Ocean Sci. Discuss.,
https://doi.org/10.5194/os-2018-170-AC2, 2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Ocean Science Discussions

Interactive comment on "Assessment of ocean analysis and forecast from an atmosphere-ocean coupled data assimilation operational system" *by* Catherine Guiavarc'h et al.

Catherine Guiavarc'h et al.

christopher.harris@metoffice.gov.uk

Received and published: 12 April 2019

We are grateful to Reviewer2 for the time and efforts assessing our work, and for providing comments. These comments are reproduced below, with our responses indicated with asterisks.

General Comments: This paper presents an evaluation of the ocean component of an atmosphere-ocean coupled prediction system that has been running in real-time at the UK Metoffice and builds upon the work presented by Lea et al. (2015). In particular, this system employs a "weakly coupled" data assimilation approach, which affects the analysis and forecast quality in a number of ways. A systematic approach is used to

Printer-friendly version

Discussion paper

provide an objective evaluation of the different state variables (sea surface temperature, mixed layer depth, currents) against accepted reference datasets, including other operational forecast products. The presentation is generally clear, but remains fairly descriptive, limiting the conclusions that can be drawn from the study. Many of the differences between the coupled and operational forecasting systems are due to differences in particular parameter choices and not directly due to the coupled model or "weakly coupled data assimilation" themselves. As a result, the value of the paper is limited to a presentation of the baseline skill of the system. Even this is put somewhat in question, as the results (for currents and sea level) are shown to be sensitive to differences in data assimilated in the two systems resulting in much larger errors in the coupled system. The paper would benefit significantly from having additional experiments to clarify the impact of the different parameter choices and a more in-depth analysis of results. Additionally, the paper would benefit from some reorganization as the differences in fluxes are described near the end of the paper, much of which is needed to understand the differences being presented in previous sections. I recommend the paper for publication but strongly encourage the authors to deepen the analysis of the results and provide additional experiments to shed-light on the source of differences.

**Experiments are currently being undertaken using an ocean-only model to look at the impact of reducing the assimilation window from 24 to 6 hours. These experiments should decompose the effect of the shortened assimilation window from the effect of the coupling and aid in the interpretation of the comparisons to the FOAM system in the manuscript. It is important to note that the scheduling of the operational CPLDA system and the availability of observations will not be investigated in these experiments. It is hoped that the results from these experiments will be available before the final resubmission of the manuscript and the text edited accordingly, but expect this would only be possible if an additional month was provided to allow time for this to be done.

We agree with reviewer2 about the organisation of the paper and we will move the

CC BY

section on surface fluxes before the velocities. We will also add a paragraph on the SLA results.**

Specific Comments:

Section 1: The introduction is limited to activities at the UK Metoffice and would benefit greatly from some context into why weakly coupled data assimilation is of interest and previous efforts. For example, a study has recently been published that should be referenced in this discussion (Browne et al., 2019). Moreover, several centers are now running coupled forecasting systems operationally, which should be mentioned to provide context into UK Metoffice efforts. Browne, P.A., de Rosnay, P., Zuo, H., Bennett, A. and Dawson, A., 2019. Weakly Coupled Ocean–Atmosphere Data Assimilation in the ECMWF NWP System. Remote Sensing, 11(3), p.234.

**To provide more context the Met Office efforts, we will update the introduction and in particular we propose to add:

" Coupled systems are used in wide range of applications (short and medium range forecasts, seasonal forecasts, climate prediction and future scenario projections) and improving the initialization of these systems can play a significant part to reduce the development of errors. Using separate atmosphere and ocean analysis to initialize a coupled system can result in an imbalanced system. The imbalance can cause an initialisation shock that could potentially increase the development of errors during the forecast. Using a coupled data assimilation approach has been shown to reduce this initialisation shock (Mulholland, 2015). In the recent years, numerous research centres have developed operational coupled data assimilation systems. For instance, the European Centre for Medium-Range Weather Forecast (ECMWF) has developed a weakly coupled ocean-atmosphere data assimilation system (Browne, 2019). Their results show that using their coupled data assimilation system reduces forecast errors compared with forecast initialised from uncoupled analysis. The National Centers for Environmental Prediction (NCEP) is using a coupled data assimilation system for

