Anonymous Referee #1

The arcticle of “On the time and length scales of the Arctic sea ice thickness anomalies: a study based on fourteen reanalyses” aims at the Arctic sea ice thickness (SIT) which contains considerable uncertainty in the popular 14 reanalyses. They evaluate the reproduced SITs from the reanalyses, and then investigate the e-folding time and length scales of the SIT anomalies. Clearly, these topic and the consequent findings are helpful to deep understanding the SIT and the concerned variability.

We thank the referee for his time and for the detailed revision of our manuscript. We appreciated very much her/his comments, which were all taken into account in the revised version of the manuscript. Below, we answer point-by-point all specific and technical comments.

It is worthwhile mentioning that in order to better incorporate the referees’ suggestions, the manuscript’s structure has changed. New figures were added, while others were replaced. Thus, all tables, figures, pages and lines referred to in this rebuttal letter are directed to the updated version of the manuscript, unless otherwise stated.

1) Undoubtedly, one conclusion is “reanalyses built with sea ice data assimilation present shorter time and length scales”. However, all the reanalyses were only assimilation of sea ice concentration, and the inferred conclusion is not based on the direct comparison of with and without assimilation in a same system frame. The more proofs based on the sea ice concentration will be helpful to increase the rationality on physics. So the counterpart analysis on sea ice concentration shown in Fig. 2 and Fig. 3 will be robust.

We fully understand the referee’s comment, since we had the same concern during our analyses: **What are the time and length scales of the Arctic Sea Ice Concentration (SIC)?** Nevertheless, after some tests, we realized that it wouldn’t be possible to apply such concepts for this sea ice parameter due to the characteristics of the sea SIC time series. The SIC is not a continuously varying property. For instance, at the region covered by the perennial ice, the SIC is expected to be (nearly) 100% for the all year round. In this case, for instance, the time scale would be “infinity”.

For the BGEP mooring’s location, notice that the SIC can be nearly constant for several months and it suddenly drops during the summer months (Fig. A, top panel). When calculating anomalies (Fig. A, bottom panel), the values remain near to zero for the months marked by nearly constant SIC. We believe the concepts of time and length scales, as they were explored in the paper, are not applicable for the SIC time series.

2) Section 3.2 illustrates the intercomprison of the reanalyses. The current main features shown by Fig. 5 is not meaningful enough: some reanalyses are very close… In this section, more other information about SIT and its anomaly need to be added.

In the second version of the manuscript, we are bringing new elements for discussion. For instance, we have now compared the reanalyses based on several other parameters and specifications (Figs. 12 and 13). In addition, an extended discussion is presented in Sec. 4.

Firstly, the ensemble mean SITs based on with and without assimilation will be useful (P 15 Line 10: “This suggests that higher length scales are associated with thicker ice”). Furthermore, it will be complementary of the previous knowns in Uotila et al. (2018) and Johnson et al. (2012).

Johnson, M., Proshutinsky A., Aksenov Y., Nguyen A. T., Lindsay R., Haas C., Zhang J., Diansky N., Kwok R., et al.: Evaluation of Arctic sea ice thickness simulated by Arctic Ocean Model

Intercomparison Project models. J. Geophys. Res., 117(C8), C00D31, doi:10.1029/2011JC007257, 2012.

Secondly, the standard deviations of the two ensembled SIT anomalies were not shown before and would be interested to the reader to know the variabilities or the distinguishes about the SIT anomaly in the reanalyses and even considering with and without assimilation.

Fig. B (top row) shows the ensemble mean from all reanalyses (top left), as well as the ensemble mean from all systems with (top center) and without (top left) data assimilation. Overall, there is a good correspondence between the patterns of mean SIT in the three panels, but with slightly thicker ice for the systems with data assimilation, which is mainly distinguished off the northern Greenland coast and around the Canadian islands. Fig. B (bottom row) displays the averaged standard deviation from the three groups of reanalyses (Ensemble, DA and NA). As for the mean fields, there is not a big difference between the three panels. Since these fields don't add much in order to distinguish systems with and without sea ice data assimilation, we preferred to leave this figure here in the discussion.

Nevertheless, we are taking referee's suggestion further in order to better understand the link between Mean State and TS/LS in the new Fig. 13, and its respective discussion in Sec. 4. For this ensemble of reanalyses, where each system considers several different parameters compared to the other systems, it is not clear what is the impact of the variability on the scales, as shown in Fig. D from the rebuttal letter for the second referee.

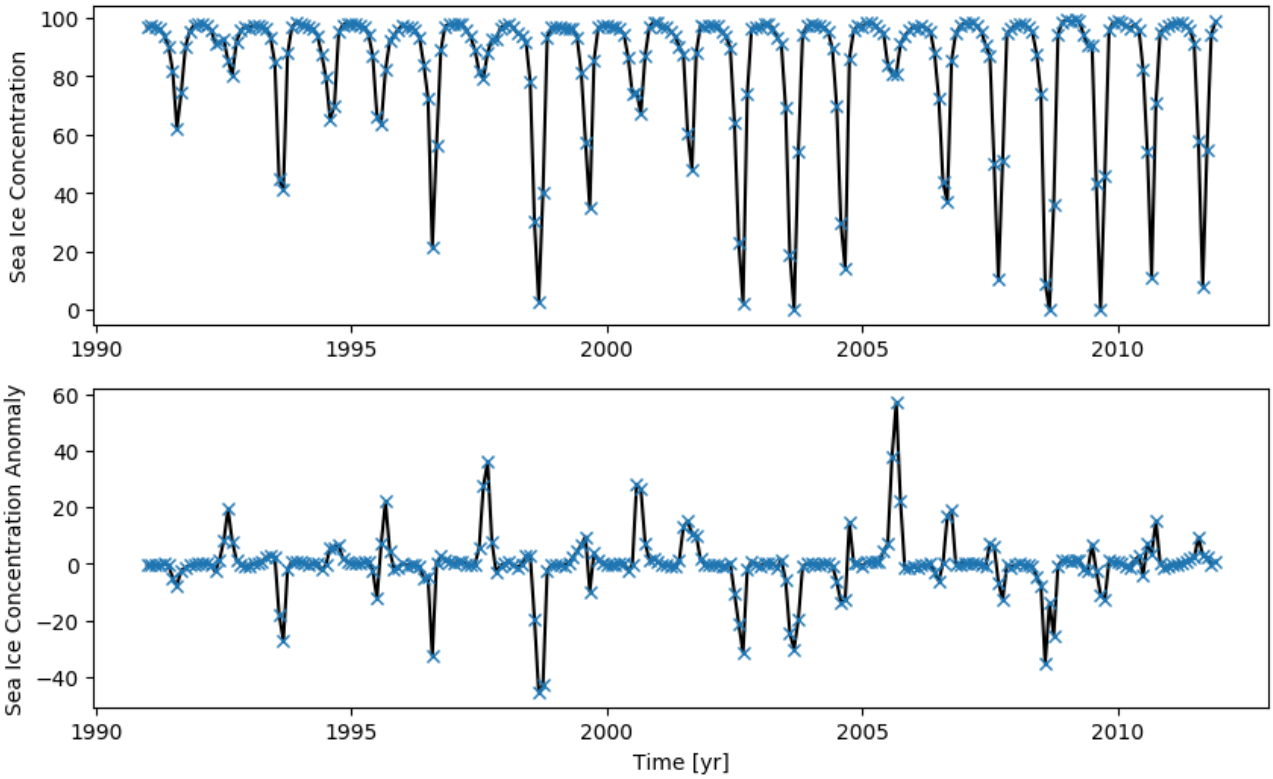

Fig A. (top) Sea Ice Concentration [%] from C-GLORS05 reanalysis at the location of the BGEP oceanographic mooring. (bottom) Same as top panel, but for the Sea Ice Concentration anomaly [%].

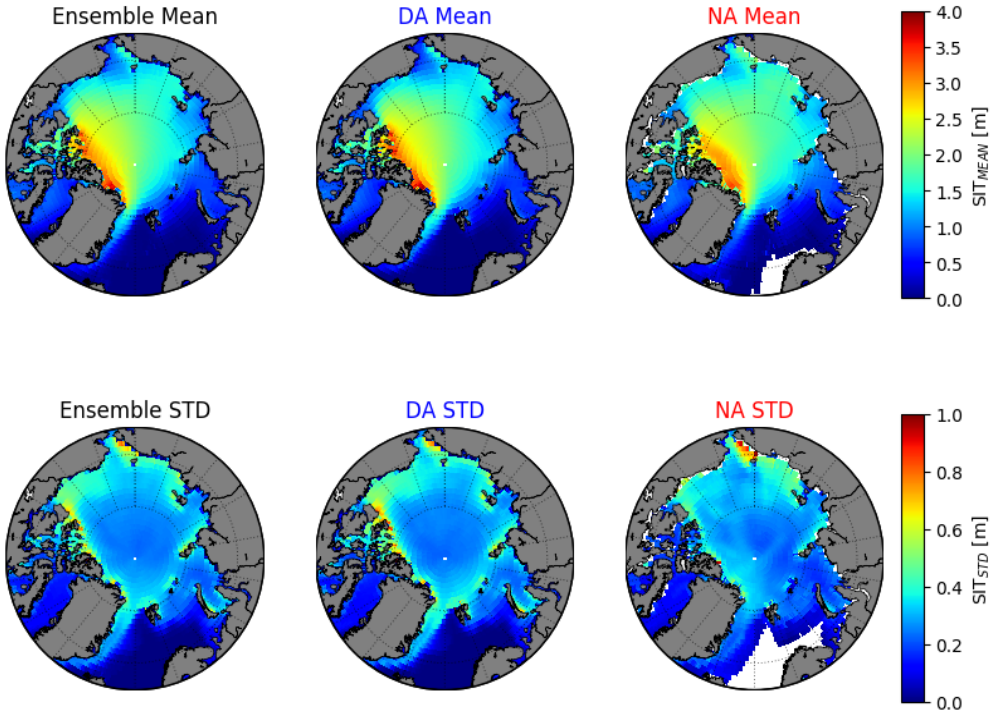

Fig B. (top) Sea ice thickness mean estimated for the (a) entire ensemble, (b) reanalyses with sea ice data assimilation and (c) reanalyses without sea ice data assimilation. (bottom) Same as (top), but for the standard deviation.

3) Figure 2 clearly shows the time scale has been extended from less than 3 months to around 4 months, which is convert with the main finding (P1, Line 9: … data assimilation present shorter time and length scales).

We absolutely agree with referee's comment. In Fig. 2 the reanalyses which do not assimilate data seem to have shorter TS, in contrast with our main finding. Nevertheless, the results that support our main finding is based on the Averaged Weighted Mean (AWM) values, as a representation of all grid cells, while Fig. 2 represents a single grid point.

Fig. 2 should be interpreted with caution. Its main goal is to show that the time scales found for the observations are somehow within the range found for the reanalyses. That's why, Fig. 2 is displayed in Sec. 2.3 ('Methods').

However, we indeed agree that this is an important point and deserves a better clarification in the text (Fig. 11; pg. 20, lines 3-4)

4) The previously compared studies show atmospheric forcing fields essentially drive the results of sea ice simulations (Gerdes and Köberle, 2007; Hunke and Holland, 2007). Can you add some comment or analysis about the impact of the forcing resolutions on the SIT time and length scales?

Gerdes, R., and C. Köberle: Comparison of Arctic sea ice thickness variability in IPCC Climate of the 20th century experiments and in ocean–sea ice hindcasts, J. Geophys. Res., 112, C04S13, doi:10.1029/2006JC003616, 2007.

Hunke, E., and M. Holland: Global atmospheric forcing data for Arctic ice-ocean mod-

eling, J. Geophys. Res., 112, C04S14, doi:10.1029/2006JC003640, 2007.

Even though the atmospheric forcing fields are reported to play a major role in the sea ice simulations, as pointed out by the referee, we could not identify distinguished patterns between the two main sources of atmospheric forcing used by the ensemble of reanalyses: Era-Interim and NCEP/NCAR (Figs. 12c-d). (pg. 20, lines 15-18)

5) The ice draft measurements from submarine have been identified an overall overestimation of +0.29 m (Rothrock and Wensnahan (2007). This dataset also is used in this study. Can you add some comments about these kinds of bias corrections (also to other related observational data) applied here or not.

The Sea Ice CDR already provides the corrected data. For the US submarine case, notice that the files were produced and made available by Dr. Mark Wensnahan http://psc.apl.uw.edu/sea_ice_cdr/Sources/US%20Submarines.html, one of the researchers who identified and reported (Rothrock and Wensnahan, 2007) the biases (pg. 5, line 7).

Technical issues:

1) As a basic index calculated by the SIT, it still is not clear how to deal with the conflicts of seawater and ice cover at each grid. For example in Fig. 2, when the observed sea ice is lager 0.1m, but the reanalyses are not all covered by sea ice. It is also not clear at P 15 Line 13 when to calculate the correlation between the two points: how to ensure the same lengths of the SIT time series.

In this work the reanalyses provide the mean sea-ice thickness of each grid cell including open water for the reanalyses (i.e. sivol variable in CMIP6). As adopted by Blanchard-Wrigglesworth and Bitz (2014), only grid points wherein the mean ice thickness at the time of summer minimum is greater than 0.1 m are taken into account.

The draft observations used in Fig. 2 represent the location where the oceanographic mooring (BGEP) was deployed (Krishfield et al., 2013), while the reanalyses' data in the same figure come from the nearest grid point to the BGEP oceanographic mooring.

On pg. 15, line 13, we are discussing about the stability of the length scale over time. In order to perform these analyses, we have used only the two long-term reanalyses (GECCO2 and MOVE-CORE), taking into account an overlapping period of 64 years, from 1948 to 2011 so that both time series have the same length.

2) P3, Line 20: "The original horizontal grids range from 0.25 to 1". It is not correct because the reanalysis of TP4 is regional product with the resolution of 12-16km (also see Xie et al. (2017)).

Xie, J., Bertino, L., Counillon, F., Lisæter, K., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991-2013, Ocean Sci., 13(1), 123-144, doi:10.5194/os-13-123-2017, 2017.

This information has been corrected in the text. It is worthwhile mentioning that the grid resolution of TOPAZ4 and of all the other reanalyses are now displayed in Table 1.

3) P3, Line 27:" … in details by Chevallier et al. (2017) (their Table 1) and Balmaseda et al. (2015)". It is not correct because they did not include the TP4 product as least so recommend of the reference: Xie et al. (2017) or Uotila et al. (2018).

Corrected.

4) P4, Line 27:" the linear relationship between both parameters given by the hydrostatic equation". It is better if clear to state the used equation or give a reference used.

This statement was tempered in the text. The main point here is the fact that we decided not to use the snow cover from the reanalyses for converting sea ice thickness to sea ice draft. This decision was taken in order to avoid adding uncertainties to the SIT fields. So, when comparing SIT from reanalyses against draft from observations, we estimate the linear relationship between both datasets by means of the coefficient of correlation (R) and the Mean Residual Sum of Squares (MRSS). (pg. 5, lines 25-32)

5) Figure 1 adds the grid lines or labels the year on each panel. It is more convenient to march the statement. So P 11, Line 13 "for instance from 2001 to 2004" looks not suitable, and can be replaced by "for instance from 2002 to 2004".

We have added yearly grid lines to Fig. 1. In the new version of the manuscript we are still referring to the period "2001 to 2004", but indeed the grid lines made the comparison between the panels much easier.

6) P 15 Line 10: "This suggests that higher length scales are associated with thicker ice" looks not so precis. It more likes around the North pole.

The relationship between the mean state and the LS is explored further in the new manuscripts' version (Fig. 13a,b; pgs. 21, lines 1-3). But indeed, the previous statement was more related to the region surrounding the North pole and this is what we tried to say with "… near central Arctic." (pg. 15, line 9 – in the first manuscript).

7) P 18 Line3:" .. scales of the sea ice thickness" replaced by "… scales of the sea ice thickness anomaly."

"anomaly" added to the text.

8) Figure 11 adds a panel to show the difference so that details would be more clear.

Fig. 11 shows the difference G2V3-G2V1, both for time (Fig. 11a) and length (Fig. 11b) scales.

Anonymous Referee #2

The paper provides a summary of the sea ice thickness anomalies found in the most recent ocean-ice analyses. It considers the impact of the assimilation of sea ice observations within some of these systems to spatial and temporal scales of the anomalies.

In my opinion the paper has lots of potential and is almost there, but at the moment it is missing some extra synthesis/analysis which would make it a really useful reference for the observation and modelling community. The impact of sea ice assimilation was considered (but did not state if any of the models assimilate anything other the concentration, one may nudge the thickness too?); it would be helpful to understand how the impact of other choices in the system may also influence the sea ice thickness anomalies. The sea ice models may have very different methods for modelling thickness. Do they have ice thickness distributions or a single category, does this have an impact?
Different atmospheric forcing sets (with different imposed sea ice cover) may influence the local energy balance and in turn affect the ice thickness. I think considering a few other key elements of these systems would make this paper very helpful in guiding the use of reanalyses and their future development (both in terms of the type of data that is used but also the set up of the assimilation systems)

We thank the referee for evaluating our manuscript with such richness of details. We appreciated very much her/his comments, which were all taken into account in the revised version of the manuscript. We have tried to identify the impact of several other reanalyses' choices (10 in total) on the time and length scales of the sea ice thickness anomalies. We did identified other parameters that potentially affect the referred scales, as for instance the air-ice drag coefficient. For the atmospheric forcing the results are not conclusive as detailed in our answer for your second specific comment. None of the reanalyses assimilate sea ice thickness.

It is worthwhile mentioning that in order to better incorporate the referees' suggestions, the manuscript's structure has changed. New figures were added, while others were replaced. Thus, all tables, figures, pages and lines referred to in this rebuttal letter are directed to the updated version of the manuscript, unless otherwise stated.

