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

Anonymous Referee #2

Received and published: 26 February 2019

GENERAL COMMENTS

This paper essentially present a method to assimilate satellite SSTs using canonical correlation analysis (CCA); a statistical technique that permits to construct an appropriate observation operator to project the state variables of a numerical model onto observed variables. In this case the method is used to correctly assimilate measurements (satellite skin sea surface temperatures) which are not included in the set of those simulated by the model (the temperature of the first model layer) but which are physically linked to them. Based on a previous paper of Pimentel et al (2018) a 1D model (GOTM) is used to simulate high-resolution temperature profiles from which it is possible to extract temperature at the model levels (the state variable of the model) and "observed" temperatures at the sub-skin level. The skin temperature is obtained, from the sub-skin temperature, using the Fairall 1996. CCA OO is based on these simulate data rather than on real measurements.

Considering that skin temperature measurements are very rare, the idea to build a matchup dataset using a specialized model is quite interesting and represents a good compromise between the two extremes of using only very few in situ data or assimilate satellite skin or sub-skin SSTs as if the they were bulk SSTs.

The validation is based on the evaluation of ability of SOSSTA in reproducing the GOTM derived skin and sub-skin temperatures (fig. 3). SOSSTA is presented in what appears to be the companion paper of this paper still submitted to Ocean Sciences by the same authors. Probably, the author should better clarify the relation between the two papers, and rather than merging the two, indicate in some way that they are part I and II of a single subject.

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.

One last, more general comment is about the fact that implications for the diurnal cycle are well discussed in the paper also in relation with satellite data in section 4 but geo-stationary satellite, and SEVIRI in particular, are never mentioned while they should represent an interesting source of information for the assessment of the proposed assimilation procedure. However Seviri SSTs are distributed as subskin-sst. This is clearly declared in the MSG SST reprocessing ATBD v1.1, 31/5/2016 Algorithm Theoretical Basis: "Since the coefficient of the SST algorithm are established using in-situ measurements, the retrieved SST is considered to be the sub-skin SST. One could apply a -0.17C (Donlon et al., 2002) to get the skin SST. However this offset is only a very rough conversion term valid at largescale for wind speed exceeding 6 m/sec." (osi-saf v1.1, 31/5/2016). If this sentence is correct one should verify if IR SSTs are to be considered skin or sub-skin SSTs.

Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J. (2002). Toward improved and validation of satellite and sea surface and skin temperature and measurements and for climate and research. Journal of Climate, 15:353–359.

Following your suggestion we have included a paragraph about SEVIRI in the introduction of Sect. 4: "One of the most important sources of SST observations is the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument onboard the Meteosat satellites of the second generation. As these are geostationary satellites, SEVIRI can provide continuous measurements of the same area with a 15-minute temporal resolution. Although the IR imager is sensitive to skin temperature, the calibration algorithm of SEVIRI corrects for the cool skin bias and the resulting SST products should be considered as subskin temperature (Saux-Picart et al., 2018). For wind speeds greater than 6 m/s the skin temperature may be calculated as Tskin = Tsubskin – 0.17˚C (Donlon, 2002), but this is only an approximation."

Overall I would say that it is a good paper that deserves to be published on Ocean Sciences doing some minor revisions as suggested in this review.

SPECIFIC COMMENTS:

Page 5, section 4 Use case: satellite SST, The authors write: "Although diurnal variability is included to some extent (Marullo et al., 2014), the vertical resolution of the OGCMs is still insufficient to fully resolve the variability of the skin and subskin ocean temperature", To resolve Skin a sub-skin is only matter of resolution or some more physics is still needed?

Vertical resolution is of key importance, but more physics is also needed. For example, the physics of the conductive skin-layer is diffusion-dominated. Pimentel et al, 2019, also explore the influence of the penetration of solar radiation within the near surface as well as the sea surface albedo. There are feedbacks in that an improved resolution of SST improves air-sea heat flux calculations.

Page 5 section 4.1: “The top 75 m of the water column is resolved using 122 vertical layers with fine resolution near the surface and gradually becoming coarser with depth. The uppermost 1 m contains a total of 21 layers, with the highest level at 1.5 cm depth”. Considering that 1.5 cm is not enough to resolve the skin and sub-skin can the author justify the choice of 122 vertical levels with the highest level at 1.5 cm depth? Is this due to computation capabilities or, for some numerical or physical reason, it makes no sense to use higher resolutions?

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.

There is flexibility within GOTM to explore even higher resolution, although it is not clear whether this would be justified. Additional model layers can be included, although this increases data handling and storage needs.

Page 6 section 4.2: “Under certain conditions the ocean skin may even cool down below the bulk temperature. “. “certain conditions” are related to latent heat loss.

Rephrased "Due to latent heat loss, the ocean skin may even cool down below the bulk temperature"

Page 10, section 5, lines 7-8. “This can be explained by the cool-skin effect that is included in GOTM and which plays a role also at nighttime”. Here you can cite figure 4 of Donlon et al 2002 (see reference above),

In the revised version the comparison with skin temperature has been removed, as SEVIRI provides only subskin.

Section 6, Discussion. The skill of CCA OO respect to some other method is measured using GOTM as a reference. As I already noted in the general comments this choice is, in some sense, obliged by the scarcity of in situ skin or sub-skin SST measurements. But, what about the possibility to use meteosat data as a reference?

In the revised version of the paper we are calculating the skill of the CCA OO compared to other methods using the SEVIRI L3C dataset of subskin temperature. To obtain an independent dataset we withhold every other profile (along the zonal direction) in the input dataset from the calculation of the CCA OO. The performance metrics are calculated using only the withheld profiles.

Bulk SSTs at about 20 cm of depth are routinely measured by drifters. Drifter SSTs are used to continuously assess the validity of satellite SST products, distributed by agencies or the Copernicus Marine Service (CMEMS). Can the proposed CCA OO method also contribute to adjust drifter SSTs to skin or sub-skin temperature making more correct the comparison with the satellite estimates?

We have shown that the CCA OO method can be used to project modelled upper-ocean temperature profiles onto the skin and subskin temperature. We believe that the same approach could also be applied to temperature profiles measured by drifters.