CC BY

seasonal and sub seasonal scale predictions (Saha et al, 2014). Penny et al (2017) proposes an overview of the efforts made on coupled data assimilation in operational centres. It shows the diversity of the approaches available. At JAMSTEC, they developed a low resolution strongly coupled system used for experimental seasonal and decadal prediction while NRL coupled model is initialised by separate analysis but uses a high-resolution ocean component (1/25). " **

Pg1, L19: Johns et al. (2012) is a technical report that appears not to be publically available. There are a number of such reports referenced here. If these reports are only available upon request, than greatly detail regarding their content should be included in the text. Also for Lea et al. (2013).

**We propose to add details for both references:

"In Johns et al (2012), a large set of global coupled atmosphere-ocean-sea ice 15-day hindcasts were initialised from separate atmosphere and ocean-ice analysis. Compared to uncoupled forecast, the coupled 15-day hindcasts show improved forecasting skills especially in the Tropics, region strongly coupled where small changes in SST can exert a major influence on patterns of convection and remote responses to diabatic heating. Both atmosphere and ocean biases benchmarked against control simulations suggest a more vigorous water cycle, improved large-scale circulation, enhanced convection, stronger teleconnections and somewhat improved representation of the Madden-Julian oscillations. They conclude that significant benefits should arise from full integration of coupled NWP system."

"Lea et al (2013) developed a coupled data assimilation (DA) system with the aim of improving the initialisation of coupled forecasts for various time ranges from short-range out to seasonal. They used a "weakly" coupled data assimilation approach whereby the coupled model is used to provide background information for separate ocean/sea-ice and atmosphere/land analyses. The increments generated from these separate analyses are then added back into the coupled model. They investigate the impact

CC BY

of weakly coupled DA on analysis and forecast skills up to 5 days. They showed that ocean and atmospheric analyses produced using coupled DA to be very similar to the analyses produced using separate ocean/atmosphere data assimilation" **

Pg2, L14: "To avoid duplication of previous work . . .we did not investigate the impact on the diurnal cycle already covered in Lea et al., (2013)". This is a fairly important point that is referred to later in terms of errors in temperature and mixed layer depth. Since Lea et al., (2013) is only a technical report these results could be included here (or at least explained more clearly).

**We propose to add: "They investigated the impact of the initialisation on the diurnal cycle of SST has been investigated by comparing to geostationary satellite data from the SEVIRI. They looked at the diurnal SST range and showed that the initialisation of the coupled forecast has very little impact on the diurnal range. Both Coupled and uncoupled forecasts appear to represent the regions of larger diurnal cycle observed by the satellite well, although they perhaps under-estimate the range in the central North Atlantic " **

Pg. 3, L1: COARE3.0, please add a reference. Also this is not noted in Table 1.

**This reference will be added Fairall, C.W., E.F. Bradley, J.E. Hare, A.A. Grachev, and J.B. Edson, 2003: Bulk Parameterization of Air–Sea Fluxes: Updates and Verification for the COARE Algorithm. J. Climate, 16, 571–591, https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2

We will update table 1 accordingly. **

Pg. 3, L9: Lea et al. (2015) note significant difficulties associated with runoff. How is runoff handled here? Are there differences between FOAM and CPLDA?

Due to the issues in Lea et al. (2015), we are for the moment using the same climatological run off as in FOAM. Since Lea et al. (2015), runoff from the atmosphere/land component may well have improved and plan to revisit using this in future.

CC BY

Pg. 3, L32: “Haney retroaction”. Please provide a reference and more details. I.e. what is the timescale of the relaxation? To what fields?

**A Haney flux correction (Haney, 1971) is applied to the sea surface salinity (SSS) based on the difference between the model and climatology. The magnitude of the restoring on SSS is -33.33 mm/day/psu.

Haney,R. 1971. Surface boundary conditions for ocean general circulation models. J. Phys. Oceanogr. 1: 241–248. **

Pg.4, L19: It says here that CPLDA uses a fixed freezing temperature, which differs from what is indicated in Table1.

Reviewer is correct the table 1 will be corrected.

Pg. 5, L27-32: It may be useful to note here that there is also an impact on the observation cutoff time that affects the number of SLA observations used.