Specific comments:

The paper provides some details of the reanalyses you have studies but I think it is missing a table summarising the reanlayses - some form of synthesis would be beneficial to the reader, the papers that do tabulate some of this (e.g. Chevallier et al) are not complete for this reanalysis set. A table with clear information about the forcing data set and types of data that are assimilated and what methods are used to do this. Given that we also know that the strength parameter impacts the thickness it would be good to tabulate P* or equivalent as well. As the paper addresses the timescales of anomalies it would also help to determine the assimilation window length and see if this has an impact.

We have now added Table 1 (pg. 4) to the manuscript. The table reproduces the info already presented by Chevalier et al. (2017), but it also brings info for the other four reanalyses not considered by these authors (GECCO2, GloSea5, PIOMAS and TOPAZ4). The following parameters and specifications are included in Table 1: nominal horizontal resolution, ocean-sea ice model, source of atmospheric forcing data, number of ice-thickness categories, EVP or VP dynamics, ice strength parameter (P*) or frictional dissipation coefficient (Cf), air-ice drag coefficient (C_W), ocean-ice drag coefficient (C_A), source of sea ice data assimilated and method used for assimilating sea ice data.

Since the atmospheric forcing and many other parameters/specifications (see further in this rebuttal letter) do not seem to play a clear role in the time and length scales, and mainly due to the fact that we are working with monthly time series, we don't expect to see large changes in the fields of time and length scales associated to the "assimilation window length". At least, not for this kind of comparison where each product has several different parameter/specification compared to the other systems.

Other parameters could be tested, nevertheless it is not always easy and straightforward to retrieve all information from all reanalyses, as well as to compare products with several varying parameters. Nevertheless, we believe that the list of parameters in Table 1 is comprehensive and **reinforce that time and length scales are mainly driven by the fact of whether or not the reanalyses assimilate sea ice concentration data** (see below).

One reason for requesting the table is when looking at your first figure. You present results but there is not much discussion about the differences that are present in the time series. Are you able to stratify the impact of certain assimilation choices other than whether sea ice data is assimilated.

I would suggest that you may see an impact of the different forcing sets that are used, many are forced with ERA-Interim but MOVE(CORE and G2), GECCO, EDCA, MERRA are not. The ice fields that the atmospheric forcing fields have "seen" will have a impact on the forcing they provide. Differences due to SST relaxation or model parameter choices may also play a role – it would be good to at least see if there are other reasons for the differences other than they include sea ice assimilation or not.

Figs. A and B (attached to this rebuttal letter) show how time and length scales (hereafter, TS and LS) are related to the reanalyses' parameters and specifications displayed in Table 1.

Fig. A shows a comparison of TS and LS against a set of 'uncountable' specifications (sea ice data assimilation, atmospheric forcing, sea ice model, dynamics, EVP or VP dynamics, and sea ice forcing). The results reinforce that TS and LS are indeed linked to the fact of whether the reanalyses assimilate (DA) or not (NA) sea ice data (Fig. A*a*,*b*,*i*,*j*), while the other specifications do not seem to have a large impact on the studied scales (Fig. A*c*–*j*). For the atmospheric forcing case, notice that we can divide the reanalyses into three groups: ERA-Interim (7 out of 14), NCEP/NCAR (3 out of 14) and Others (4 out of 14). So, an effective comparison can be made only between ERA-Interim and NCEP/NCAR. Nevertheless, from NCEP/NCAR 1 system assimilates sea ice data, while 2 other systems do not assimilate sea ice data. This makes it difficult to evaluate the impact of the atmospheric forcing on the studied scales.

Fig. B*g*-*p* compares TS and LS against the 'countable' reanalyses parameters. A part of the horizontal grid resolution and the air-ice drag coefficient (C_A), which show a certain correlation with LS (Fig. B*h*) and TS (Fig. B*m*), respectively, the number of thickness categories, the ice strength parameter (P*), and the ocean-ice drag coefficient (C_W) do not show a strong correlation with the studied scales.

In addition, Fig. B*a*–*f* also compares TS and LS against the mean state (mean Sea Ice Volume; SIV), the interannual of variability (std SIV anomaly), and the Sea Ice Drift (SID). All these parameters show a certain correspondence with the TS and/or LS. The most pronounced is the case Mean Sea Ice Drift *x* TS shown in panel Fig. B*e*.

The most meaningful results from Figs. A and B are incorporated in the new version of the manuscript (Figs. 12 and 13; Sec. 4)

Section 2.3:

It was not entirely clear to me how you were treating draft and SIT differently from the reanalyses when comparing to observations. Did you use the snow cover from reanalyses to compute this when comparing to observations or just disregard for both reanalyses and observations?

We decided not to use the snow cover from the reanalyses for converting sea ice thickness to sea ice draft. This decision was taken in order to avoid adding uncertainties to the SIT fields. So, when comparing SIT from the reanalyses against Draft from the observations we made use of **two metrics**:

(i) **the correlation coefficient (R)**, as a measure of the **linear correlation** between SIT (reanalyses) and Draft (observations);

(ii) **the Mean Residual Sum of Squares (MRSS)**, as an indicator of whether SIT values from the reanalyses are good **predictors** for Draft observations;

We have improved this info in the text in order to make it clearer to the reader (pg. 5, lines 25-33).

Section 3:

Figure 4: some of the scatter diagrams look like the model thickness stops at a particular value e.g. ECDA are you missing some data from thickness categories? Where the draft data do not show a similar relationship to the SIT is there something different about the way snow on ice is treated in the systems? Is this dependent on the assumptions you made about how you compare draft and SIT?

The data used in this work was previously compiled and published by Chevalier et al. (2017) and Uotila et al. (2018). These authors collected the data as they were made available by the providers. So, it is really unlike that some thickness categories are missing.

In ECDA this feature (Fig. 4) takes place due to the fact that this system is characterized by relatively thin ice. See Fig. C at the end of this rebuttal letter. Notice that the same is also valid for GloSea5 and UR025-4.

Do you have an understanding of why the largest differences are near the Greenland coast and Canadian Archipelago? is this down to model physics differences?

The regions near Greenland coast and Canadian Archipelago are marked by the thickest sea ice over the studied domain (please, see Fig. B from the rebuttal letter to the first reviewer). So, proportionally, the same differences in these regions are amplified when calculating the errors. This is a simple, but interesting, effect that deserves some explanation in the text. Thanks for your observation. (pg. 19, lines 8-10)

Section 3.3:

You note that the GloSea systems have shorter timescales than others - is this persistence also linked to mean state? If you have thinner ice on average you may lose it over the summer and it will reduce your persistence.

Overall, the systems which do assimilate sea ice are marked by a certain correspondence between TS (persistence) and the mean state as shown in Fig. 13a (or Fig. Ba in this rebuttal letter). The same for the LS and mean state (Fig. 13b or Fig. B*b*). Nevertheless, this is an overall observation

for the cloud of points. Looking carefully at the GloSea systems, GloSea5 is indeed characterized by thin ice (second thinner mean state of the ensemble). On the other hand, GloSea5-G05 has thicker ice compared to GloSea5. Compared to the ensemble of the reanalyses which do assimilate sea ice, the GloSea5-G05 presents an intermediate mean state. In Sec. 4 (pg. 22, lines 11-20), we raise some aspects that could potentially impact the scales of the GloSea systems.

pg 12, lines 5-6: these sentences were not enitrely clear - would suggest you consider rephrasing to make sure the meaning is clear. do you mean that G2V3, ORAP5, PIOMAS, TP4 and UR025-4 are just similar and if so why hightlight this group - is it for a particular region of the Arctic. It wasn't totally clear what you wanted to point out here.

We agree that this sentence was confusing, and we have reconsidered it. (pg. 14, lines 5-6)

pg 12, lines 17-18: do you mean years or months?

We meant months. Corrected in the text.

Does your wavelet analysis give an indication how robust your findings might be based on the limited time series? If so I think this would be worth commenting on in the text.

The "cone of influence" shown in Fig. 7c,g (cross-hatched areas) highlights the region of the wavelet spectrum where edge effects become important due to the length of the time series. In this region of the spectrum, the results should be interpreted carefully, since the time series is short for solving some periods (y-axis) in some specific time spans (x-axis). This information is found in the Fig. 7's caption.

Section 4:

pg 17, lines 3-6: You are comparing c_GLORS05-G2V3 - do both of these systems assimilate sea ice data in the same way? You also note that G2V3 and ORAP5 could be considered similar – they may be similiar in some respects but potentially have different forcing G2V3 its not clear to me if G2V3 uses operational analysis from ECMWF rather than ERA-Interim - which may lead to some differences. ORAP5 uses a non-standard value of P* compared to standard LIM2 setup.

Indeed, C-GLORS05 and G2V3 assimilate sea ice data in a different way (Table 1), and we agree with referee's comment that this may be a reason for the relatively high RMSE between these two systems. The three referred reanalyses (C-GLORS05, G2V3 and ORAP5) apply the same atmospheric forcing (ERA-Interim). C-GLORS05 and G2V3 assume $P^* = 2.0 \times 10^4$, while ORAP5 uses $P^* = 1.5 \times 10^4$.

Based on referee's observations, we brought this discussion to the text (pg. 18, lines 5-12).

Figure 11 shows the reduction in time scale in one system with the application of sea ice assimilation but the change seems somewhat smaller than the differences with the no ice assimilation across the ensemble - would you expect this given you results? where would G2V1 be in your fig 13?

Following referee's 1 suggestion, the new version of Fig. 11 shows the TS and LS differences between the systems (G2V3 – G2V1). Fig. 11 suggests that the fact that time and length scales are shorter for systems with data assimilation is valid in terms of pan-Arctic averages, but not necessarily at every grid cell. The pan-Arctic AWM-TS and AWM-LS from G2V3 are 5 months

and 728.2 km, respectively. Without sea ice data assimilation (G2V1), the AWM-TS and AWM-LS increase to 5.5 months and 745.3 km, respectively. For this case, TS increases by 10% from G2V3 to G2V1, though the LS scale isn't strongly impacted (pg. 20, lines 1-6). In Sec. 4, we are discussing that TS is more sensitive to the systems' parameters and specifications.

Fig. 14 (empty red circle; highlighted in the gray rectangle) displays where G2V1 takes place in the diagram TS *x* LS.

Figure 13 and discussion pg19, lines 15-17. I was not convinced that the ice volume anomaly correlated well with the time and length scales - without the GloSea systems it seems less clear.

We agree that in the way the figure was plotted, it wasn't easy to identify the correlation between mean state and the scales (TS and LS). So, we have now plotted two different diagrams: Mean SIV *x* TS (Fig. 13a) and Mean SIV *x* LS (Fig. 13b). Besides the spread among the points and the relatively weak correlations (which are not significant at the 5% level), an increase in the mean SIV generally leads to longer time and length scales, taking into account that each reanalysis has different sets of specifications and parameters (pg. 21-22).

Technical comments:

Abstract: line 4: intend rather than intent

Corrected.

section 2.2:

pg4, line3: not sure what you mean by ponctual - is this "point" measurements?

Corrected.

pg4, line 14: run by Environment.…

Corrected.

section 3.1: pg8, line 2: Don't you mean MERRA rather than MOVE?

Indeed, we meant MERRA. Corrected in the new version of the manuscript.

section 3.2: pg8, line 22 and pg11, line8: do you mean "error" or "difference" ?

We meant error (RMSE). Clarified in the text.

Section 3.4:

the last sentence of the section "..for what a quantification is presented in the conclusions of this study" - the phrasing is somewhat awkward and hard to understand exactly what you mean

Corrected.

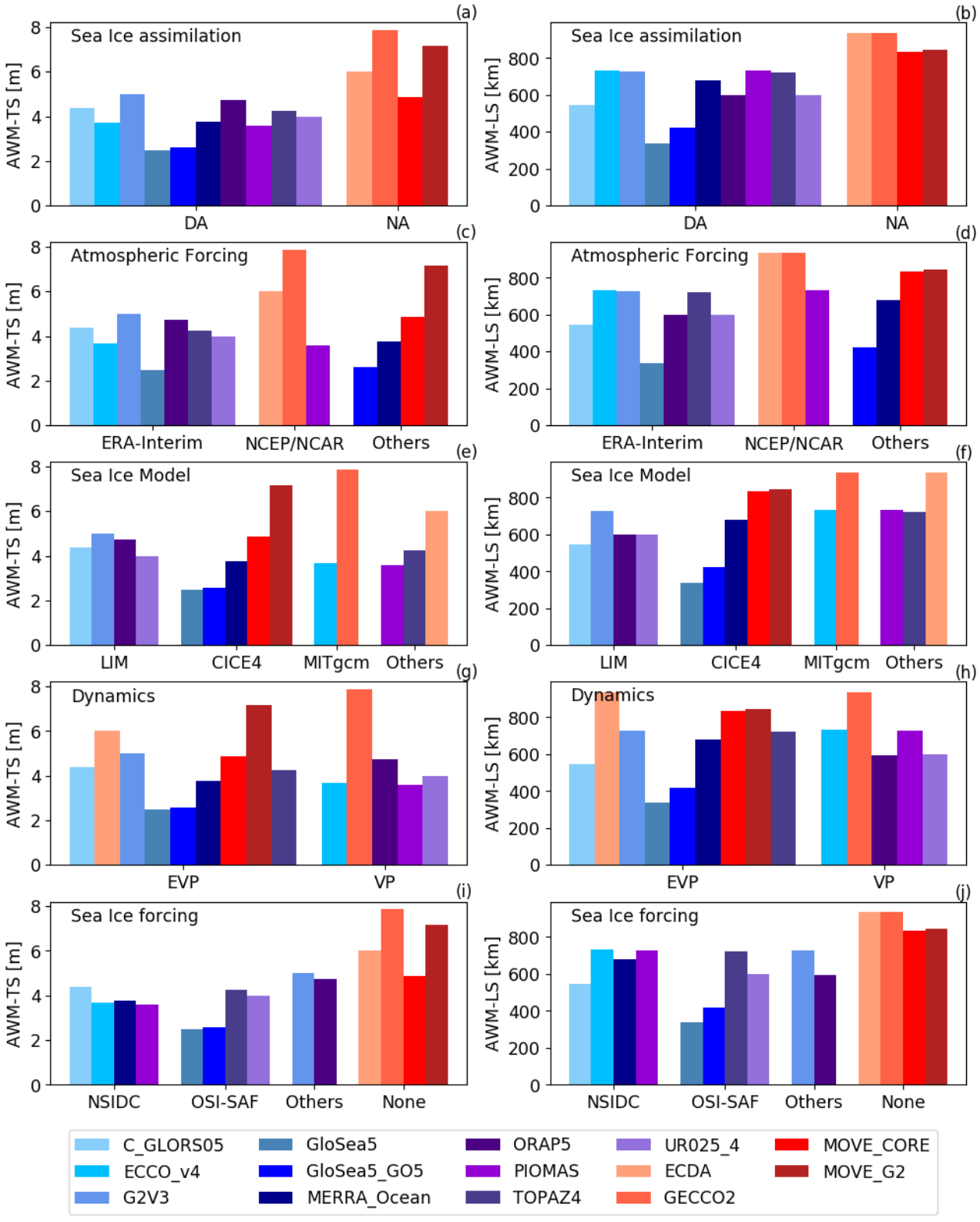

Fig. A: Histograms showing how the AWM-TS (left panels) and AWM-LS (right panels) are related to other "uncountable" reanalyses' parameters and/or specifications, as follows: (a-b) whether or not the system assimilates sea ice data; (c-d) the source of atmospheric forcing data; (e-f) the used sea ice model; (g-h) the used dynamics (Viscous-Plastic or Elastic-Viscous-Plasctic) to account for ice-ice interactions that control ice deformation'; and (i-j) the source of assimilated sea ice concentration data.

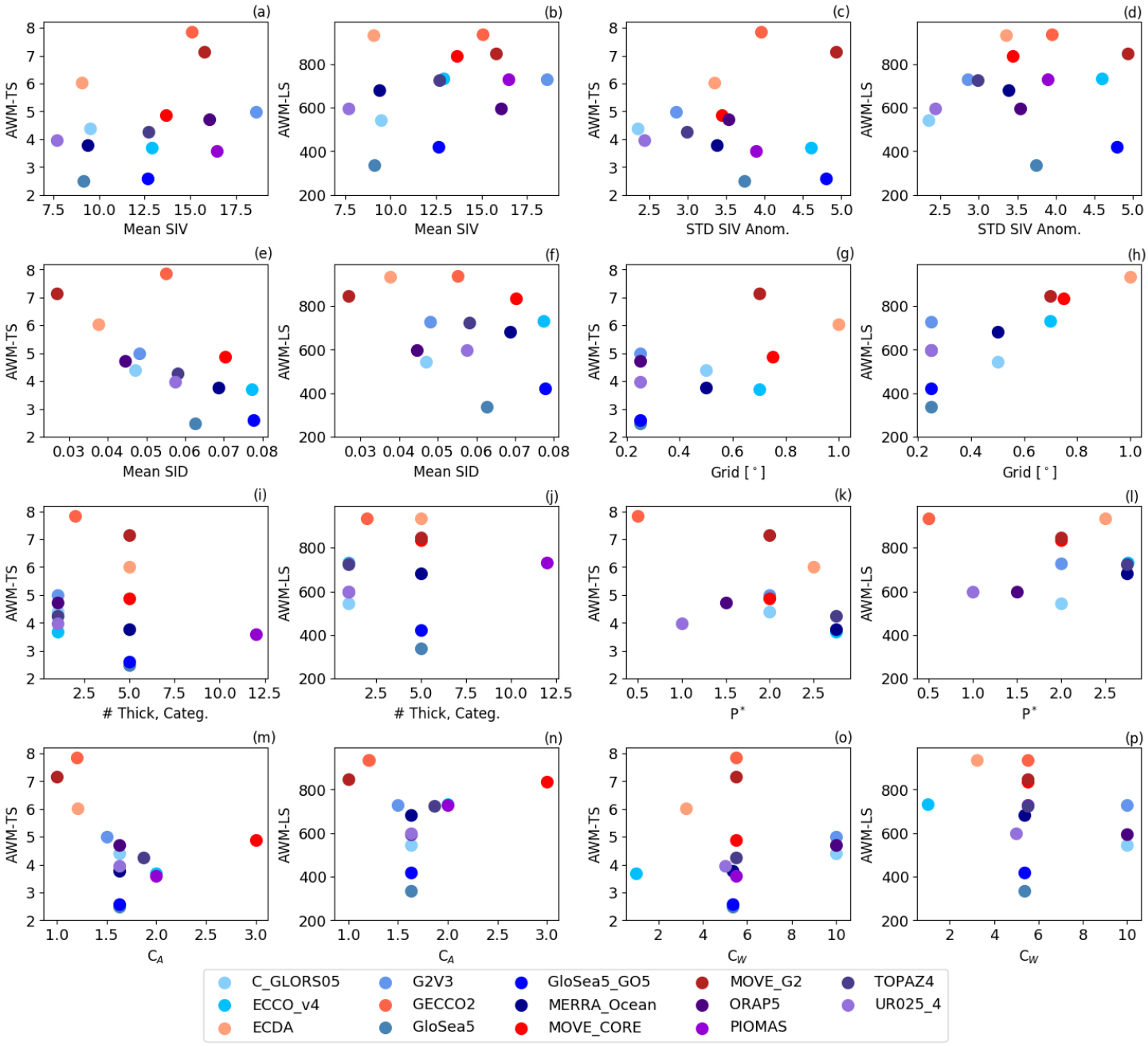

Fig. B: Scatter plots showing how the AWM-TS [months] (first and third columns) and AWM-LS [km] (second and fourth columns) are related to other reanalyses' parameters, as follows: (a-b) Mean Sea Ice Volume [m] (SIV); (c-d) standard-deviation of the SIV Anomaly [m]; (e-f) Mean Sea Ice Drift [m/s] (SID); (g-h) grid resolution [degrees]; (i-j) number of thickness categories for discretization of ice thickness; (k-l) ice strength parameter (P*) [N/m] for the ice strength formulation following Hibler (1979); (m-n) air-ice drag coefficient [10^{-3}] (C_A); and (o-p) ocean-ice drag coefficient [10^{-3}] (C_W).

