CCA OO is straightforward to implement in this scheme, since for H' and its adjoint H'T it follows that: H' = MT, H'T = M"

6. The GOTM dataset used for training the CCA is a model and not based on observations. The work of Pimentel et al. 2018 describes how the model represents the skin and subskin SST, but some more information is required here to justify its use. A brief description of how the model has been made to represent the skin and subskin would help (the highest level of 1.5 cm is not at the same depth as the skin or subskin). A summary of the assessment of the model compared with real observations is also needed here, otherwise there is no link to the real world.

We have expanded Sect. 4.1 and included the following: "The subskin SST represents the temperature at the base of the conductive laminar sub-layer of the ocean surface; for practical purposes it is represented by the temperature of the top model layer of GOTM (1.5 cm). The conductive sub-layer of the air-sea interface, associated with the cool-skin effect, is parameterised and dynamically computed within GOTM to produce a modelled skin SST. Further details are provided in Pimentel et al., (2019)"

Technical comments

1. Pg 2, line 5. The wording "are not or not sufficiently modelled" is a bit confusing on first reading. This lack of process representation in the model is often included in the representation error in data assimilation systems. A discussion on the relationship between the complexity of the observation operator, and the inclusion of representation error in the R matrix would be good here.

We did not include a discussion of the representation error as we felt that this would be beyond the scope of this paper. You are right that missing or insufficiently modelled processes should be taken into account in the covariance matrix. However, increasing the uncertainty is not an alternative to actually modelling the missing processes. This is especially true for processes that are not random but that create strong systematic biases, such as the diurnal cycle of SST that we discuss in Sect. 4.

We have included the following paragraph in Sect. 4: "Errors due to e.g. limited spatial resolution or unrepresented processes are generally included in the representation error. Representation errors have been extensively discussed within ocean applications (Oke and Sakov, 2008; Janjíc et al., 2018). However, the diurnal variability of skin SST represents a potentially systematic error that requires a proper treatment rather than just the increase the representation component of the observational error."

2. Pg 2, line 6 -7. Does the cost of the "second" model depend on the observation number as implied here, or the (horizontal) model grid size? The cost of these models, e.g. a diurnal model, is cited as one of the justifications for implementing the CCA method. It is not obvious that a simple diurnal model is that expensive compared to the cost of the full GCM.

The use of an observation operator scales with the number of observations to be assimilated. When talking about the computational cost we are referring to state of the art diurnal models such as the presently used GOTM. Of course simpler models/parameterisations exist that are less costly to run, in fact the parameterisation used in the SOSSTA operator could be regarded as such a model.

3. Pg2, line 7. "needs" to "need".

Thank you, it has been corrected.

4. Pg 3, eq (3). Normally matrices would be in uppercase but here you start using lowercase letters. This is particularly confusing when you use uppercase and lowercase of the same letters (e.g. u and v).

Unfortunately, the common notation for canonical coordinates and for singular value decomposition both use the letter v. We understand that this may be confusing and have decided to rename the canonical variables F and G. This removes the double usage of letters and allows us to capitalise all matrices.

5. What are the implications of eq (11)? It is taking into account the biases in the training observations and model. These presumably are not constant in time so how can this be applied in practice? There is no description of how these values are calculated in section 4, or their magnitude.

As the aim of CCA is to find correlations between datasets, it only considers variations of the variables with respect to their mean value. The CCA procedure subtracts the mean of each level, so the matrix M by itself would only relate temperature anomalies to each other. The offset factor K adds the mean values in order to relate temperature values instead of anomalies. It does not represent a form of bias, even in a perfect world K is not expected to be 0. The calculation is done according to Eq. 11, using the mean temperature values of the two observation levels ($\bar{Y}$) and the mean values of each model level ($\bar{X}$).

6. Pg 7. Where does eq. (14) come from?

H' is the tangent-linear version of H, as defined in Eq. 13.

7. Pg 5, 1st paragraph. There is no description of the near-surface temperature structure to introduce the reader to terms like "skin" and "subskin".