**We propose to add this: “It is important to note that difference in observation cutoff due to difference in operational scheduling impact the number of observations used in each system.” **

Pg. 6, Sec2.6: If the CPLDA system has been running operationally since October 2016, why is it necessary to run additional experiments? If the aim here is to benchmark the skill of CPLDA as compared to FOAM why not use the operational runs? If iit possible to make additional runs, could a longer run with the SLA observation cutoff corrected be produced? This would make the evaluation much more straightforward.

Since October 2016, the operational CPLDA suite had several updates and bug fixes so we could not have a long enough period with consistent settings to be able to properly assess the system. We will not be able to produce a longer run with the corrected cutoff but results from the Operational CPLDA system with the updated scheduling allowing more SLA observations to be assimilated now shows similar SLA statistics than the FOAM system.

CC BY

Pg. 6, L17: “There are a few differences. . .”. What is the impact of these differences? If they are not too significant an analysis of the operational run compared to FOAM for Oct. 2016 onwards could add additional insight to the paper.

**There are differences between the operational system in 2016 and the same system in 2018 (several upgrades and bug fixes). The 2015 experiment was made to assess the 2018 version of the operational system. As noted in the manuscript some differences are unavoidable (different observations available) and the other main difference is the scheduling that was update in the operational system in June 2018 as a result of the assessment of the 2015 experiment. **

Pg. 7, L7: SLA is an important variable providing information regarding the circulation and mesoscale activity. It is also relevant here to provide context to the errors in 15m velocity. I would strongly encourage you to include these results, despite the poor results. Ideally, an additional run could be made with the cutoff issue corrected.

**The current CPLDA operational system with the updated scheduling allowing more observations to be assimilated has SLA statistics similar to FOAM. A section on the SLA results will be added to the manuscript.

”CPLDA 2015 and FOAM sea surface height are assessed against observations using class4 statistics (Ryan et al., 2015). The observations used are provided by CMEMS and include data from Altika, Cryostat and Jason-2 satellites. Altimeter bias correction is applied to the observations. For each model comparison against the satellite observations, it is important to use the model own altimeter bias. The altimeter bias contains information from the model mesoscale so correcting observations using the altimeter bias from one model to assess a second model penalised this second model. Figure 1 shows a timeseries the sea level anomaly (SLA) difference statistics assessed against CMEMS observations. In the 2015 experiment with the old scheduling, CPLDA SLA root-squared-mean error (RMSE) is significantly larger than FOAM SLA RMSE (Figure 1a). The larger RMSE in CPLDA can be attributed to the difference in the number

CC BY

of SLA observations assimilated by both systems. Figure 2 shows the number of observations assimilated by both systems in 2015 the number assimilated by CPLDA is significantly smaller than the number assimilated by FOAM. Differences in scheduling can explain this. In comparison with FOAM, the CPLDA 2015 experiment "best analysis" runs earlier in the day so fewer observations are available. Following these results, the scheduling of the CPLDA operational system was updated in April 2018 in order to improve CPLDA performance. The best analyses at 0600Z, 1200Z and 1800Z are now delayed allowing more observations to be assimilated. This change along with a change in the Met Office database (MetDB) that now allows a more frequent ingestion of SLA observations, has resulted in a significant reduction of the CPLDA RMSE (see figure 1b).

Figure 1: SLA class4 statistics with respect to CMEMS satellite product (Jason2, Cryosat and Altika) for CPLDA and FOAM. a. For the long CPLDA experiment in 2015 with the old scheduling. b. from the operational systems for 2018, the CPLDA scheduling was updated to allow more observations to be assimilated in April 2018.

Figure 2: Number of SLA observations assimilated in FOAM and CPLDA in the 2015 experiment" **

Pg. 7, L16: It is argued that sea ice is not evaluated because of a difference in freezing temperature. However, there are differences in many other aspects as well (bulk formulae, SLA observation cutoff, assimilation window). If the aim is to benchmark the performance of the system, than an evaluation of the sea ice component should be included as well.