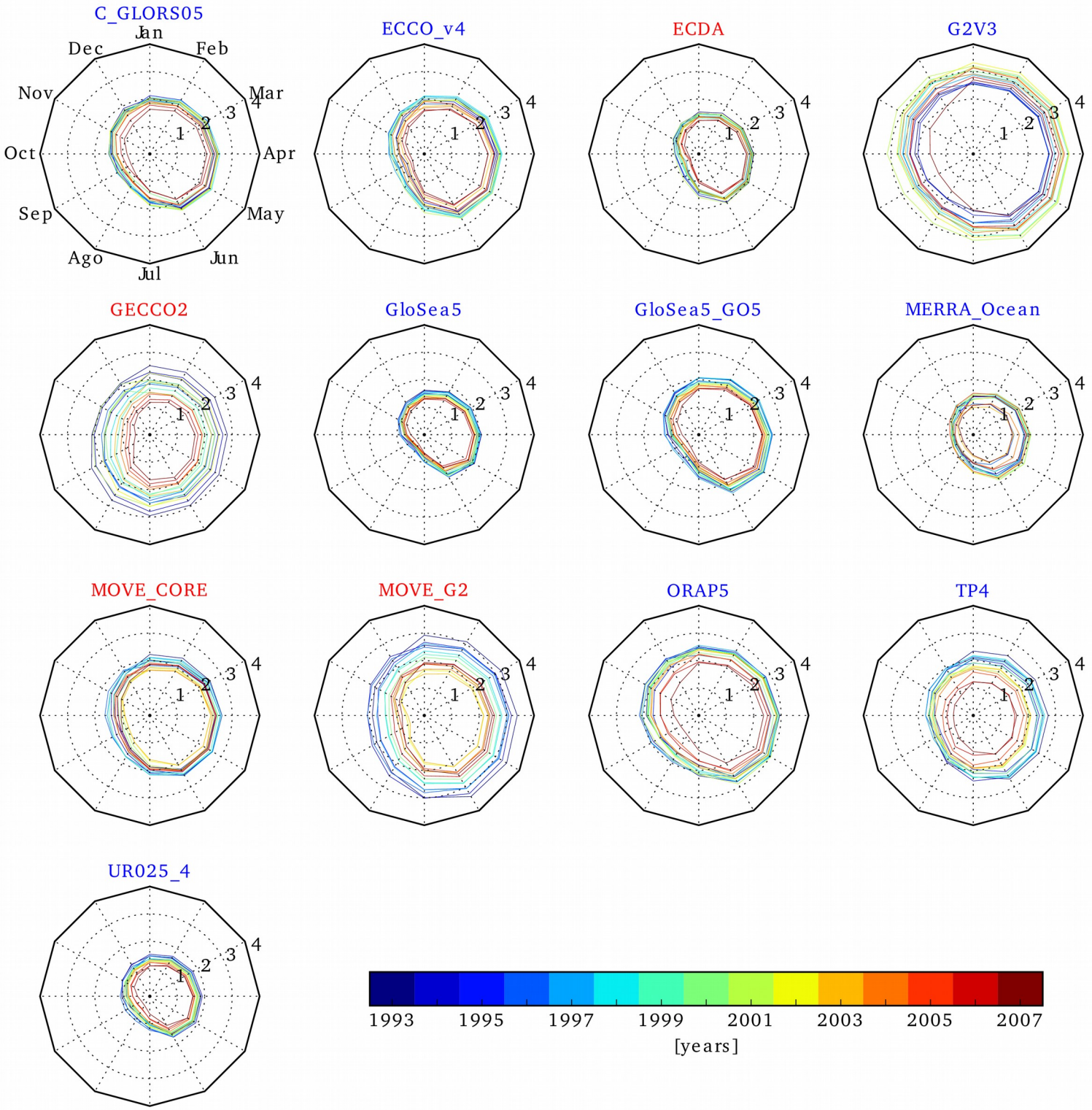

Fig. C: Radar plots for the Averaged-Weighted-Mean Sea Ice Thickness. Every color represents a different year from 1993 to 2007.

On the time and length scales of the Arctic sea ice thickness anomalies: a study based on fourteen reanalyses

Leandro Ponsoni[1], François Massonnet[1], Thierry Fichefet[1], Matthieu Chevallier [2], and David Docquier[1]

[1]Georges Lemaître Centre for Earth and Climate Research (TECLIM), Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium

[2]Centre National de Recherches Météorologiques (CNRM), Météo France/CNRS UMR3589, Toulouse, France

Correspondence to: Leandro Ponsoni (leandro.ponsoni@uclouvain.be)

Abstract. The ocean–sea ice reanalyses are one of the main sources of Arctic sea ice thickness data both in terms of spatial and temporal resolution, since observations are still sparse in time and space. In this work, we first aim at comparing how the sea ice thickness from an ensemble of fourteen reanalyses compares with different sources of observations, such as moored upward-looking sonars, submarines, airbornes, satellites and ice boreholes. Second, based on the same reanalyses, we ~~intent~~ intend to characterize the time (persistence) and length scales of sea ice thickness anomalies. We investigate whether data assimilation of sea ice concentration by the reanalyses impacts the realism of sea ice thickness as well as its respective time and length scales. The results suggest that reanalyses with sea ice data assimilation do not necessarily perform better in terms of sea ice thickness compared with the reanalyses which do not assimilate sea ice concentration. However, data assimilation has a clear impact on the time and length scales: reanalyses built with sea ice data assimilation present shorter time and length scales. The mean time and length scales for reanalyses with data assimilation vary from 2.5–5.0 months and 337.0–732.5 km, respectively, while reanalyses with no data assimilation are characterized by values from 4.9–7.8 months and 846.7–935.7 km, respectively.

1 Introduction

The variability of the Arctic sea ice has received increasing attention from the scientific community over the past years (e.g., Chevallier and Salas-Mélia, 2012; Stroeve et al., 2014; Blanchard-Wrigglesworth and Bitz, 2014; Guemas et al., 2016). The main reason lies in the fact that Arctic sea ice plays a key role in the Earth's climate system (Budyko, 1969; Manabe and Stouffer, 1980b; Maykut, 1982). Among other contributions, it has been suggested that a decline of the Arctic sea ice extent and volume leads to a weakening of the Atlantic Meridional Overturning Circulation (Sévellec et al., 2017) and, therefore, potentially impacts the global distribution of heat (Drijfhout, 2015; Hansen et al., 2016). At the same time, the Arctic is one of the most sensitive regions to climate changes due to a phenomenon known as Arctic amplification (Manabe and Stouffer, 1980a; Holland and Bitz, 2003; Serreze et al., 2009). For instance, the current observed warming in the Arctic is reported to be nearly twice as large as other regions of the globe (Anisimov et al., 2007).

Other multiple specific interests from different stakeholders have reinforced the importance of the sea ice projections, both at regional and larger scales, ~~that~~ which include: shorter shipping lanes (Lindstad et al., 2016), travel and tourism industry

(Handorf, 2011), hunting and fishing activities (Nuttall et al., 2005), mineral resource extraction (Gleick, 1989), potential impact on the weather at mid-latitudes (Walsh, 2014), environmental hazards (Nelson et al., 2002) and loss of weather predictive power by indigenous communities (Krupnik and Jolly, 2002). In this context, the sea ice thickness (SIT) is likely the most relevant state variable for monitoring, forecasting and understanding ~~both~~ recent and future changes ~~in terms~~ of the Arctic sea ice. First, because this parameter provides predictive information for the sea ice extent anomalies (Lindsay et al., 2008; Holland et al., 2011) and, second, due to the fact that SIT anomalies persist longer than sea ice extent anomalies, the former being reported as a forcing of the latter (Blanchard-Wrigglesworth et al., 2011).

However, direct observations of SIT and/or related parameters, namely draft and freeboard, are still sparse in time and space, besides the continuous efforts for compiling former and recent datasets from a range of sources (Lindsay, 2010; Lindsay and Schweiger, 2015). Some recent observational programmes, such as the Year Of Polar Prediction (YOPP) (Jung et al., 2016) and the MOSAiC International Arctic Drift Expedition (http://www.mosaicobservatory.org/), aim to enhance the Arctic observational system, being especially useful for improving our future modeling and forecasting skills.

Due to this lack of direct measurements in the past and present-day, the ocean–ice reanalyses deserve special attention. A reanalysis product consists of models' outputs, which are generated over a certain time span by the same model, configurations and procedures, and so distributed onto regular grids, evenly stepped in time. These products are often built with assimilation of observational dataset(s) in order to improve the estimate of a certain parameter. For instance, SIT is often estimated by assimilating atmospheric, oceanic and, eventually, sea ice concentration data. The ocean–ice reanalyses are likely the main and more robust source of SIT data in terms of spatio-temporal resolution, being also broadly used for initialization and assimilation in other climate models (e.g., Guemas et al., 2016). Additionally, long-term reanalyses are crucial for understanding the past Arctic sea ice characteristics, in a period when *in situ* observations of ice parameters were inexistent.

In this work we make use of fourteen state-of-the-art reanalyses in order to study two important aspects of the SIT predictability: the time scale (or persistence) and the length scale of SIT anomalies (see Sec. 2.3). Their importance is reinforced by the fact that the predictability of the SIT field depends on how long the anomalies persist over time and on ~~the spatial scale of the respective anomalies~~ how these anomalies spread in space. Notice that, hereafter, besides the traditional definition of time predictability, we adopt this term also for the spatial scale. In addition, time and length scales may also be useful for designing an optimal observation system, when selecting ideal locations for deploying instruments, as well as for defining a frequency sampling strategy (Blanchard-Wrigglesworth and Bitz, 2014).

Blanchard-Wrigglesworth and Bitz (2014) reported SIT anomalies with typical time and length scales of about 6–20 months and 500–1000 km, respectively. These results reinforce the fact that SIT anomaly persists longer compared to sea ice area anomaly, which is reported with a time scale of 2–5 months (Blanchard-Wrigglesworth et al., 2011; Day et al., 2014). Blanchard-Wrigglesworth and Bitz (2014) suggested that the SIT anomalies from models characterized by a thinner mean ice state tend to present shorter persistence, but larger spatial scales. Blanchard-Wrigglesworth et al. (2011) reported a decline in the time scale of sea ice volume anomalies, as a result of the ice thinning induced by the recent climate changes.

The first aim of this study is to evaluate the performance of different reanalysis products regarding their SIT realism by comparing these reanalyses against observational datasets. A point of main interest is to identify whether or not the assimilation

of sea ice concentration by the reanalyses improves the representation of SIT. Second, we seek to characterize the time and length scales of the Arctic SIT anomalies. Again, we verify whether or not sea ice data assimilation plays a role in the temporal and spatial scales of SIT anomalies. Furthermore, we investigate the long-term evolution of time and length scales, as well as the relationship between these two parameters.

The manuscript is organized as follows: Section 2 introduces the reanalysis products, the observational datasets and the respective methods applied in this research; Section 3 compiles all results, including the comparison between observations and reanalyses (Section 3.1), the comparison of reanalyses themselves (Section 3.2), and the patterns of time (Section 3.3) and length (Section 3.4) scales. Lastly, Section 4 draws discussion and conclusions on the findings reported in the previous sections.

2 Data and methods

2.1 Sea ice reanalyses

Monthly fields of SIT from fourteen state-of-the-art ocean–ice reanalyses are used in this work. They were all but one compiled in the context of the ORA-IP project (Balmaseda et al., 2015; Chevallier et al., 2017; Uotila et al., 2018). The ORA-IP reanalyses (and their respective author/provider institution) are: C-GLORS05 (CMCC, Storto et al. (2014)), ECCO-v4 (JPL/NASA, MIT, AER, Forget et al. (2015)), ECDA (GFDL/NOAA, Zhang et al. (2013); Chang et al. (2013)), G2V3 (Mercartor Océan, Ferry et al. (2010)), GECCO2 (University of Hamburg, Köhl (2015)), GloSea5 (UK Met Office, Blockley et al. (2014)), GloSea5-GO5 (UK Met Office, Megann et al. (2014)), MERRA-Ocean (GSFC/NASA/GMAO, Rienecker et al. (2011)), MOVE-CORE (MRI/JMA, Danabasoglu et al. (2014)), MOVE-G2 (MRI/JMA, Toyoda et al. (2013)), ORAP5 (ECMWF, Zuo et al. (2015); Tietsche et al. (2017)), TOPAZ4 (~~TP4,~~ ARC MFC, ~~Sakov et al. (2012)~~Sakov et al. (2012); Xie et al. (2017)) and UR025-4 (University of Reading, Valdivieso et al. (2014)). The fourteenth reanalysis is the Pan-Arctic Ice-Ocean Modeling and Assimilation System – PIOMAS (Zhang and Rothrock, 2003). For the acronyms, the reader is referred to Appendix A. The original horizontal grids range from ~~0.25°~~ 12 km to 1°. For comparison, all reanalyses are interpolated onto a common grid of $1° \times 1°$ spatial resolution ~~(Chevallier et al., 2017)~~following Chevallier et al. (2017).

Specific characteristics of each reanalysis regarding horizontal resolution, ocean–sea ice model, ~~spanning period, data output frequency,~~ atmospheric forcing data, ~~type of vertical discretization,~~ subgrid-scale ice thickness distribution, ice dynamics (*VP* viscous-plastic or *EVP* elastic-viscous-plastic), parameters for the ice strength formulation, air-ice drag coefficient, ocean-ice drag coefficient and the presence (and respective method) of ice data assimilation were ~~described in details by Chevallier et al. (2017) (their Table 1) and Balmaseda et al. (2015). For PIOMAS, a complete description is provided by Zhang and Rothrock (2003).~~ summarized in Table 1. For additional information the reader is referred to Chevallier et al. (2017) and/or to the respective providers.

Table 1. Ensemble of reanalyses, their respective configurations and set of selected parameters.

Reanalysis/ Parameter	C-GLORS05	ECCO-v4	ECDA	G2V3	GECCO2	GloSea5	GloSea5-GO5	MERRA-Ocean	MOVE-CORE	MOVE-G2	ORAP5	PIOMAS	TOPAZ4	UR025-4
Nominal horizontal resolution	0.5°	0.4°–1.0°	1.0°	0.25°	~40 km in the Arctic	0.25°	0.25°	0.5°	0.5°×1.0°	0.3–0.5°×1.0°	0.25°	~22 km in the Arctic, Barents, and GIN Seas	12–16 km	0.25°
Ocean–sea ice model	NEMO3.2-LIM2	MITgcm	GFDL-MOM4.4.1-SIS	NEMO3.1-LIM2	MITgcm	NEMO3.2-CICE4.0	NEMO3.4-CICE4.1	MOM4.1-CICE4.0	MRI.COM3-Mellor & Kantha (1989) + CICE4.0	MRI.COM3-Mellor & Kanta (1989) + CICE4.0	NEMO3.4-LIM2	POP-TED	HYCOM2.2-Drange & Simonsen (1996), Hunke & Dukowicz (1997)	NEMO3.2 LIM2
Atmospheric forcing data	ERA-Interim	ERA-Interim	Coupled run constrained to NCEP/NCAR-NCEP/DOE	ERA-Interim	NCEP/ NCAR	ERA-Interim	CORE2	Coupled run constrained to MERRA	CORE	JRA55	ERA-Interim	NCEP/ NCAR	ERA-Interim	ERA-Interim
Number of Ice-thickness categories	1	1	5	1	2	5	5	5	5	5	1	12	1	1
Dynamics	EVP	VP	EVP	EVP	VP	EVP	EVP	EVP	EVP	EVP	VP	VP	EVP	VP
P^* (10^4; N/m^{-2}) or Cf (–)	$P^* = 2.0$	$P^* = 2.754$	$P^* = 2.5$	$P^* = 2.0$	$P^* = 0.5$	Cf = 17	Cf = 17	$P^* = 2.75$	$P^* = 2.0$	$P^* = 2.0$	$P^* = 1.5$	Cf = 17	$P^* = 2.75$	$P^* = 1.0$
Drag air–ice (10^{-3})	1.63	2.00	1.21	1.50	1.2	1.63	1.63	1.63	3.00	1.00	1.63	2.00	1.6–2.14	1.63
Drag ocean–ice (10^{-3})	10.00	1.00	3.24	10.00	5.5	5.36	5.36	5.36	5.50	5.50	10.00	5.50	5.50	5.00
DA sea ice system	Linear nudging	Adjoint	None	2D local analysis SEEK filter	None	3D-VAR	3D-VAR	EnOI	None	None	3D-VAR-FGAT	Nonlinear nudging	EnKF	OI
DA sea ice data	NSIDC	NSIDC	None	CERSAT	None	OSI-SAF	OSI-SAF	NSIDC	None	None	OSTIA	NSIDC	OSI-SAF	OSI-SAF

2.2 Observational references

We use a compilation of sixteen observational datasets available in the Unified Sea Ice Thickness Climate Data Record (Sea Ice CDR; Lindsay (2010) – http://psc.apl.uw.edu/sea_ice_cdr). The Sea Ice CDR is a concerted effort to bring together a range of datasets in a consistent format, but originally sampled by different methods and spatio-temporal scales, as well as stored in a variety of formats. We use the post-processed version of the Sea Ice CDR data, that is distributed in monthly mean for moored upward-looking sonar (ULS) ~~ponctual~~ point measurements or 50-km averages for submarine, airborne and satellite observations. If applicable, the Sea Ice CDR already provides the files corrected from data biases (e.g. Rothrock and Wensnahan, 2007b).