We have rephrased and expanded the second paragraph of Sect. 4 to make the definition of skin and subskin more clear: "The different types of sensors used on satellites probe the ocean temperature at different depths. Infrared (IR) sensors measure the temperature at about 10um, a layer that is referred to as the ocean skin. Microwave (MW) sensors on the other hand measure the temperature of the layer below that, the subskin, with a depth of about 1mm. This is well below the vertical resolution of an OGCM, while these layers are strongly affected by the atmospheric conditions. [...]"

8. Pg 5, line 11. "At the same time. . .". Not always at the same time.

Rephrased: "At the same time, wind can mix ..."

9. Pg 5, line 13. "straightforward assimilation". I think you mean here that it is a problem when the observations contain significant diurnal cycle changes at the skin or subskin, and that is not accounted for when comparing the observation with the model. A straightforward approach could be to remove those observations from the assimilation as you mention later on.

By "straightforward assimilation" we mean assimilating the observations into the model without rejecting or correcting observations that are affected by the diurnal cycle.

10. Pg 6, line 8. You take the daily mean value for wind and insolation. How good is this for determining the magnitude of the diurnal cycle?

Using the mean values is of course an approximation, but one that has been used throughout the literature (e.g. Gentemann, 2003) for describing the diurnal cycle. The dependence of the magnitude of the diurnal signal on the wind and insolation categories of the CCA OO is shown in Fig. 1.

The parameterisation of the CCA OO bases the (sub)skin temperature estimate mostly on the shape of the temperature profile in the upper ocean layers. The categorisation in wind and insolation serves to group together similar profiles to allow for a better parameterisation. The use of the mean values is of course an approximation, but as can be seen from Fig. 1 it works reasonably well to separate the different magnitudes of the SST diurnal cycle.

11. Fig 1. What depth is the diurnal cycle that is plotted?

Added "at the subskin level".

12. Pg 6, line 11. How do you get skin and subskin estimates from GOTM?

The subskin SST represents the temperature at the base of the conductive laminar sub-layer of the ocean surface; for practical purposes we have represented this by the temperature of the top model layer of GOTM (1.5cm). The conductive sub-layer of the air-sea interface, associated with

the cool-skin effect, is parameterized and dynamically computed within GOTM to produce a modelled skin SST. Further details are provided in Pimentel et al., 2019.

13. Sec 4.3, second paragraph. Wouldn't it be fairer to do the validation of the CCA on a different time-period to the one used to build the statistics for the CCA OO?

The results in this section have been redone using profiles that were withheld from the calculation of the CCA. See also the answer to specific comment 3.

14. Section 5 is a comparison of the CCA OO to other methods. I think its title should be changed.

The title of the section has been changed to "Performance and discussion"

15. Pg 10, line 2. Waters et al (2015) assimilate data during the day where the wind speed is high.

Indeed, Waters et al. (2015) assimilates at nighttime and also during high winds. We have clarified this in the text and included also a figure showing the comparison between SOSSTA and the upper model level method in the afternoon.

16. Last paragraph section 5. The last sentence of this paragraph describes the method that should be used in the paper to assess the performance of the CCA OO.

In the revised version of the manuscript the validation and comparison with other methods is only done using profiles that are withheld from the operator calculation.

17. Pg 11, line 24. Can you include a reference for the magnitude of the diurnal cycle?

In this location we have added a reference to Pimentel et al. 2019. The discrepancy discussed at this point is between the upper model level and the (sub)skin temperature, i.e. $\max(T_{\mathrm{skin}} - T_{1.5\mathrm{m}})$. As the model captures some of the variability, this is typically smaller than the magnitude of the diurnal cycle discussed in literature $T_{\mathrm{skin,max}} - T_{\mathrm{skin,min}}$.

Furthermore, we have added to Sect. 4 where the amplitude is discussed: "Under favourable conditions this amplitude is typically of the order of a few degrees (see e.g Flament (1994)), but values as high as 6˚C have been observed (Merchant, 2008)."