**We propose include a figure illustrating the sea ice results. "Sea ice:

Figure 3 shows the sea ice extent and volume respectively for CPLDA and FOAM best analyses. Figure a shows that the sea ice extent is similar between CPLDA and FOAM and is also comparable to OSTIA. Figure b shows that the sea ice volumes are similar in the Arctic but in the Antarctic CPLDA has a reduced volume compared to FOAM

CC BY

in the through Southern Hemisphere winter and spring. Differences in volume can be attributed to differences in sea ice thickness probably caused by differences in freezing temperature treatment (see section 2.3 Sea Ice component). Figure 4 shows the sea ice extent forecast for the melting season in Arctic (August) and Antarctic (February) for CPLDA and FOAM. Both models have the tendency to melt too much ice during the forecast especially in Arctic, in the Antarctic the exaggerated melting in CPLDA is slightly reduced in FOAM."

Figure 3: Time series of Sea Ice extent (left) and sea ice volume (right) for CPLDA (2015 experiment) and FOAM

Figure 4: Sea ice extent forecast for CPLDA (2015 experiment) and FOAM during the melting season (February in Antarctic and August in Arctic) **

Pg. 7, L20: Is it the instantaneous SST or a daily average?

We use the daily average it will be clarify in the text.

Pg. 8, L9: Should this be 0.01K? The global mean difference in bias between CPLDA and PSY4 appears to be less than 0.05K. Also, if OSTIA is cooler than CPLDA and PSY4 assimilates OSTIA, why is PSY4 warmer than CPLDA?

**Global average difference between CPLDA and OSTIA is 0.06K the text has been updated. Using Class4 statistics from the GODAE model inter-comparison (see figure 5 - not for inclusion in paper) PSY4 (purple) is generally biased warm and OSTIA(grey) is generally biased cold relative to surface drifters, this is consistent with the results presented for 2015. This reasons for the difference in SST characteristics between the systems despite PSY4 assimilating OSTIA could be due to model biases, atmospheric forcing etc but diagnosing these is outside the scope of this paper.

Figure 5. Global SST class4 bias compared to drifting buoy observations **

Pg. 8, L19: The Fiedler reference is only "in prep". I recommend providing a more thorough explanation with a "personal communication" to Fiedler.

CC BY

We propose to add:"The SST spatial variability for each of the experiments was determined by wavenumber auto spectra analysis. All spectra results were calculated on the native grid of the analysis (ORCA025). For each region of interest, the spectra were calculated along the horizontal coordinates closest to a given. The power spectral density of each latitude was calculated using the Welch's average periodogram method, with 28 degrees of freedom (DOF) and Hanning window.The overlapping length for each band averaging was half the band length (Fiedler et al, submitted, personal communication)"

Pg. 9, Fig. 2d: What is the global mean value of this? It would be helpful to include this on all four panels. It appears to me that it is positive, which seems in conflict with Fig. 3b that shows CPLDA has a more negative global near-surface temperature increment.

**Global mean values have been added to all of the panels (at present to the legend but will be added to the figure titles before the final manuscript is submitted) The global mean value of the difference between CPLDA and FOAM surface T increments is -0.01K which is consistent with Fig 3b. **

Pg.10, L5-14: This asymmetrical effect of SST increments is quite an important difference between CPLDA and FOAM and warrants further explanation. Perhaps an example figure showing profiles and how the increments deteriorate the profile. However,the argument is based on the fact that large increments in CPLDA then create a large asymmetrical effect on SST. But why are the increments larger in the first place?

**This section has been re-written, see below. The larger increment is likely trying to correct a warm temperature bias due to the under estimation of wind stress in CPLDA.

"Temperature and Mixed layer depth The temperature of CPLDA is assessed against Argo profile observations provided by CMEMS. The class 4 global temperature statistics for the best analysis are presented in Fig 3a. The results for the forecast (not shown) are similar to the best analysis with the mean bias staying unchanged and the RMSE slightly increasing. CPLDA has a cold bias in the subsurface which is maximum