From these sixteen datasets, eleven provide draft measurements, while the remaining five provide sea ice thickness data. Seven draft datasets were sampled by means of moored ULSs, namely “North Pole Environmental Observatory" (NPEO; Drucker et al. (2003); Rothrock and Wensnahan (2007b)), “Beaufort Gyre Exploration Project" (BGEP), “Institute of Ocean Science (IOS) - Eastern Beaufort Sea" (IOS-EBS) and “- Chuck Sea (IOS-CHK)", “Alfred Wegener Institute - Greenland Sea" (AWI-GS; Harms et al. (2001)), “Bedford Institute of Oceanography Lancaster Sound" (BIO-LS; Pettipas et al. (2008); Prinsenberg and Pettipas (2008); Prinsenberg et al. (2009)), and “Polar Science Center - Davis Strait" (Davis_St; Drucker et al. (2003)). Four other draft datasets are also based on ULS measurements, but installed on US and UK submarines: “US Navy Submarines - Analog" (US-Subs-AN), “US Navy Submarines - Digital" (US-Subs-DG; Tucker III et al. (2001); Wensnahan and Rothrock (2005); Rothrock and Wensnahan (2007b)), “UK Navy Submarines - Analog" (UK-Subs-AN), and “UK Navy Submarines - Digital" (UK-Subs-DG; Wadhams and Horne (1980); Wadhams (1984)).

From the ensemble of sea ice thickness datasets, the “Ice Thickness Program" ~~ran~~ run by Environmental Canada (CanCoast) is the only dataset providing direct measurement of ice thickness by means of ice boreholes. The “NASA Operation IceBridge datasets" (IceBridge-V2 and IceBridge-QL; Kurtz et al. (2013)) are derived from aircraft-mounted laser altimeter. Finally, two datasets come from satellite campaigns: the laser-altimeter ~~derive~~ derived “ICESat Mission-Goddard" (ICESat1-G; Zwally et al. (2008)) from National Aeronautics and Space Administration (NASA) and the radar altimeter-derived “CryoSat satellite data" (CryoSat-AWI; Ricker et al. (2014)) from the European Space Agency (ESA).

2.3 Methods

Reanalyses are compared against observations by selecting SIT values from the nearest grid points to the respective observational sites, during the same respective months. Complementary metrics are employed to evaluate the relationship between ~~both datasets~~observations and reanalyses. When directly comparing SIT from reanalyses and observations, we estimate the Root Mean Squared Error (RMSE), the correlation coefficient (R) and the Mean Residual Sum of Squares (MRSS) from the linear fit between both datasets, by having the reanalysis values as predictors and the observational values as predicted variables. ~~When~~ Since SIT and draft are different variables, we here evaluate the strength of the linear relationship between them. Thus, when comparing SIT from reanalyses ~~with~~ against draft from observations, we only estimate R and MRSS. ~~Though SIT~~

~~and draft are different variables, we consider the linear relationship between both parameters given by the hydrostatic equation, but~~ In this work we do not account for snow variation in order to avoid adding uncertainties ~~. RMSE and R are also used for a comparison among all pairs of reanalyses.~~

to the SIT fields.

SIT anomalies are derived by eliminating the trend and the seasonal cycle present in the time series. To do so, the trend is estimated separately for every month (Jan, Feb, ..., Dec) by means of a 2nd-order polynomial fit and subtracted from the respective month. A 2nd-order fit seems to better reproduce the trends when compared to the linear fit, although the results are very similar (not shown). The same method is applied for the analyses conducted with the pan-Arctic sea ice volume anomaly, derived from the SIT data, as illustrated in Fig. 1.

Grid point comparisons of SIT anomalies among all reanalyses are ~~quantified by RMSEs~~ performed by means of RMSE and R maps, calculated over an overlapping period of 15 years, from Jan/1993 to Dec/2007. This time span corresponds to the period during which data is available from all reanalyses. Furthermore, as adopted by Blanchard-Wrigglesworth and Bitz (2014), only grid points wherein the mean ice thickness at the time of summer minimum is greater than 0.1 m are taken into account. This condition is valid for all reanalysis-based results, unless otherwise stated.

The time scale (or persistence) is derived from individual time series by calculating the lagged autocorrelation stepped forward by one measurement, equivalent to one month. The e-folding reference is used so that the persistence is assumed to be the time when the lagged autocorrelation curve crosses the $1/e$ (~ 0.3679) value, as proposed in previous works (e.g., Blanchard-Wrigglesworth and Bitz, 2014; Guemas et al., 2016). As an example, Fig. 2 displays the time scale derived from the mooring-based draft anomaly sampled in the framework of the BGEP at 150°W, 75°N, from Aug/2003 to Aug/~~2013.~~ 2013 (Krishfield et al., 2013). Fig. 2 also shows the time scales from the allocated reanalysis-based SIT anomalies. For this geographical location and time span, the draft anomaly from BGEP persists for about 3.7 months, while the SIT anomalies from the different reanalyses persist from 2.4 to 8 months. The persistence is estimated both from a regional and pan-Arctic perspective. First, it is calculated at each grid point, for all SIT anomaly time series. Second, it is estimated for the long-term (GECCO2 and MOVE-CORE) pan-Arctic ice volume anomalies. For the latter case, we evaluate how stable the e-folding time scale is over time by applying a moving (stepped by 1 month) and length-variable window (from 5 to 59 years). Here, we also investigate whether the moving time scale is marked by significant band(s) of variability. To do so, we applied wavelet analysis as proposed by Torrence and Compo (1998).

The length scales of the SIT anomalies are estimated for the reanalysis datasets. The first step is to determine one-point correlation maps. In other words, we calculate the cross-correlation between the SIT anomaly from each grid point with the anomaly from all other points. Subsequently, we make use of the e-folding reference and, for every map, we select all grid cells with correlation coefficient higher than $1/e$. The radius of a circle that yields the area covered by these selected cells is defined as the length scale of the SIT anomaly. This methodology is detailed and graphically presented by Blanchard-Wrigglesworth and Bitz (2014). Fig. 3a shows an example where the length scale is calculated for the SIT anomalies from CryoSat seasonal data: Spring (Mar-Apr) and Autumn (Oct-Nov), from Autumn 2010 to Spring 2017. In turn, ~~Fig~~Figs. 3b-c reveal that a similar length scale pattern is also present in PIOMAS. It is worthwhile mentioning that this illustrative example

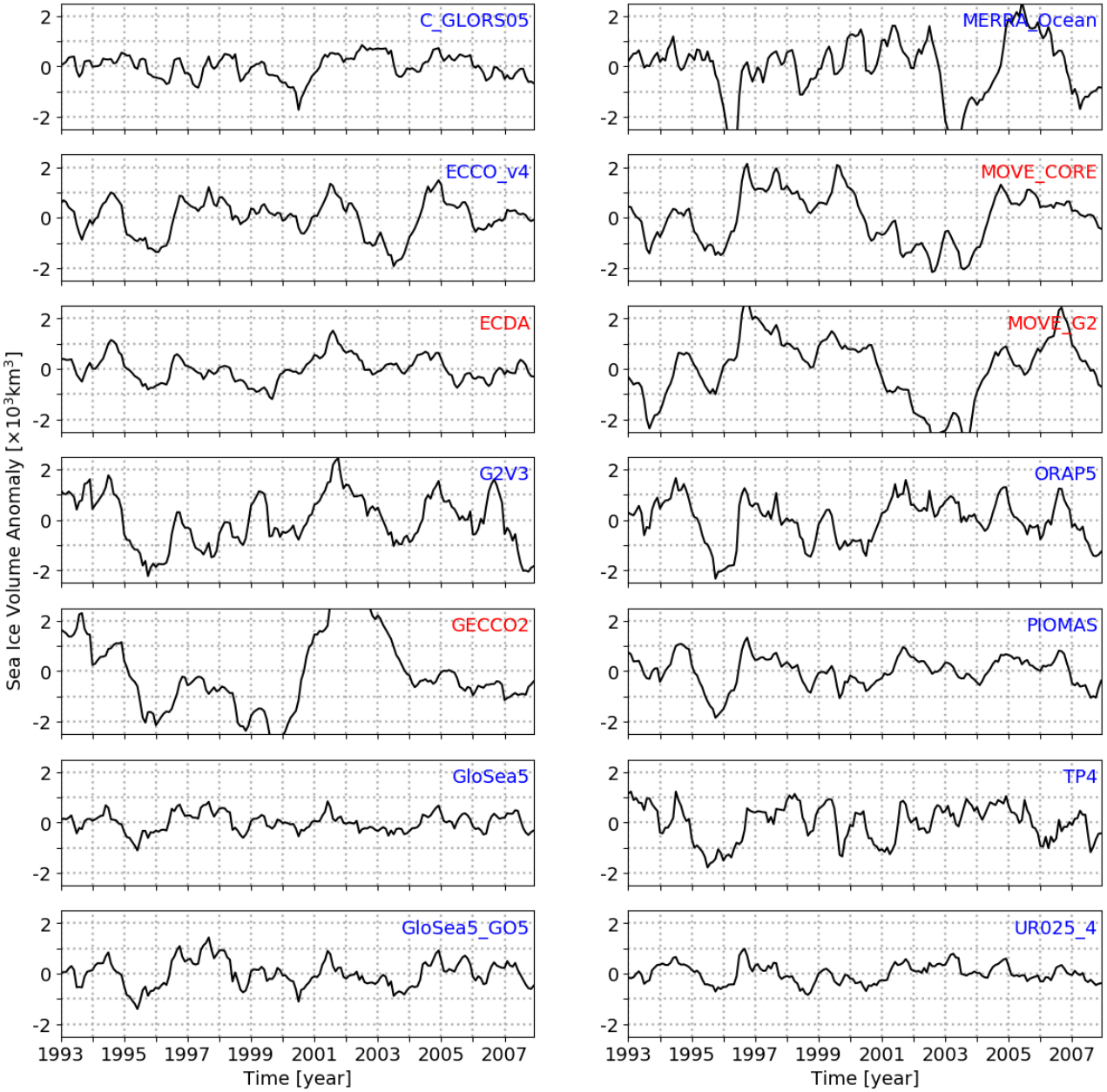

Figure 1. Sea ice volume anomalies estimated from all reanalyses. Anomalies are calculated by excluding the trend and the seasonal cycle. Ticklabels are placed at the first day of the respective year. Reanalyses labeled in blue and red highlight whether the datasets were built with or without sea ice data assimilation, respectively.

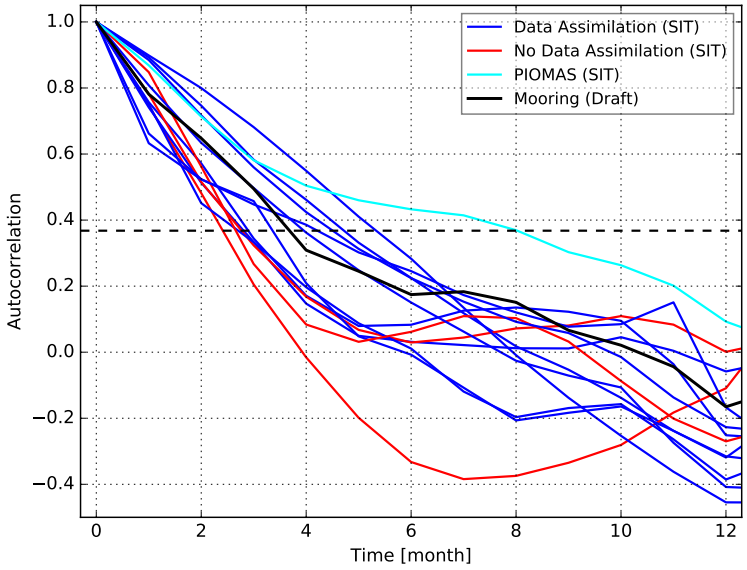

Figure 2. Autocorrelation curves for the draft time series sampled *in situ* by upward-looking sonars deployed in the BGEP oceanographic mooring (black line). This mooring was placed at 150°W, 75°N (see location in Fig. 3c) and the data spans from Aug/2003 to Aug/2013. The blue and red lines display the autocorrelation estimated from the SIT anomalies time series for the ORA-IP reanalyses, at the nearest grid point to the mooring and same time span, for the reanalyses built with and without sea ice data assimilation, respectively. The cyan line indicates the autocorrelation estimated for the PIOMAS reanalysis. The time in which the curves cross the black dashed line is defined as their respective e-folding time scales.

allows a first assessment on how length scales from observations and reanalyses compare to each other. However, it can not be compared to the spatial scales of monthly anomalies further studied in Sec. 3.4.

3 Results

3.1 Comparison of reanalyses with observations

The scatter plots shown in Fig. 4 combine SIT from each reanalysis and the observational datasets from all sources. The latter are separated into two parameters: draft (black dots) and SIT (green dots). The comparisons indicate that all reanalyses are significantly correlated to the observations, whether these are draft or SIT. By comparing SIT and draft, four reanalyses have correlation coefficients larger than 0.7: ~~TP4~~ TOPAZ4 (R = 0.76), C-GLORS05 (0.74), MOVE-CORE (0.74) and UR025-4 (0.73). On the other hand, GECCO2 (0.17) and ~~MOVE-G2~~ MERRA-Ocean (0.10) are marked by the weakest correlations. If we evaluate the reanalyses' statistical capability for predicting the observational values, the MRSS from the linear fit indicates that ~~TP4~~ TOPAZ4 (MRSS = 0.39 m^2), UR025-4 (0.42 m^2) and C-GLORS05 (0.49 m^2) are the best predictors, while ~~MERRA-OCEAN~~ MERRA-Ocean (1.42 m^2) and GECCO2 (1.27 m^2) provide the lowest agreement.

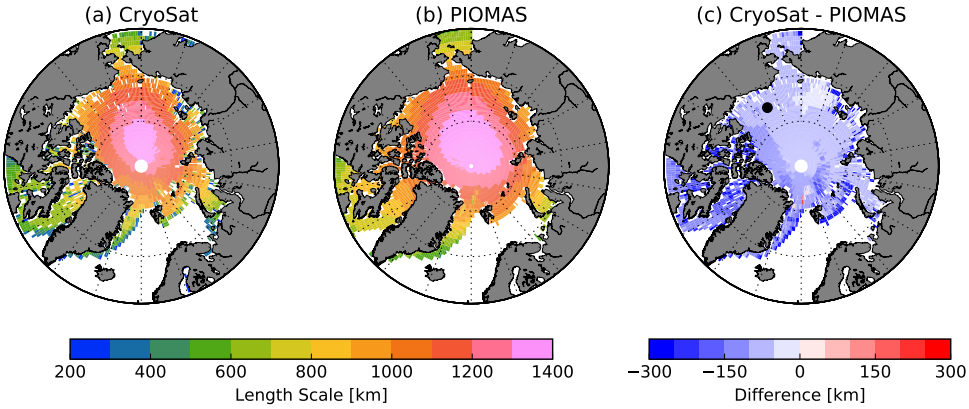

Figure 3. (a) E-folding length scale estimated from the CryoSat seasonal data of sea ice thickness. This dataset contains fourteen Spring (Mar-Apr) and Autumn (Oct-Nov) fields, starting in Autumn 2010 and finishing in Spring 2017. (b) Same as (a), but using the equivalent temporal averages from the PIOMAS data. The difference between the fields shown in (a) and (b) is plotted in (c). The black circle in (c) indicates the location of the mooring from which data is used in Fig. 2.

When comparing SIT from both datasets, the reanalyses with higher correlation coefficient are PIOMAS (R = 0.66), GECCO2 (0.64) and ~~TP4~~ TOPAZ4 (0.61), while ECDA (0.43), ECCO-v4 (0.40) and MOVE-G2 (0.30) are the reanalyses with poorest correlation. In terms of linear fit, PIOMAS (MRSS = 0.41 m^2), ~~TP4~~ TOPAZ4 (0.41 m^2), GECCO2 (0.42 m^2), ORAP5 (0.46 m^2) and C-CGLORS05 (0.49 m^2) are the best performing predictors (MRSS < 0.5 m^2), while MOVE-CORE (0.71 m^2) and ECCO-v4 (0.7 m^2) provide the lowest prediction capability. In addition, a direct comparison by means of RMSEs indicates which reanalyses are closer to the ensemble of observations, as follows: PIOMAS (RMSE = 0.7 m), C-GLORS05 (0.8 m), GloSea5-GO5 (0.8 m), ORAP5 (0.8 m), GECCO2 (0.9 m), GloSea5 (0.9 m), MOVE-CORE (0.9 m), ~~TP4~~ TOPAZ4 (0.9 m), ECCO-v4 (1.0 m), ECDA (1.0 m), MERRA-Ocean (1.0 m), G2V3 (1.1 m), MOVE-G2 (1.1 m) and UR025-4 (1.1 m).