CC BY

(around 0.1K) at 10 m and present in the thermocline down to approximately 50 m. A smaller warm bias is present at around 100 m, but this is over a greater range of depths so represents a large amount of heat. A subsurface cold bias is present in the ocean-only FOAM system but with a smaller amplitude. For the CPLDA results shown, the number of profile observations being assimilated was smaller than in FOAM due to differences in scheduling of assimilation cycles. Even in tests with the operational scheduling (referred to in section 2.6) the sub-surface bias in CPLDA persists despite the increased number of observations being assimilated. The sub-surface bias can be attributed to the vertical propagation of the surface temperature increments through the water column. King et al (2018) shows that the succession of positive and negative temperature increments has an asymmetric effect on the vertical temperature structure due to the way the temperature increment at the surface is propagated to the bottom of the mixed layer. A negative surface increment weakens the stratification and so deepens the mixed layer, this means that a subsequent positive surface increment is projected deeper. Due to the shorter assimilation cycle in CPLDA (6 hours), relative to FOAM (24 hours), the increments exhibit more temporal noise which leads to the larger subsurface bias observed. In both CPLDA and FOAM the temperature RMSE is largest in the thermocline which highlights its variability, the RMSE is larger in CPLDA due to the over-deepening of the mixed layer leading to a misplacement of the thermocline. CPLDA has slightly larger RMSE than FOAM in the sub-surface due to the propagation of surface increments through the mixed layer described above. Some options for reducing the asymmetric effect of the SST increments in the sub-surface were described in King et al (2018). These may be tested and implemented in future versions of the CPLDA system. Mixed Layer Depth statistics confirm that CPLDA has a deeper MLD than the assimilated profile observations. The mean error against those assimilated observations is -5.2 m while the RMSE is 34.7 m. As expected, the CPLDA MLD is deeper than FOAM (Fig 4). As for CPLDA, the MLD in FOAM is deeper than in the observations but both the mean error and the RMSE are reduced (-2.1 m for the mean error, 32.6 m for the RMSE). The deeper MLD in CPLDA could be caused by

CC BY

the asymmetric effect of the sub-surface temperature increment on vertical temperature structure but also by differences in wind stress between the two system. Further experiments running FOAM with a 6 hour assimilation time window (consistent with CPLDA) are needed to help to separate the impact of the assimilation time window from the impact of the wind stress. In CPLDA the large negative increment applied at the surface (Fig 3b) is propagated down to approximately 50 m, below this a small warm increment is applied down to approximately 150m. The fact that the negative increment is projected deeper in CPLDA than in FOAM and the dipolar structure in the vertical is consistent with idealised experiments into vertical propagation of temperature increments, (King, pers. comm. 2018). CPLDA has a larger negative increment at the surface than FOAM, this is likely due to the under-estimation of the wind stress in CPLDA, described in section 3.4, which causes a warm bias that the assimilation is trying to correct. Below approximately 200-400 m the magnitude of the average temperature increment is small (Fig 3b) and the increments applied by CPLDA and FOAM are similar." **

Pg. 10, Fig. 3: What is the explanation for the large increments between 50-100m and associated increase in RMSE in CPLDA? Is it related to the asymmetric effect of SST increments? If so, this should be described.

This section has been re-written (see above) and the increased RMSE addressed.

Pg. 10, L16: How is the mixed layer depth calculated? The Fig. 4 caption quotes Kara (2000), but some explanation of the methodology used would be appropriate.

**For the stats presented the Kara mixed layer uses a density based criteria. It takes the top most observed profile density (calculated from T and S using the Jackett and McDougall 1994 equation of state) (as long as it is shallower than 10m) and calculates the depth of the first observation where the density is (first) greater than the density resulting from an decrease in temperature of 0.8 degrees. C compared to the surface value. A linear interpolation of the depth of this observation and the one above gives

CC BY

one observation of MLD from one temperature and salinity profile observation. (No observation results from a temperature only profile). The same procedure is done for the model equivalent temperatures and salinities which gives the model equivalent MLD. The results of many profiles are gathered to give the statistics of MLD.**

Pg. 10, L17: "mean error .. is -5.2m". The text (Pg. 10, line2) notes that CPLDA has a deeper mixed layer than FOAM, shouldn't this value be positive?

The error calculated is observation minus model that why the values are negative. We agree it is not clear and will change the text accordingly (5.2 m too deep instead of -5.2).

Pg. 11, Fig. 4b: Please state whether positive values are deeper or shallower in CPLDA.

Positive value mean CPLDA is deeper. We will add this to the legend of the figure.

Pg. 11, L2-4: It is difficult to follow this reasoning when we haven't yet seen the differences in wind stress. This section would be easier to follow if Sec3.4 was presented first providing context regarding differences surface fluxes.

We agree with the reviewer and we will change the order of the sections by putting the surface fluxes section before the velocities.

Pg. 11, L7: "Predictions of ocean current are important for marine activities". It would be helpful to expand somewhat since this statement is vague. Also, it should be "currents".

We propose to add: "Predictions of ocean currents are important for marine activities. They are used for a number of practical applications such as ship routing, marine search and rescue, pollution monitoring, offshore oil and gas industry, marine renewable energy."

Pg. 11, L19: Are differences in tropical Pacific along capable of affecting global statis-

CC BY

tics this much? Or are there large differences elsewhere as well?

**Yes, there is no statistically significant differences in the other regions. We propose to add :

“In other regions than the Tropical Pacific no significant differences are observed, CPDLA and FOAM statistics are similar.” **

Pg. 11, L21: If the shorter window affects the increments in a direct manner that influences the results to this degree, I think it warrants some further description. Perhaps some examples of increments in this area. This is extremely relevant to the overall effort to produce a “weakly coupled” data assimilation system and should be elaborated upon.

**It is thought that this problem is less to do with the shorter window and more to do with the availability of SLA observations due to the timeliness of the obs sources assimilated, the updated scheduling somewhat alleviates the problem. This will be confirmed by the assimilation window experiments detailed above. **

Pg. 11, L21: By how much are the number of observations limited? This too could be expanded upon to clarify the impact of the assimilation window on results.

Plot of the number of SLA observations has been added to the SLA section, there is also a slight reduction in the no of profile observations assimilated in CPLDA (5-10%).

Pg. 12, L3: If the operational system uses the updated scheduling than why not evaluate it instead of the experiment with the reduced number of observations?

**The schedule of the operational system has been updated as a result of the assessment made with the 2015 experiment as it highlighted than the scheduling at the time limited the number of observations impacting especially the SLA and currents. At the time of writing we only had a short period of the operational system with the updated schedule. **

Printer-friendly version

Discussion paper

CC BY

Pg.12, L7: If the unrealistic currents are not always caused by the reduced number of SLA observations, than perhaps the observations are simply covering up an existing issue in the coupled model. This issue warrants further investigation, it its central to many of the results presented.

The issue is to do with not having the profile observations to constrain the changes in density structure caused by the assimialtion of SLA observations, it is hoped the experiments looking at the assimilation window length will shed some light on this.

Pg. 14, L8-9: A “reduced heat loss” should lead to warmer SSTs shoudn’t they?

This is a mistake it should say:” The reduced heat gain compared to FOAM between 30S and 30N contributes to the colder SST observed in CPLDA”

Pg. 14, L9-10: While the differences in short wave radiation and latent heat clearly contribute to differences in net heat flux difference. These fields both have large scale patterns, whereas there is a notable small scale pattern in the net heat flux difference not explained.

The difference is due to the difference in the representation of the mesoscale between the 2 models, eddies are not located in the same exact positions as there are not completely constraint by the observations (DA).

Pg. 17, L25-27: As noted above, the argument about the noisier increments is quite important to the overall conclusions. I would recommend a deeper analysis be made of this issue.

See above, this is being investigated.

Pg. 19, L8: should be “well-placed”

It will be corrected

Interactive comment on Ocean Sci. Discuss., https://doi.org/10.5194/os-2018-170, 2019.