For a detailed overview on how each reanalysis is linked to each observational dataset, in terms of RMSE, MRSS and R, the reader is referred to the tables presented in Appendix B.

3.2 Comparison of reanalyses to each other

As a first assessment of how well the reanalyses compare to each other, we estimate the RMSE and R between time series of SIT anomaly, at every grid point and between all pairs of products. The results are organized as a square matrix in Fig. 5, where the number on the top of each panel represents the respective global value estimated by considering the data from all grid points. The lower triangular part of the matrix reveals that the smallest RMSE is found for the pair ECDA–UR025-4 (RMSE = 0.21 m). Only four other pairs present RMSE $\leq$ 0.25, they are the match between the two GloSea5 products (0.23 m), and the combination of UR025-4 against C-GLORS05 and ECCO-v4 (0.25 m). The largest ~~error~~ RMSE takes place when comparing GECCO2–MOVE-G2 (0.61 m).

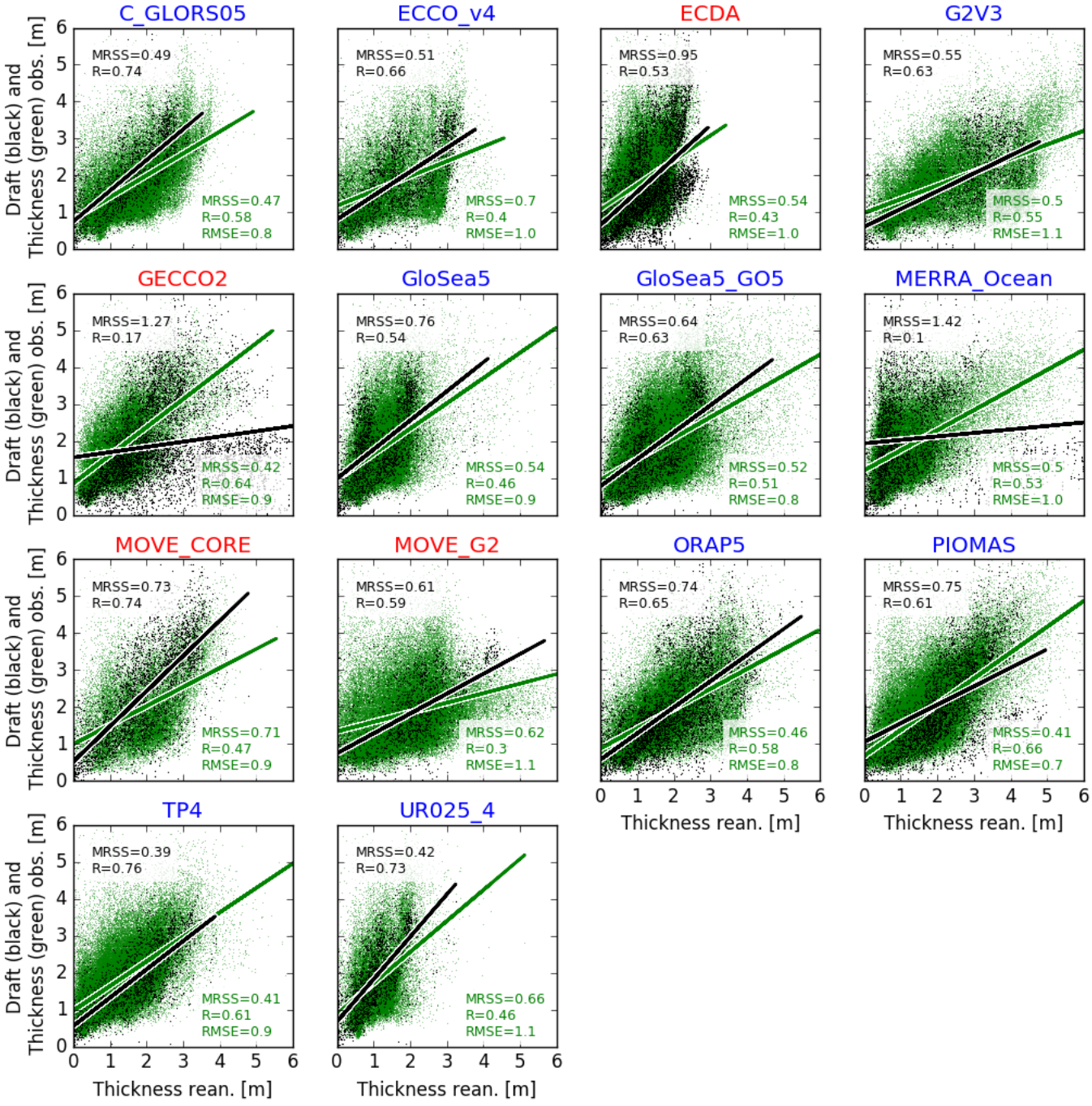

Figure 4. Comparison between sea ice thickness from reanalyses and sea ice thickness (green points) or draft (black points) from observational datasets. The lines represent the linear fits having the reanalysis as the predictor and the observations as predicted variables. The Mean Residual Sum of Squares (MRSS) from the fit, the correlation coefficient (R) and the Root Mean Squared Error (RMSE) are also displayed for each comparison. RMSE is calculated only when comparing SIT from both sources (green), but not when comparing SIT and draft (black). Reanalyses labeled in blue and red highlight whether the datasets were built with or without sea ice data assimilation, respectively.

From Fig. 5 (lower triangle), the averaged RMSE for each individual reanalyses indicates that UR025-4 is the reanalysis closer to the ensemble, while MOVE-G2 has the largest errors compared to its counterparts: UR025-4 (0.30±0.06 m), ECCO-v4 (0.33±0.06 m), ECDA (0.33±0.06 m), GloSea5 (0.34±0.07 m), C-GLORS05 (0.35±0.06 m), PIOMAS (0.35±0.06 m), MOVE-CORE (0.36±0.06 m), GloSea5-GO5 (0.36±0.07 m), ~~TP4~~ TOPAZ4 (0.37±0.06 m), ORAP5 (0.38±0.06 m), G2V3 (0.41±0.06 m), MERRA-Ocean (0.44±0.04 m), GECCO2 (0.45±0.06 m), MOVE-G2 (0.47±0.06 m).

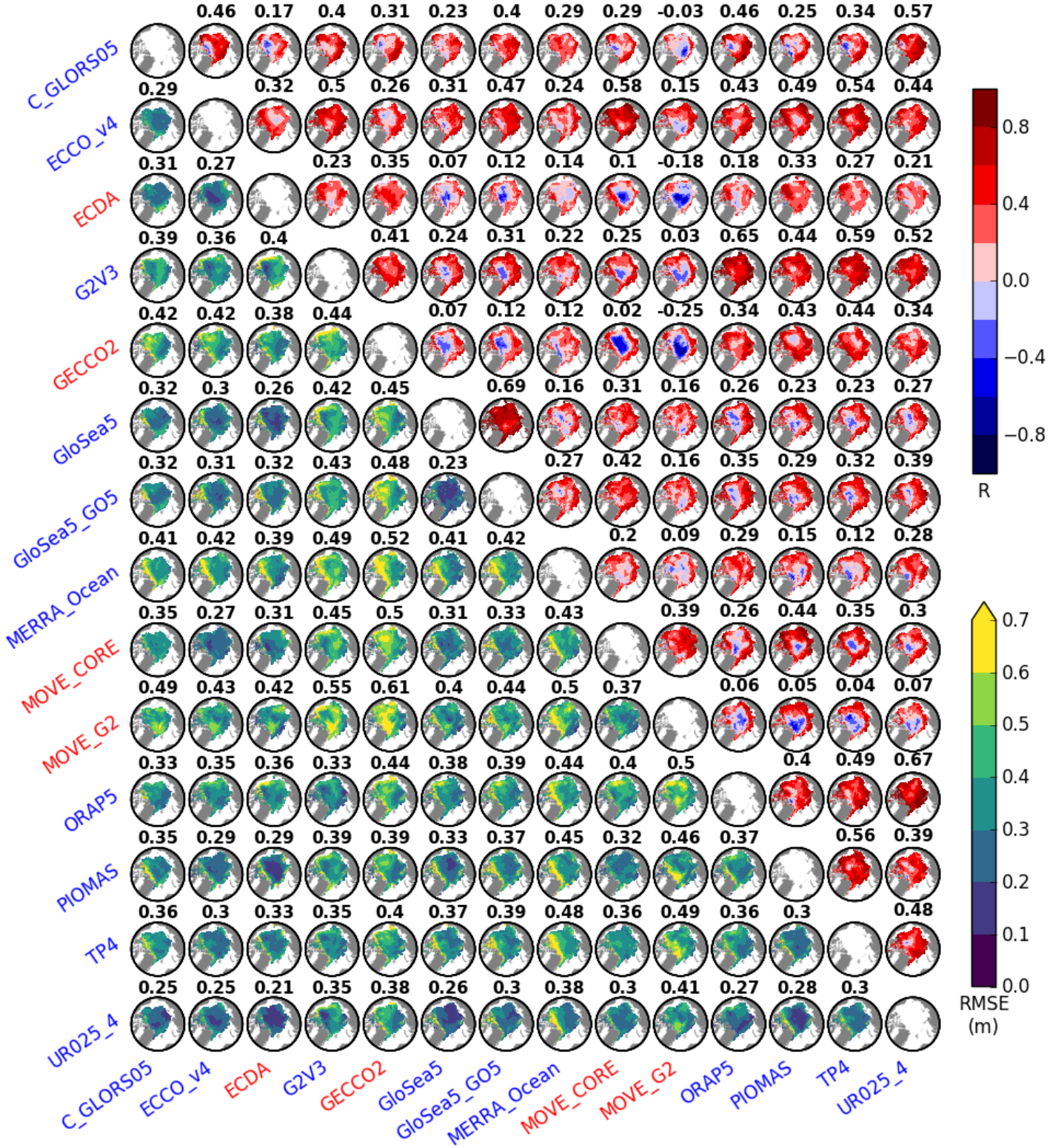
C_GLORS05
ECCO_v4
ECDA
G2V3
GECCO2
GloSea5
GloSea5_GO5
MERRA_Ocean
MOVE_CORE
MOVE_G2
ORAP5
PIOMAS
TP4
UR025_4
0.46 0.17 0.4 0.31 0.23 0.4 0.29 0.29 -0.03 0.46 0.25 0.34 0.57
0.29 0.32 0.5 0.26 0.31 0.47 0.24 0.58 0.15 0.43 0.49 0.54 0.44
0.31 0.27 0.23 0.35 0.07 0.12 0.14 0.1 -0.18 0.18 0.33 0.27 0.21
0.39 0.36 0.4 0.41 0.24 0.31 0.22 0.25 0.03 0.65 0.44 0.59 0.52
0.42 0.42 0.38 0.44 0.07 0.12 0.12 0.02 -0.25 0.34 0.43 0.44 0.34
0.32 0.3 0.26 0.42 0.45 0.69 0.16 0.31 0.16 0.26 0.23 0.23 0.27
0.32 0.31 0.32 0.43 0.48 0.23 0.27 0.42 0.16 0.35 0.29 0.32 0.39
0.41 0.42 0.39 0.49 0.52 0.41 0.42 0.2 0.09 0.29 0.15 0.12 0.28
0.35 0.27 0.31 0.45 0.5 0.31 0.33 0.43 0.39 0.26 0.44 0.35 0.3
0.49 0.43 0.42 0.55 0.61 0.4 0.44 0.5 0.37 0.06 0.05 0.04 0.07
0.33 0.35 0.36 0.33 0.44 0.38 0.39 0.44 0.4 0.5 0.4 0.49 0.67
0.35 0.29 0.29 0.39 0.39 0.33 0.37 0.45 0.32 0.46 0.37 0.56 0.39
0.36 0.3 0.33 0.35 0.4 0.37 0.39 0.48 0.36 0.49 0.36 0.3 0.48
0.25 0.25 0.21 0.35 0.38 0.26 0.3 0.38 0.3 0.41 0.27 0.28 0.3
0.8
0.4
0.0
−0.4
−0.8
R
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
RMSE
(m)

Figure 5. Square matrix-plot displaying the Root Mean Squared Error (RMSE) (lower triangular part) and the Correlation Coefficient (R) maps (upper triangular part), estimated from the sea ice thickness time series, at every grid point and between all pairs of reanalyses. The numbers on the top of each panel indicate the respective value calculated with data from all grid points. All maps have the 0°-longitude placed at 6-o'clock, while the bounding latitude is 67°N. Reanalyses labeled in blue and red highlight whether the datasets were built with or without ice data assimilation, respectively.

At the regional scale, most of the pairs of reanalyses have larger differences off the northern Greenland coast and to the north of the Canadian Archipelago, which are more pronounced in the MERRA-Ocean product. Almost all systems present minimum errors in the central Arctic Basin.

In turn, the upper triangular part of the matrix in Fig. 5 displays the linear relationship between pairs of reanalyses, quantified by the correlation coefficient. The strongest pan-Arctic correlations are observed for GloSea5–Glosea-G05 (R = 0.69), ORAP5–UR025-4 (0.67) and G2V3–ORAP5 (0.65). MOVE-CORE and MOVE-G2 present a marked anti-correlation with several other reanalyses, mainly in the central Arctic Ocean. Such anti-correlation is also reflected in the sea ice volume anomalies shown in Fig. 1. Notice that negative anomalies in MOVE-CORE and MOVE-G2, for instance from 2001 to 2004, ~~are associated with~~ occur at the same time that strong positive anomalies in reanalyses as GECCO2, G2V3 and ECDA, as well as C-GLORS05, ORAP5, PIOMAS, ~~TP4~~ TOPAZ4 and UR025-4, though in a less pronounced way (Figs. 1 and 5). We do not have a clear understanding of why these anti-correlations take place.

3.3 Time scales

An important property inherent to time series in general concerns their time scale and/or persistence, as defined in Sec. 2.3. In other words, we aim to infer for how long the SIT anomaly maintains a good correlation with future measurements at the same grid cell. Persistence can also be perceived as ~~part~~ the skill of a self-prediction scheme where past data is used to predict future values. In addition, it is a relevant variable to be taken into account when designing the sampling frequency of observational programmes, specially if these programmes target the understanding of the SIT ~~variability~~time-variability.

Fig. 6 displays the e-folding time scales for the SIT anomaly at every grid point, and for all reanalyses. The Area Weighted Mean (AWM) time scales (in months) sorted in ascending order are: 2.5 (GloSea5), 2.6 (GloSea5-GO5), 3.6 (PIOMAS) 3.7 (ECCO-v4), 3.8 (MERRA-Ocean), 4.0 (UR025-4), 4.3 (~~TP4~~TOPAZ4), 4.4 (C-GLORS05), 4.7 (ORAP5), 4.9 (MOVE-CORE), 5.0 (G2V3), 6.0 (ECDA), 7.2 (MOVE-G2), and 7.8 months (GECCO2). These values were calculated taking into account only grid points with a valid SIT value from all reanalyses.

The results reveal that the thickness anomalies from reanalyses with no ice data assimilation (NA; Fig. 6, red labels) present a longer persistence, mainly distinguished in MOVE-G2 and GECCO2. Potential reasons to explain why the thickness anomalies persist longer in NA systems are suggested and discussed in Sec. 4. On the contrary, the thickness anomalies from the GloSea5 systems (GloSea5 and GloSea5-GO5) have a much shorter persistence.

From a regional point of view, Fig. 6 shows that GloSea5 and GloSea5-G05 are the only reanalyses in which the SIT anomaly persistence is remarkably short all over the Arctic, presenting e-folding time scales higher than 4 months only in a

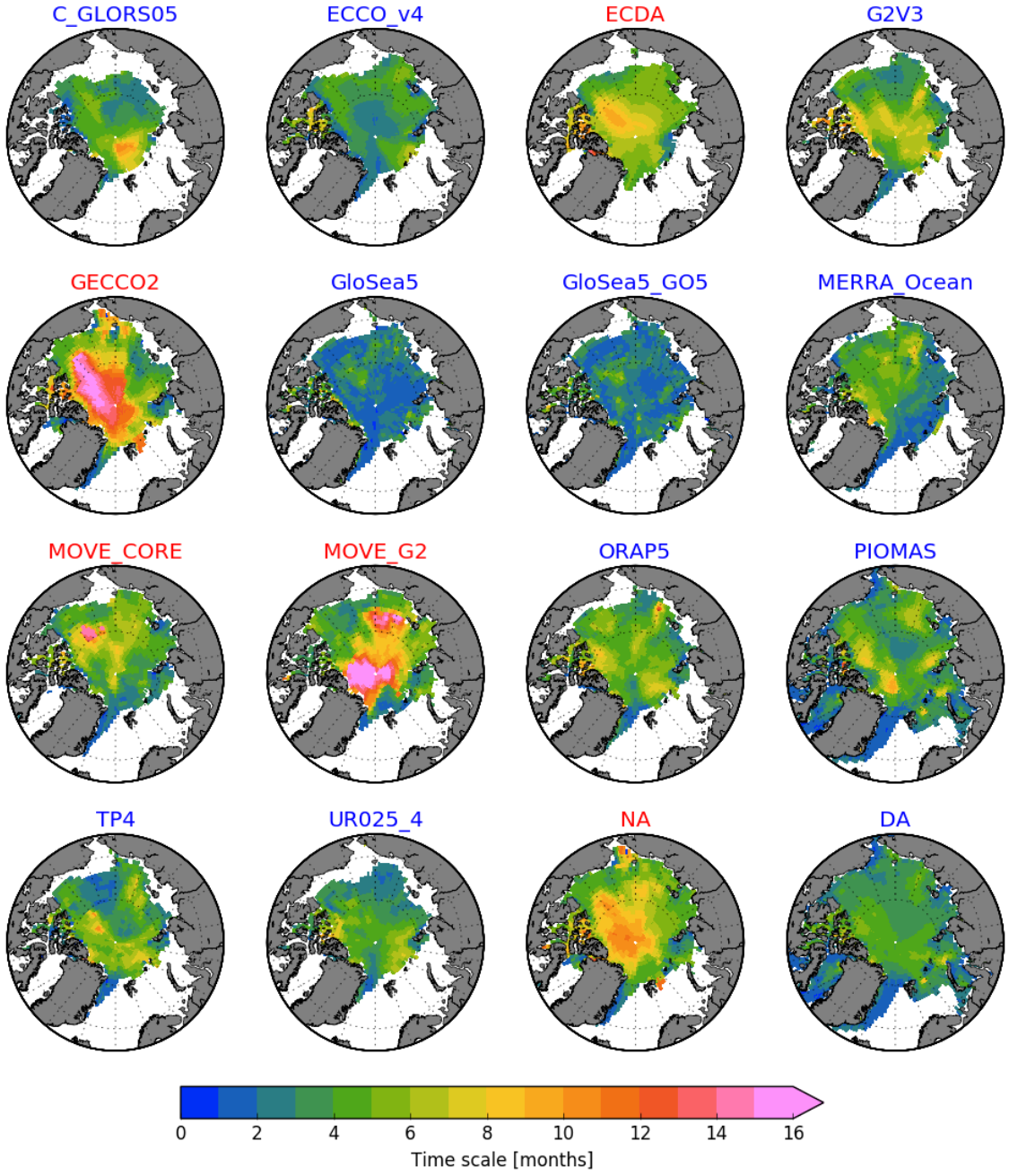

Figure 6. E-folding time scales (or persistence) estimated for the SIT time series. Only grid cells in which the time-mean (for the Jan/1993–Dec/2007 period) SIT at the time of summer minimum is greater than 0.1 m are taken into account for the computations. Averages for the systems with ice Data Assimilation and No data Assimilation are represented by the NA and DA panels. All maps have the 0°-longitude placed at 6-o'clock, while the bounding latitude is 67°N. Reanalyses labeled in blue and red highlight whether the datasets were built with or without ice data assimilation, respectively.

few, not evenly distributed, grid points. By contrast, the SIT from GECCO2 has a marked longer persistence (>15 months) extending from the region off the northen coast of Greenland to the north of the Canadian Archipelago and mid Arctic Ocean. The ECDA product presents a relatively similar pattern of time scale over the region mentioned above, but persisting for a shorter period (∼8 months). SIT anomalies from MOVE-G2 also indicate long persistence off the northern Greenland coast, extending to the central Arctic and East Siberian Sea. ~~The~~ For the remaining reanalyses, there is not a common regional pattern of persistence outstanding from their respective time scale maps~~from G2V3, ORAP5, PIOMAS, TP4 and UR025-4 present a slight resemblance, but not so clear as between GloSea5 and GloSea5-G05~~.

~~E-folding time scales (or persistence) estimated for the SIT time series. Only grid cells in which the time-mean (for the Jan/1993–Dec/2007 period) SIT at the time of summer minimum is greater than 0.1 m are taken into account for the computations. Averages for the systems with ice Data Assimilation and No data Assimilation are represented by the NA and DA panels. All maps have the 0°-longitude placed at 6-o'clock, while the bounding latitude is 67°N. Reanalyses labeled in blue and red highlight whether the datasets were built with or without ice data assimilation, respectively.~~

Nevertheless, the results above should be interpreted with caution. The e-folding time scale is a metric that depends on the shape of the lagged-autocorrelation curve, which in turn may differ according to the period and time span of the original time series being analyzed. In order to evaluate how stable the time scale is by varying the time span and also by evolving over time, we applied a time-moving and length-variable window to calculate the e-folding time scale of the ice volume anomaly (detrended in the same way as the SIT time series) from the two longest reanalyses (GECCO2 and MOVE-CORE), as shown in Figs. 7a,e. The window length varies from 5 to 59 years (stepped by 1 year) and it moves over time stepped forward by one month. Here, we use the ice volume anomaly, rather than SIT anomaly, for two reasons. First, because it is computationally more efficient than calculating the time scale for the SIT anomalies at every grid cell, considering the large number of interactions for a time-moving and length-variable window. Second, because the volume provides a pan-Arctic perspective of the SIT persistence.

Notice in Figs. 7a,e that the persistence overall grows to longer than ∼20 ~~years~~ months when taking into account long time spans, remarkably for GECCO2 in which ice volume anomaly persists for longer than 25 ~~years~~ months at several center times. As for the thickness anomalies, MOVE-CORE presents a shorter persistence compared to GECCO2.

As a measure of stability, we estimate the standard deviations for all computations displayed in Figs. 7a,e. Results show that MOVE-CORE has a more stable time scale, with standard deviation of 3.0 months from its mean (9.7 months), while GECCO2 presents average and standard deviation of 15.0±6.5 months.

Figs. 7b,f show the case where the window length is 15 years, as it is for the overlapping period Jan/1993–Dec/2007. For this case, the average (standard deviation) time scales for GECCO2 and MOVE-CORE are 11.4±2.6 and 9.1±2.5 months, respectively. Minimum to maximum ranges are 6.2–16.5 months for GECCO2 and 4.9–13.5 months for MOVE-CORE. If we take into account the same center time of the time span Jan/1993–Dec/2007, that is mid-Jun/2000, the ice volume anomaly persistences are 13.6 and 9.2 months (red stars in Figs. 7b,f, respectively). Note that the time scales of the ice volume anomalies are a few months longer compared to the persistence of the AWM-thickness anomalies (9.2 months, GECCO2; 5.4 months, MOVE-CORE).

We make use of wavelet analysis (Torrence and Compo, 1998) to evaluate whether the time series displayed in Figs. 7b,f exhibit significant band(s) of variability. Figs. 7c,d reveal that the ice volume anomaly from GECCO2 presents two bands of significant variability, as highlighted by the horizontal gray bars in Fig. 7d. The first spans from 4.4 to 6.1 years, and it is present in the first half of the time series but does not persist over time (black contours in Fig. 7c). The second is marked by periods longer than 10.7 years, which seems to be recurrent over time, but should be interpreted with caution since it is placed near the “cone of influence", where edge effects become important, as indicated by cross-hatched areas overlapping the black contours in Fig. 7c. The ice volume anomaly from MOVE-CORE, in turn, is marked by a single band of significant variability, with periods longer than 12.7 years (Figs. 7g,h). Again, this band should be interpreted with caution since it is also placed near the “cone of influence".

3.4 Length scales

The e-folding length scale is a metric used for indicating how well a variable from a certain grid cell compares to the neighboring cells. In other words, it shows how the anomalies spread in space. As for the time scale, the length scale is a promising parameter to be explored when designing observational systems, but in terms of spatial coverage of instruments. Simplistically, regions marked with high length scales would require less instruments to be better monitored.

Fig. 8 shows the length scales for the SIT anomaly at every grid point. The AWM-length scales, in kilometers (km) and ascending order, for each system are: 337.0 (GloSea5), 420.7 (GloSea5-GO5), 544.6 (C-GLORS05), 681.5 (MERRA-Ocean), 724.3 (~~TP4~~TOPAZ4), 728.2 (G2V3), 596.9 (ORAP5), 597.4 (UR025-4), 730.2 (PIOMAS), 732.5 (ECCO-v4), 846.7 (MOVE-G2), 835.8 (MOVE-CORE), 934.0 (ECDA), 935.7 km (GECCO2).

A similar pattern to the time scale is here observed, with GloSea5 and GloSea5-GO5 presenting the minimum length scales, rarely higher than 500 km, while the reanalyses without sea ice data assimilation are characterized by higher length scales, sometimes higher than 1200 km. In all systems the length scales are relatively longer near the central Arctic. This suggests that higher length scales ~~are~~ could be somehow associated with thicker ice. The relationships between mean ice thickness *vs.* time scale *vs.* length scale will be explored in details in Sec. 4.

The stability of the length scale over time (Fig. 9) was tested by means of a moving window with 15 years length, as follows: first, we calculate the one-point correlation maps for every grid point; second, we estimate the length scale for each one-point correlation map; third, the AWM-length scale was calculated taking into account only grid points with a valid SIT value from all reanalyses; fourth, the process was repeated by stepping forward the 15-year window by 12 months. It is worthwhile mentioning that, computationally, it is much more expensive to calculate the length scale than the time scale. This is the reason why, here, we just use a time-moving but not length-variable window. The results suggest that the length scale is relatively more stable than the time scale (~~Fig~~Figs. 7b,f), ~~for what a quantification is presented in the conclusions of this study~~as further discussed in Sec. 4.

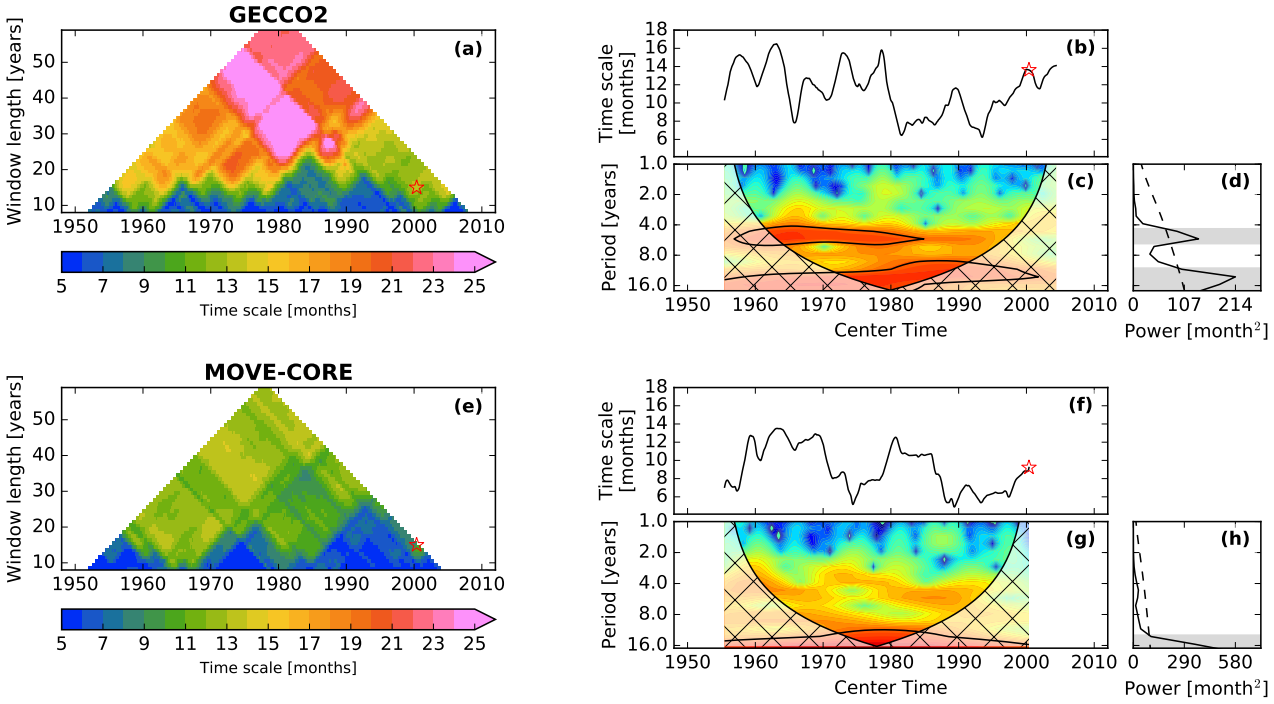

Figure 7. (a) Moving e-folding time-scales estimated for the ice volume anomaly time series from the GECCO reanalysis. The window length varies from 5 to 59 years and it is stepped forward by one month over a total period of 64 years (Jan/1948–Dec/2011). (b) Moving e-folding time-scales for the 15-year window length case. The red stars in (a) and (b) indicate the 15-year overlapping period (Jan/1993–Dec/2007, center time mid Jun/2000). (c) Wavelet power spectrum of (b), with Morlet as wavelet mother. The black lines denote the 95% significance levels above a red noise background spectrum, while the cross-hatched areas indicate the "cone of influence" where edge effects become important. The colorbar is omitted in panel (c) since we are not interested in the power's magnitude but in the frequencies outstanding as significant in the spectrum. (d) Time-integrated power spectrum from the wavelet analysis, where the dashed line corresponds to the 95% significance level. The bands of significant periods (4.4–6.1 years and $>$10.7 years) are highlighted by the gray horizontal bars. (e–h) Same as (a–d), respectively, but for the MOVE-CORE ice volume anomaly which has a spanning period of 60 years (Jan/1948–Dec/2007). The horizontal gray bar in (h) highlights the only period band of significant variability, being defined by periods longer than 12.7 years.

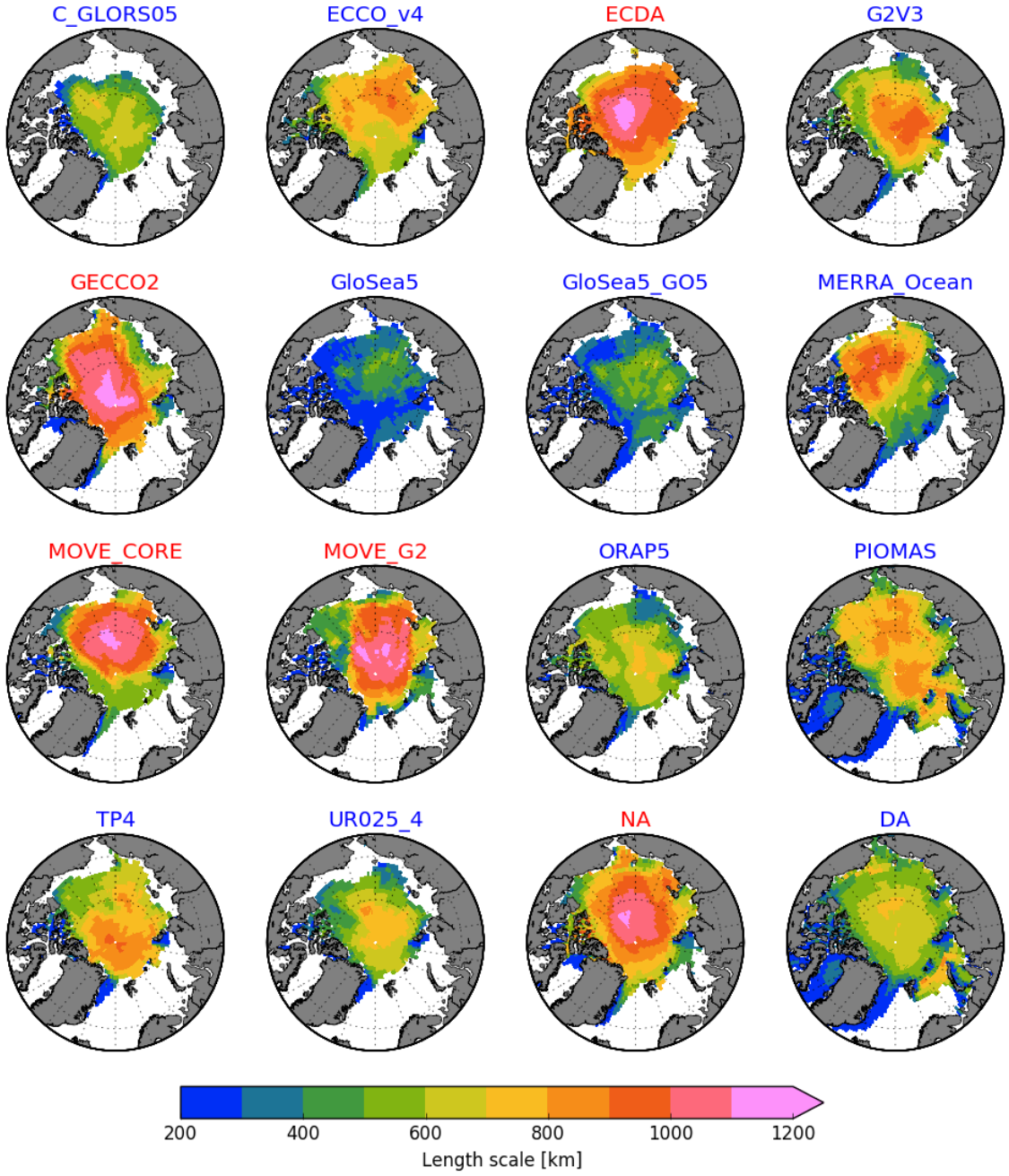

Figure 8. E-folding length scales estimated for the SIT time series. Only grid cells in which the time-mean (for the Jan/1993–Dec/2007 period) SIT at the time of summer minimum is greater than 0.1 m are taken into account for the computations. Averages for the systems with ice Data Assimilation and no Data Assimilation are represented by the NA and DA panels. All maps have the 0°-longitude placed at 6-o'clock, while the bounding latitude is 67°N. Reanalyses labeled in blue and red highlight whether the datasets were built with or without ice data assimilation, respectively.