CC BY

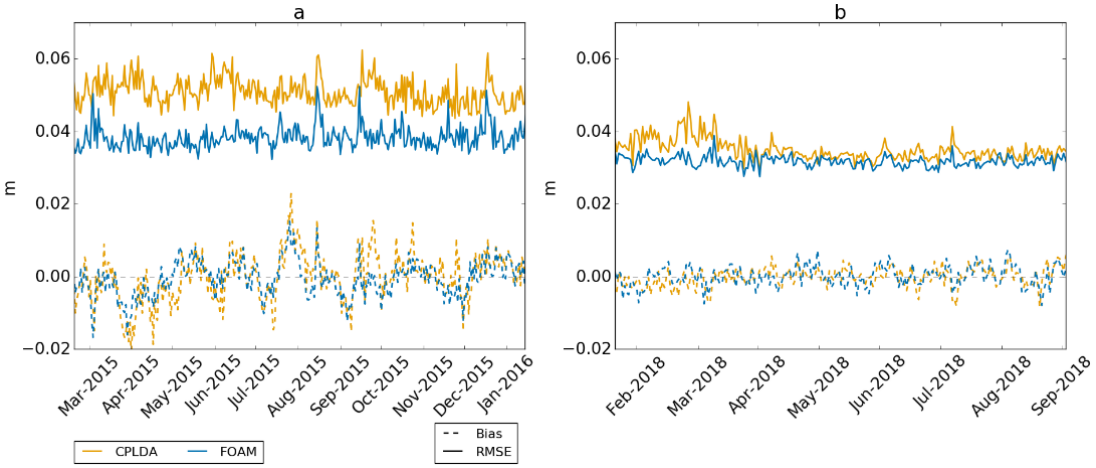

Fig. 1. SLA class4 statistics with respect to CMEMS satellite product (Jason2, Cryosat and Altika) for CPLDA and FOAM. a. For the long CPLDA experiment in 2015 with the old scheduling. b. from the operational

Printer-friendly version

Discussion paper

CC BY

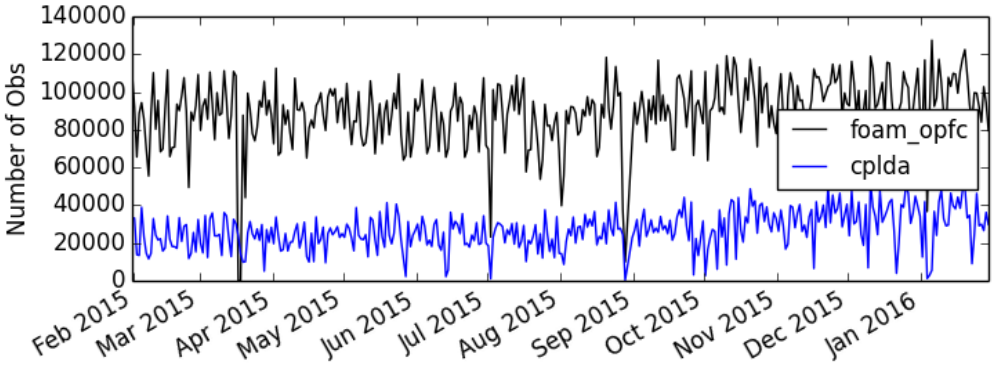

Fig. 2. Number of SLA observations assimilated in FOAM and CPLDA in the 2015 experiment

Printer-friendly version

Discussion paper

CC BY

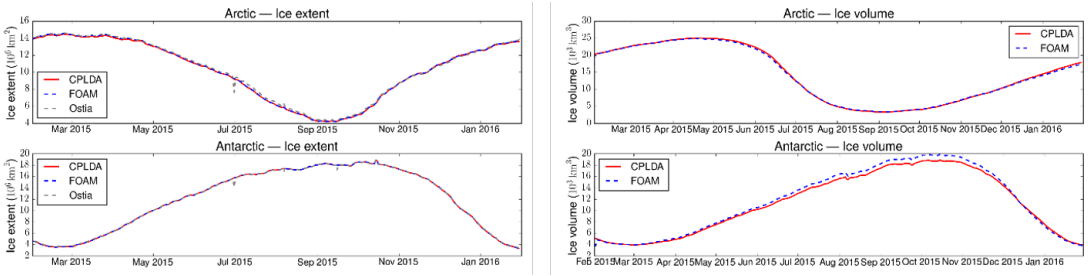

Fig. 3. Time series of Sea Ice extent (left) and sea ice volume (right) for CPLDA (2015 experiment) and FOAM