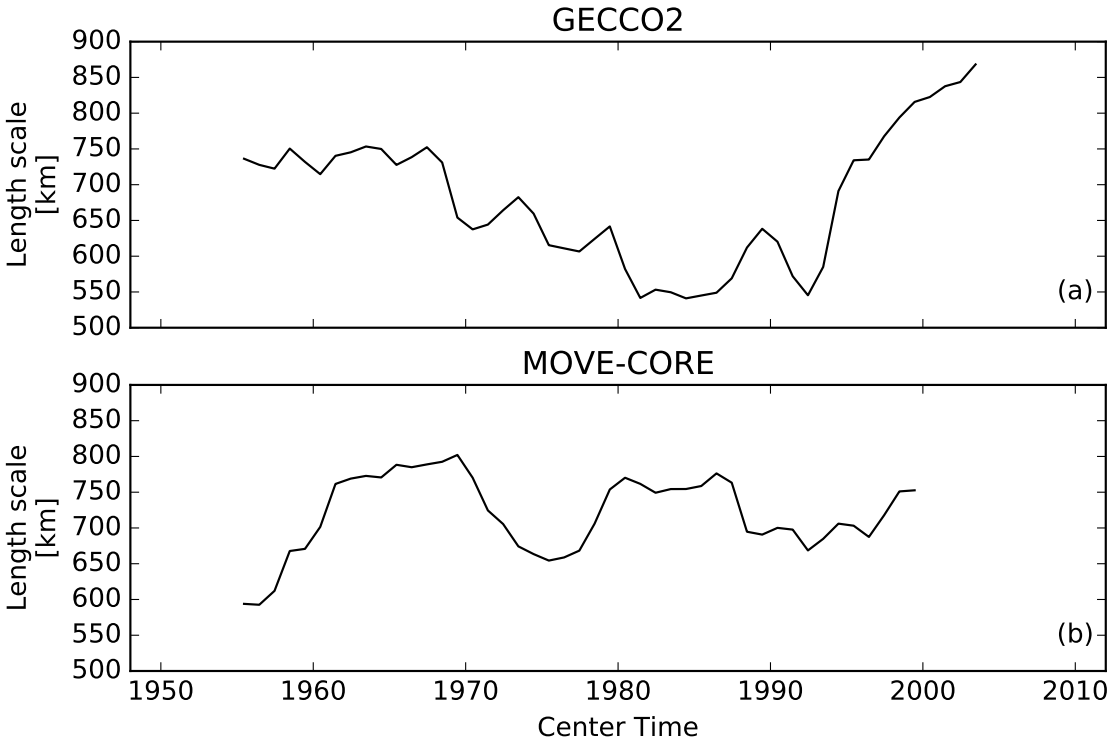

Figure 9. (a) Moving e-folding length scales estimated for the ice volume anomaly time series from the GECCO reanalysis. The window length is 15 years and it is stepped forward by 12 months over a total period of 64 years. (b) Same as (a), but for the MOVE-CORE reanalysis, which has a time span of 60 years.

4 Discussion and conclusions

The first aim of this study was to evaluate how the SIT from the reanalyses compares against observational datasets, either draft or SIT. We have used three different metrics to perform this comparison: the correlation coefficient (R), as a measure of the linear correlation between datasets; the Mean Residual Sum of Squares (MRSS), as an indicator of whether reanalysis values are good predictors for the observations; and the Root Mean Square Error (RMSE), which directly compares how the SIT from the reanalyses approaches the SIT from observations. The results show that some of the reanalyses have a relatively good correspondence either comparing SIT and draft or SIT from both sources of data. This is the case, for instance, for the ~~TP4~~ TOPAZ4 product. A direct comparison between SIT from ~~both sources~~ all reanalyses and observations indicates RMSEs ranging from 0.7 to 1.1 m. PIOMAS has the best agreement with ~~observations. An interesting~~ the observational datasets. A particular case is GECCO2, which presents a relatively small RMSE, as well as good correlation and linear relationship to the SIT observational datasets. However, this same product is weakly correlated and has poor predictive skill to the draft observational datasets.

One of our main goals in performing such a comparison was to identify whether or not systems built with assimilation of sea ice concentration data are closer to observations, compared to the products built with no sea ice data assimilation. The results ~~are not conclusive and~~ suggest that reanalyses with sea ice data assimilation do not necessarily perform better. One could speculate that some reanalyses do not take the best advantage of the covariances between sea ice concentration and SIT.

~~It is worthwhile mentioning that the~~ We have compared the mean state (mean SIV) and respective variability (std SIV) of all reanalyses against the specifications and parameters displayed in Table 1. Nevertheless, for such a comparison, where each system has its own configuration with several varying parameters, we were not able to distinguish the effect that the selected parameters may have on the mean state and variability. The comparison among SIT from the different reanalyses (Sec. 3.2) is not straightforward and does not necessarily improve due to common specifications and key parameters from the two systems being compared. For instance, the pair C-GLORS05–G2V3 ~~presents relatively high RMSE (0.39 m), and not too strong correlation (R = 0.4), even though both systems~~ share a set of ~~specifications and parameters such as~~ common assumptions (ocean–sea ice model, atmospheric forcing, vertical discretization, number of ice thickness categories, dynamics-EVP, ocean-ice drag coefficient, analysis window and ~~also~~ both assimilate sea ice data), but they still present relatively high RMSE (0.39 m) and not so strong correlation (R = 0.4), as shown in Fig. 5. Only a few different assumptions and parameters, as well as their non-linear interactions, may result in systems with considerably distinct mean state and variability. Although the pair C-GLORS05–G2V3 shares several common aspects, these two systems assume different air-ice drag coefficient and also assimilate the sea ice data in a different way, for instance. The same ~~is valid for~~ statement could be applied to other pairs of systems, e.g. ~~, G2V3–ORAP5~~G2V3-ORAP5, which also share some similarities but are still distinct in terms of mean state and variability.

The pair with smallest RMSE (Fig. 5), ECDA–UR025-4, has at the same time a ~~relatively~~ weak linear relationship (R = 0.21). This reinforces the importance of looking at different metrics, when comparing different products. If we average the RMSEs that one specific reanalysis ~~present~~ presents against all the others (Fig. 5), and so compare with the pan-Arctic mean ice volume of this same ~~reanalyses~~reanalysis, it becomes clear that products with relatively low sea ice volume (i.e., thin ice) present small RMSEs when compared with their counterparts (Fig. ~~??). This~~ 10), although MERRA-Ocean is an outlier in this pattern by presenting a thin sea ice but large RMSE compared to the other reanalyses (Fig. 10, left upper corner). Fig. 10 helps to explain why ECDA and UR025-4 have a small RMSE~~(Fig. ??)~~, though their anomalies are marked by a ~~not so strong correlation .~~ relatively weak correlation as suggested in Fig. 5. This is also evident in the respective sea ice volume anomalies from these ~~both~~ two reanalyses shown in Fig. 1. ~~MERRA-OCEAN is an outlier in this pattern, presenting a thin sea ice but large RMSE compared to the other reanalyses~~In the same way, Fig. 5 also indicates that the large differences in the SIT field take place near the northern Greenland coast and Canadian Archipelago, which are regions marked by the thickest sea ice over the studied domain.

Another main goal of this work was to characterize the time and length scales of the sea ice thickness anomaly, as well as to report whether these parameters are influenced by the fact that a respective reanalysis assimilates or not sea ice ~~concentration~~ data. In this case, sea ice data assimilation plays a clear role in the referred scales: systems with sea ice data assimilation are characterized by shorter time and length scales compared to the systems which do not assimilate sea ice data. Nevertheless, a

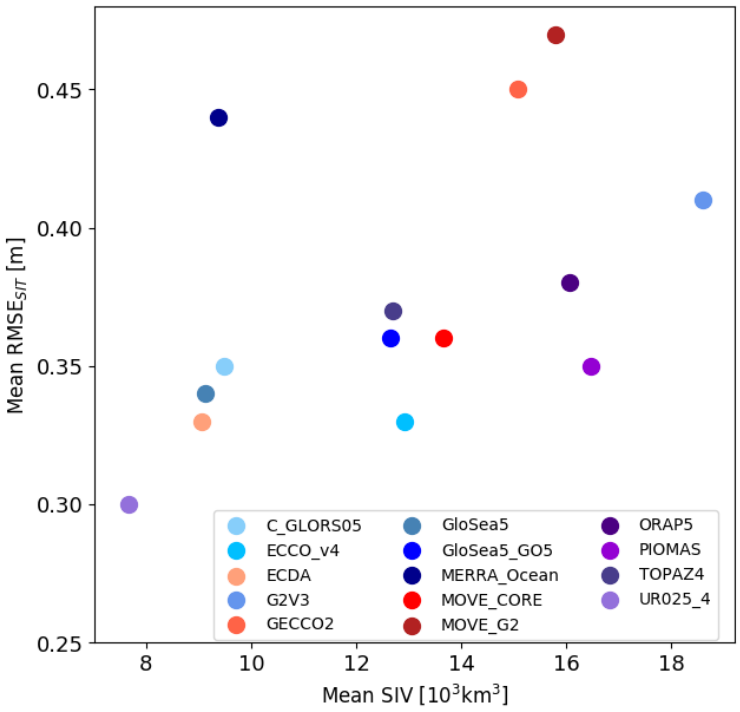

Figure 10. Time-mean sea ice volume *versus* the mean RMSE. This last parameter is an average of the RMSEs that each reanalysis has when its SIT field is compared individually to the other thirteen reanalyses~~. Reanalyses~~ , as shown in Fig. 5. Shades of blue and purple indicate the reanalyses which do assimilate sea ice data, while shades of red ~~highlight whether~~ indicate the ~~datasets were built with or~~ reanalyes without sea ice data assimilation~~, respectively~~.

comparison between the same system but built with (G2V3) and without (G2V1; not included in the 14 reanalyses of the present study) assimilation of sea ice concentration data (Fig. 11) suggests that this finding is valid in terms of pan-Arctic averages, but not necessarily at every grid cell. This may explain why in the specific location addressed in Fig. 2 the reanalyses with data assimilation showed relatively longer time scales compared to the reanalyses without data assimilation. The pan-Arctic AWM-TS and AWM-LS from G2V3 are 5 months and 728.2 km, respectively. Without sea ice data assimilation (G2V1), the AWM-TS and AWM-LS increase to 5.5 months and 745.3 km, respectively.

Likely, the main reason why ~~this happen~~ the assimilation of sea ice concentration data impacts the time and length scales of the SIT field is linked to the fact that when a reanalysis assimilates sea ice information, the system is forced towards the assimilated conditions, differently from what occurs with free-running models. Eventually, data assimilation introduces SIT increments that are not necessarily physical, and so contributes to an attenuation in the correlation of this variable at a certain grid cell both in time, with their future estimations, and in space, with the neighboring grid points. ~~Fig. ?? shows an interesting example where the time scales are estimated for the same system, but built with (G2V3)and without (G2V1)assimilation~~

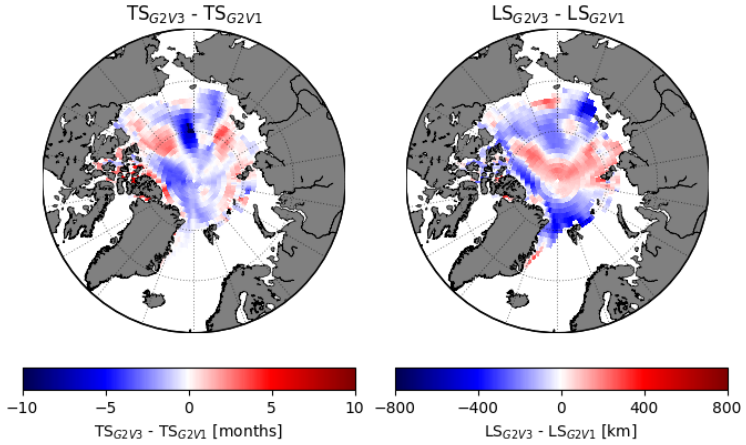

Figure 11. Grid-point differences (G2V3 − G2V1) of time scale (a) and length scale (b) between two versions of the GLORYS system: G2V3 which assimilates sea ice data and G2V1 which does not assimilate sea ice data.

We have shown that time and length scales are clearly influenced by whether or not the reanalyses assimilate sea ice data, as graphically represented in Figs. 12a-b. However, are these two properties also clearly influenced by other specifications and parameters? Figs. 12c-h show how time and length scales are linked to the choices of atmospheric forcing, sea ice model and dynamics for ice-ice interactions that control ice deformation (VP or EVP). Even though the atmospheric forcing fields are reported to play a major role in the sea ice simulations (Gerdes and Köberle, 2007; Rothrock and Wensnahan, 2007a), we could not identify distinguished patterns between the two main sources of atmospheric forcing used by the ensemble of reanalyses: Era-Interim and NCEP/NCAR (Figs. 12c-d). Likewise, time and length scales are not clearly linked to the choices of sea ice ~~concentration.Again, the results show relatively longer time scales for the system with no data assimilation~~model (Figs. 12e-f) and ice deformation dynamics (Figs. 12g-h), although a certain coherence in the time and length scales is observed for the systems that use the Louvain-la-Neuve Sea Ice Model (LIM; Figs. 12e-f).

~~The~~ Besides the spread among the points, the scatter plots displayed in Figs. 13a-b indicate a certain correlation between the time-mean SIV (mean state) and the studied scales, where relatively thin ice leads to shorter scales, in agreement with Massonnet et al. (2018). In contrast, the time scales have a marked anti-correlation with the sea ice ~~velocities~~ drift as shown in Fig. ~~??~~13c: reanalyses with faster sea ice present a short time scale~~and the other way around. The same is observed~~. Such correlation is less pronounced for the length ~~scales, though not so marked as for time scales (not shown~~scale (Fig. 13d). Different ~~specifications~~ parameters from the reanalyses could potentially influence the sea ice velocity. ~~High~~ For instance, high air–ice and low ocean–ice drag coefficients contribute to faster ice velocities, respectively (Tandon et al., 2018). ~~For instance~~As an example, ECCO-v4 has the second highest air–ice (smaller only compared to MOVE-CORE) and the smallest ocean–ice drag coefficients (see Table ~~1 from Chevallier et al. (2017)~~1). This may explain why ECCO-v4 has relatively high ice velocities (Fig. ~~??~~13c) and, therefore, low time and length scales (Figs. 6 and 8). ~~In addition, the~~ For our ensemble of reanalyses, Fig. 13e

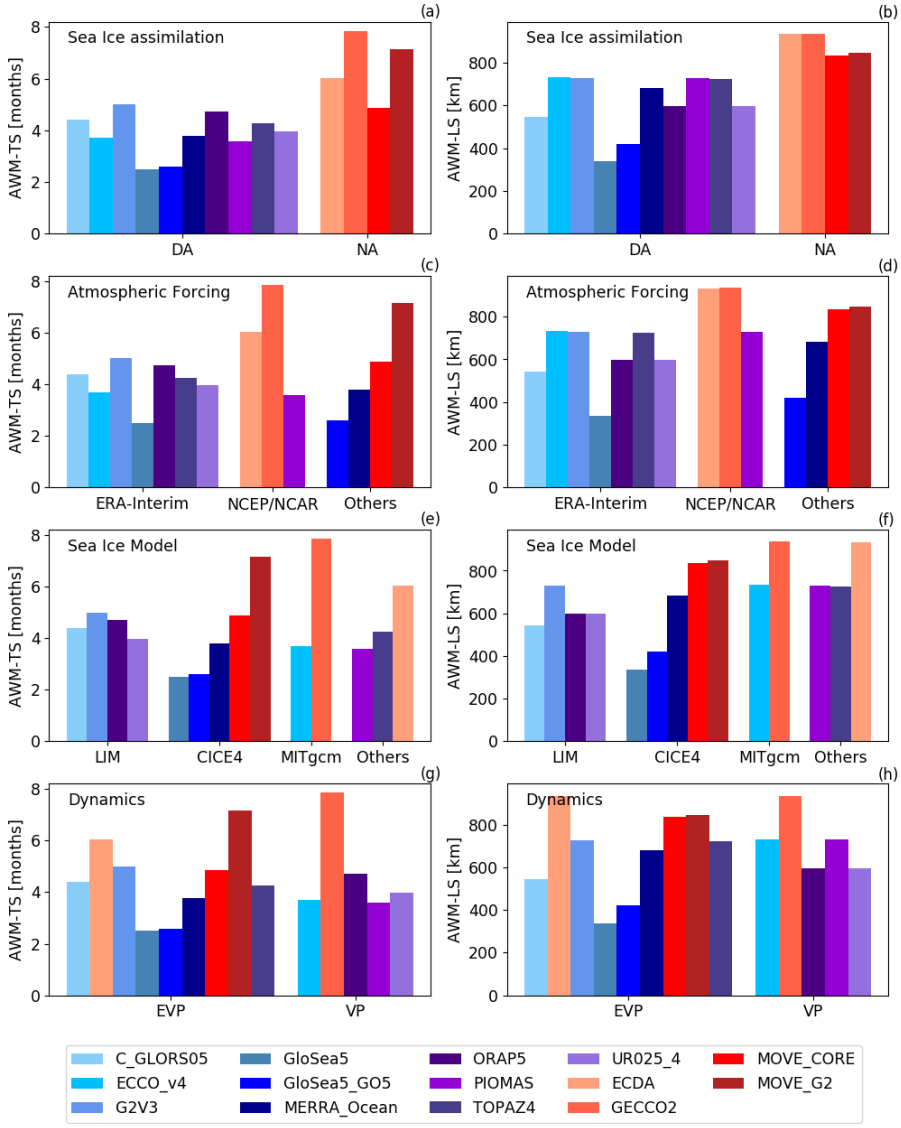

Figure 12. ~~Time scale estimated~~ Histograms showing how the AWM-TS (left panels) and AWM-LS (right panels) are related to different reanalyses specifications: (a-b) whether or not the ~~same reanalysis, but built with~~ system assimilates sea ice data; (~~G2V3~~c-d) the source of atmospheric forcing data; (e-f) the used sea ice model; and (~~G2V1~~g-h) the used dynamics (Viscous-Plastic or Elastic-Viscous-Plasctic) for ice-ice interactions that control ice deformation. Shades of blue and purple indicate the reanalyses which do assimilate sea ice data, while shades of red indicate the reanalyes without sea ice data assimilation.

shows a tight correlation between sea ice velocity and the drag air–sea coefficients. Again, this correlation is less pronounced for the length scale Fig. 13f.

The ice strength formulation is a major player in the sea ice velocity (Ungermann et al., 2017). All reanalyses follow the linear parameterization proposed by Hibler (1979), except the GloSea5 products and PIOMAS, which employ the ice strength formulation following Rothrock (1975). A higher ice strength parameter P^* in Hibler's formulation leads to thicker and slower-moving ice, what would potentially lead to larger scales. Nevertheless, a relation between P^* and the scales are not so clear for the ensemble of reanalyses (not shown). In addition, Ungermann et al. (2017) presented a detailed study comparing ~~both~~ Hibler's and Rothrock's methods and have shown that, for systems characterized by relatively thinner ice, model simulations with Rothrock (1975) formulation result in lower ice strength, and therefore faster ice velocities, compared to Hibler (1979) formulation. ~~The~~ We do not have a clear understanding of why the GloSea5 products present such short time and length scales, however, we could speculate that the combination between relatively thin ice (~~see colorbar in~~ Fig. ~~??) and~~ 13a) and the use of Rothrock (1975) ice strength formulation ~~is a potential explanation of why GloSea5 and GloSea5-GO5 have such a short time and length scales~~could play in that direction.