Printer-friendly version

Discussion paper

CC BY

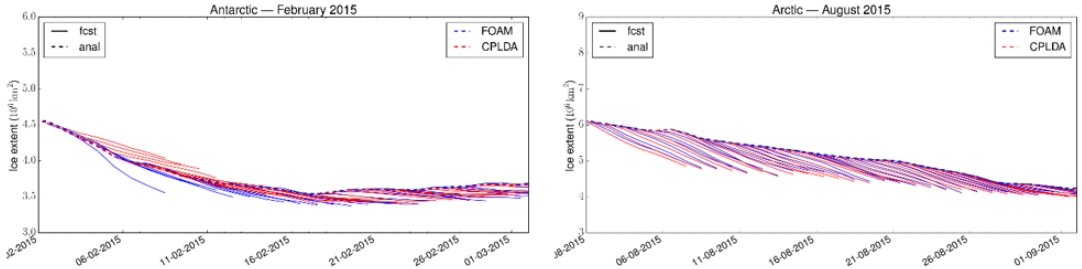

Fig. 4. Sea ice extent forecast for CPLDA (2015 experiment) and FOAM during the melting season (February in Antarctic and August in Arctic)

Printer-friendly version

Discussion paper

CC BY

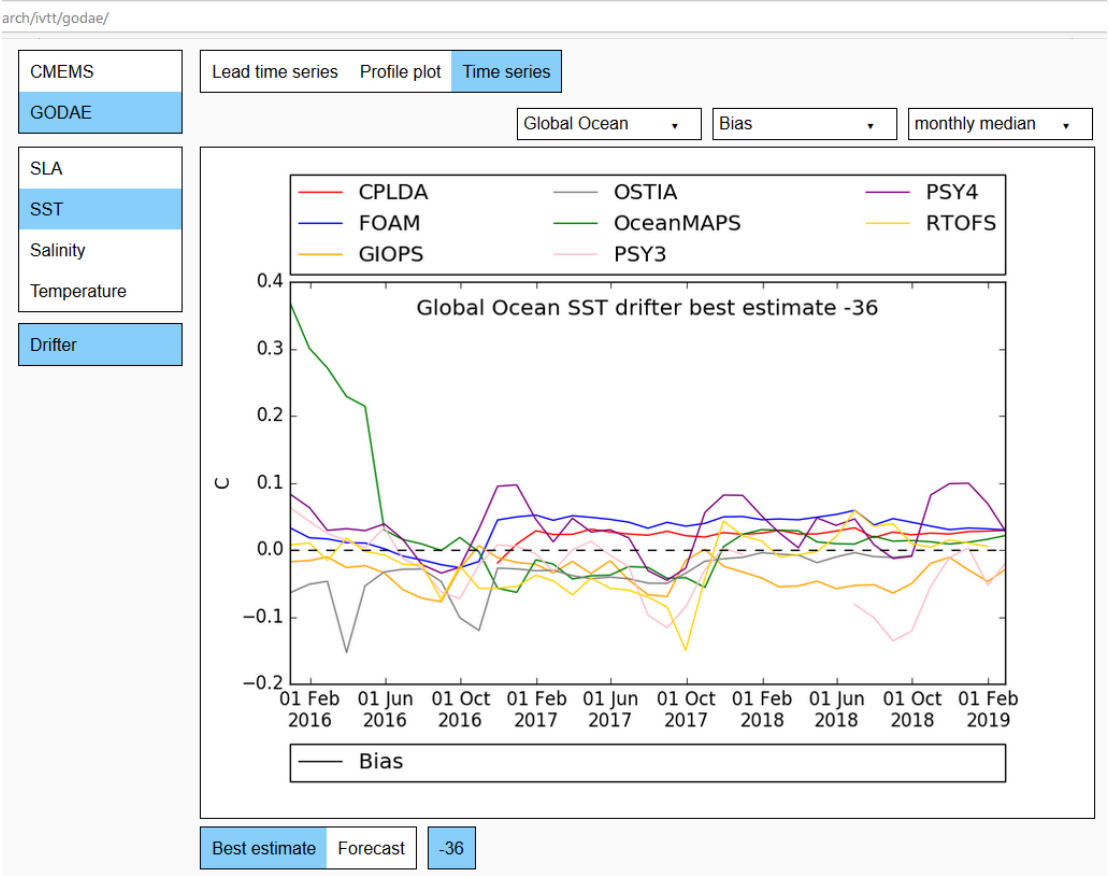

Fig. 5. For information not inclusion

Printer-friendly version

Discussion paper

CC BY