~~We have shown~~ The discussion above indicates that time and length scales are ~~influenced by~~ mainly driven by the fact of whether or not the reanalyses assimilate sea ice data~~. However, are these two properties correlated to each other? Also, are they somehow related to the mean SIT of the system?~~ , but also influenced by the air-ice drag coefficient and sea ice drift. Fig. ~~??~~ ~~shows that~~ 14 shows a strong correlation between both time and length scales, where long time scales are associated with large length scales. ~~This figure also compares the respective scales against the time-mean ice volume. Interestingly, for the reanalyses with ice data assimilation, an increase in mean ice thickness seems to lead to an increase in time and length scales, as it is suggested by the points in~~ Notice in this plot the difference in the ~~bottom-left quadrant defined by the gray dashed lines (Fig. ??). This is in agreement with the higher~~ time and length scales ~~associated with slower ice shown in Fig. ??, suggesting that by increasing ice thickness, ice moves slower, resulting in higher length and time scales. The extreme example is~~ from G2V1 and G2V3 ~~which has the thickest ice and also relatively high length and time scales. This does not seem to be the case for the systemswhich do not assimilate ice data as indicated by the points in the upper-right quadrant~~systems, highlighted by the gray rectangle placed near the center of the figure.

As mentioned before, the ice thickness time and length scales are interesting properties to be explored when designing and ~~planing~~ planning an optimal observation system both in terms of temporal sampling and spatial placement of instruments. Nevertheless, ~~Sec. 3.3 (Fig. 9) and Sec. 3.4 (Fig. 7) showed that~~ these properties also vary over time ~~. The over-time stability (Figs. 7 and 9. The stability over time~~ of time and length scales can be ~~quantified~~ estimated by the coefficient of variation (C_v=std/mean~~, where std is the standard deviation~~). The C_v is a non-dimensional metric used to evaluate the extent of a certain variability in relation to its mean, allowing to compare different properties. The C_v for the GECCO2 and MOVE-CORE moving time scales, estimated from the time series shown in ~~Fig~~Figs. 7b,f, are 0.23 and 0.27, respectively, while the C_v for moving length scales (Fig. 9) are 0.13 and 0.07.

Lastly, it is worthwhile mentioning that ~~we still consider~~ both time and length scales are promising properties to support the design of an optimal observing system. As suggested by the C_v presented above, and by the fact that the time scale is more

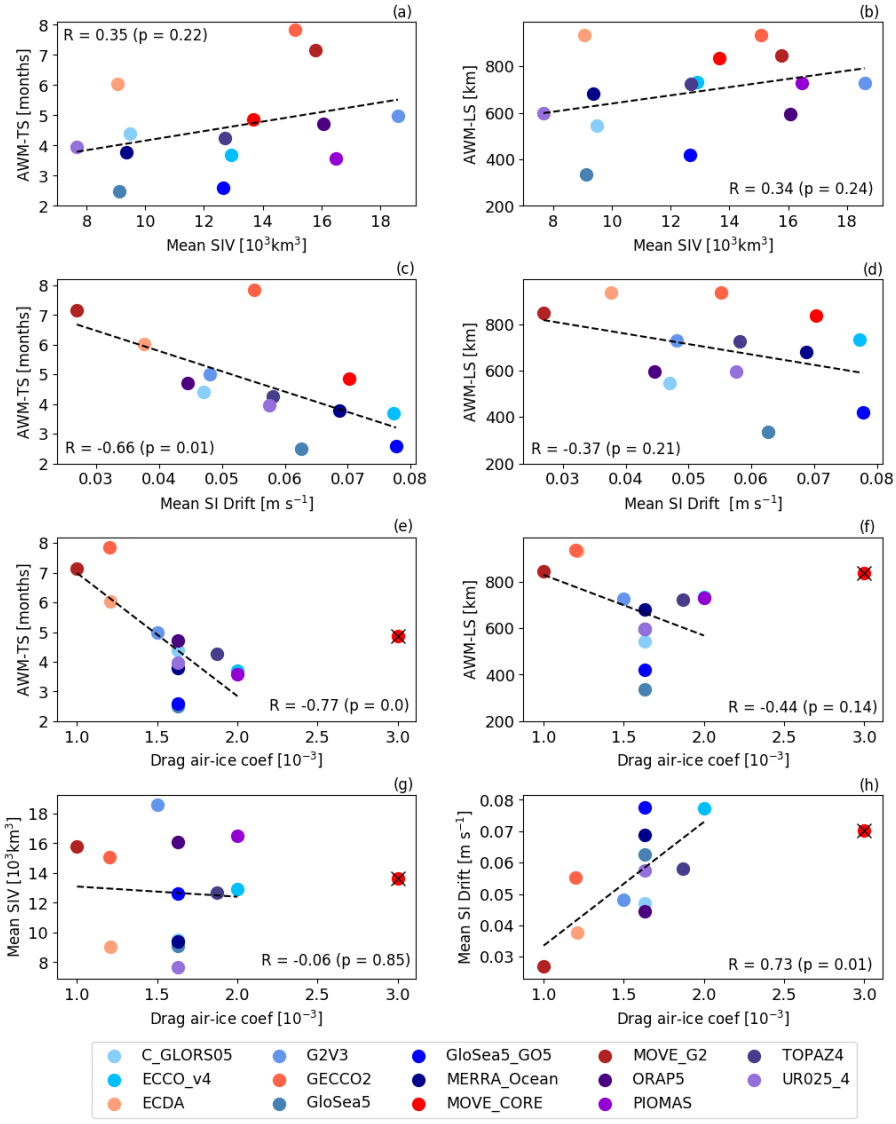

Figure 13. Scatter ~~plot of~~ plots showing how the ~~Average Weighted Mean~~ (~~AWM~~a, c, e) AWM-TS and (b, d, f) AWM-LS are related to the (a, b) mean state, (c, d) mean sea ice ~~velocity versus~~drift and (e, f) drag air-ice coefficient. (g) Relation between drag air-ice coefficient and mean state. (f) Relation between drag air-ice coefficient and mean sea ice drift. Black dashed lines indicate the ~~AWM time scale~~linear fit while the coefficient of correlation R (and its respective p-value) is also displayed in each panel. ~~Reanalyses~~ The black cross in panels e-h indicate that MOVE-CORE was not used in the respective regressions (black dashed lines), since this reanalysis adopts a much higher drag air-ice coefficient compared to the other thirteen reanalyses. Shades of blue and purple indicate the reanalyses which do assimilate sea ice data, while shades of red ~~highlight whether~~ indicate the ~~datasets were built with or~~ reanalyes without sea ice data assimilation~~, respectively~~.

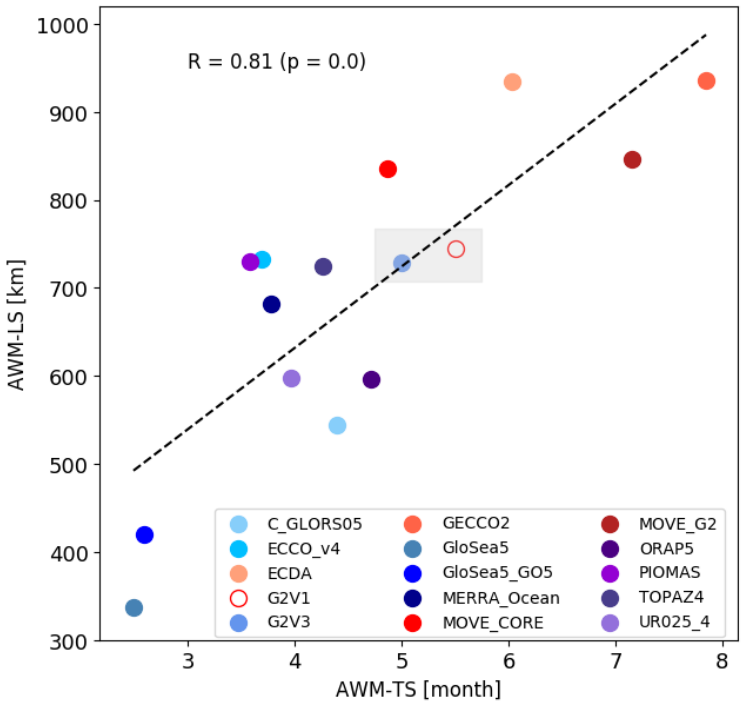

Figure 14. Scatter plot of the Average Weighted Mean (AWM) time scale *versus* the AWM length scale. The black dashed line indicates the linear regression between both parameters, ~~with colors representing~~ while the ~~time-mean sea ice volume. Reanalyses labeled~~ coefficient of correlation (R) is also displayed in ~~blue and red highlight whether~~ the ~~datasets were~~ plot. The gray rectangle compares the two GLORYS systems: G2V1, built ~~with or~~ without sea ice data assimilation, ~~respectively. The dashed gray lines have just an illustrative purpose, in order to highlight the longer time scale~~ and ~~larger length scale for systems without~~ G2V3, built with sea ice data assimilation~~.~~

sensitive to the reanalyses specifications and parameters (see Figs. 13c-f), the length scale is considerably more stable than the time scale so that it is a more reliable variable to be taken into account for deploying observing systems. For instance, the multiple linear regression model used by Lindsay and Zhang (2006), for determining optimal locations to predict sea ice extent from SIT, could be combined with the length scale information, avoiding that two or more stations placed into the same radius of correlation (length scale) are selected. Time scale would be more useful if used in combination with the knowledge of its variability. Further ~~specific~~ studies are required to evaluate the performance of time and length scales in providing support for the optimal design of observational programmes, though this work already shows some promising results in ~~such~~ that direction.

Appendix A: List of acronyms

This appendix displays all acronyms and their respective long names referred to in the text and used in the figures. The long names of some acronyms were previously omitted in order to preserve the readability of text, while others were already defined. All of them will be mentioned below so that the reader can easily consult their meaning at any time. As follows:

AER	Atmospheric and Environmental Research
AWI	Alfred Wegener Institute
AWM	Area-Weighted Mean
BIO-LS	Bedford Institute of Oceanography Lancaster Sound
BGEP	Beaufort Gyre Exploration Project
C-GLORS05	CMCC - Global Ocean Reanalysis System
CDR	Climate Data Record
CHK	Chuck Sea
CMCC	Centro Euro-Mediterraneo sui Cambiamenti Climatici
CryoSat	CRYOgenic SATellite
C_v	Coefficient of variation
DA_v	Data Assimilation
EBS	Eastern Beaufort Sea
ECCO-v4	Estimating the Circulation and Climate of the Ocean - version 4
ECDA	Ensemble Coupled Data Assimilation
ECMWF	European Centre for Medium-Range Weather Forecasts
ESA	European Space Agency
EVP	Elastic-Viscous-Plastic
GECCO2	German - Estimating the Circulation and Climate of the Ocean
GFDL	Geophysical Fluid Dynamics Laboratory
GLORYS	Global Ocean reanalysis and Simulation

GloSea5	Global Seasonal forecasting system
GloSea5-GO5	Global Seasonal forecasting system - Global Ocean 5.0
GMAO	Global Modeling and Assimilation Office
GS	Greenland Sea
GSFC	Goddard Space Flight Center
G2V3	GLORYS 2 - version 3
ICESat	Ice, Cloud,and land Elevation Satellite
IOS	Institute of Ocean Science
JMA	Japan Meteorological Agency
JPL	Jet Propulsion Laboratory
MERRA	Modern Era Retrospective-Analysis for Research and Applications
ARC MFC	Arctic Marine Forecasting Center
MIT	Massachusetts Institute of Technology
MOVE-CORE	Multivariate Ocean Variational Estimation - Coordinated Ocean-ice Reference Experiment
MOVE-G2	Multivariate Ocean Variational Estimation - Global version 2
MRI	Meteorological Research Institute
MRSS	Mean Residual Sum of Squares
NA_v	No data Assimilation
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NPEO	North Pole Environmental Observatory
ORA-IP	Ocean Reanalysis Intercomparison project
PIOMAS	Pan-Arctic Ice-Ocean Modeling and Assimilation System
R	Correlation Coefficient

RMSE	Root Mean Squared Error
SIT	Sea Ice Thickness
TP4	TOPAZ4
UK-Subs-AN	UK Navy Submarines - Analog
UK-Subs-DG	UK Navy Submarines - Digital
ULS	Upward-Looking Sonar
US-Subs-AN	US Navy Submarines - Analog
US-Subs-DG	US Navy Submarines - Digital
UR025-4	University of Reading, 1/4° deg - version 4
VP	Viscous-Plastic
YOPP	Year Of Polar Prediction

Appendix B: Tables

This appendix presents the comparison of all reanalysis products with all different observational datasets. Such comparison is based on three different metrics: Root Mean Square Error (RMSE), Mean Residual Sum of Squares (MRSS) and Correlation Coefficient (R).

Table B1. Root Mean Square Error (RMSE; in m) calculated between the SIT from observations and reanalyses. The RMSE was not calculated for the draft datasets (*).

(Reanalyses	NPEO*	BGEP*	IOS-EBS*	IOS-CHK*	SSUB-AN*	USSUB-DG*	UKSUB-AN*	UKSUB-DG*	AWI-GS*	BIO-LS*	Davis-St*	CanCoast*	IceBridge-V2	IceBridge-QL	ICESat1-G	CryoSat-AWI
C-GLORS05	–	–	–	–	–	–	–	–	–	–	–	–	0.7 (390)	–	0.9 (28413)	0.6 (29058)
ECCO-V4	–	–	–	–	–	–	–	–	–	–	–	–	0.9 (264)	–	1.1 (29147)	0.6 (5934)
ECDA	–	–	–	–	–	–	–	–	–	–	–	–	1.6 (926)	1.3 (439)	1.1 (29102)	1.0 (76909)
G2V3	–	–	–	–	–	–	–	–	–	–	–	–	1.0 (390)	–	1.5 (28413)	0.7 (29058)
GECCO2	–	–	–	–	–	–	–	–	–	–	–	–	1.2 (399)	–	0.9 (27560)	0.8 (30891)
GloSea5	–	–	–	–	–	–	–	–	–	–	–	–	1.5 (670)	1.3 (201)	1.1 (29071)	0.8 (53005)
GloSea5-GO5	–	–	–	–	–	–	–	–	–	–	–	–	1.0 (910)	1.0 (671)	1.0 (29071)	0.8 (101407)
MERRA-Ocean	–	–	–	–	–	–	–	–	–	–	–	–	1.2 (666)	0.9 (202)	1.1 (28937)	1.0 (45491)
MOVE-CORE	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.9 (24966)	–
MOVE-G2	–	–	–	–	–	–	–	–	–	–	–	–	1.4 (651)	1.4 (199)	1.1 (28413)	1.0 (52074)
ORAP5	–	–	–	–	–	–	–	–	–	–	–	–	0.9 (683)	0.9 (204)	0.9 (29366)	0.7 (53293)
PIOMAS	–	–	–	–	–	–	–	–	–	–	–	–	1.1 (933)	0.8 (860)	0.9 (29452)	0.6 (127270)
TOPAZ4	–	–	–	–	–	–	–	–	–	–	–	–	1.2 (914)	0.9 (442)	1.0 (28882)	0.8 (76046)
UR025-4	–	–	–	–	–	–	–	–	–	–	–	–	1.5 (258)	–	1.1 (29071)	0.7 (5922)

Table B2. Mean Residual Sum of Squares (MRSS; in m^2) estimated from the linear fit between reanalysis SIT (predictor) and observations (predict), either draft or SIT, parameters, respectively. Draft observational datasets are distinguished from the SIT datasets by the (*).

Reanalyses	NPEO*	BGEP*	IOS-EBS*	IOS-CHK*	SSUB-AN*	USSUB-DG*	UKSUB-AN*	UKSUB-DG*	AWI-GS*	BIO-LS*	Davis-St*	CanCoast*	IceBridge-V2	IceBridge-QL	ICESat1-G	CryoSat-AWI
C-GLORS05	0.3 (64)	0.1 (309)	0.5 (340)	0.1 (26)	0.1 (216)	0.3 (766)	1.5 (40)	–	0.3 (112)	–	0.2 (54)	0.2 (14)	0.4 (390)	–	0.7 (28413)	0.2 (29058)
ECCO-V4	0.4 (64)	0.1 (265)	0.4 (340)	0.1 (26)	0.2 (216)	0.4 (701)	–	–	0.2 (110)	0.4 (36)	0.2 (48)	0.2 (195)	0.5 (264)	–	0.8 (29147)	0.2 (5934)
ECDA	0.4 (64)	0.1 (369)	0.5 (382)	0.1 (26)	0.4 (794)	0.4 (1000	2.3 (149)	0.1 (27)	0.3 (131)	0.4 (36)	0.3 (53)	0.2 (2664)	0.9 (926)	0.7 (439)	0.8 (29102)	0.4 (76909)
G2V3	0.6 (64)	0.2 (310)	0.4 (324)	0.2 (26)	0.7 (216)	0.4 (637)	–	–	0.2 (108)	–	0.2 (54)	–	0.5 (390)	–	0.6 (28413)	0.3 (29058)
GECCO2	0.6 (64)	0.2 (308)	0.6 (357)	0.2 (23)	0.9 (840)	0.8 (100)	2.0 (149)	0.2 (27)	0.3 (115)	0.4 (36)	0.4 (46)	0.3 (2864)	0.5 (399)	–	0.6 (27560)	0.2 (30891)
GloSea5	0.6 (64)	0.2 (348)	0.4 (382)	0.1 (26)	0.5 (253)	0.4 (816)	0.9 (40)	–	0.3 (134)	–	0.1 (67)	0.2 (399)	0.7 (670)	0.6 (201)	0.8 (29071)	0.4 (53005)
GloSea5-GO5	0.5 (64)	0.2 (371)	0.4 (382)	0.1 (26)	0.5 (253)	0.4 (777)	0.8 (40)	–	0.3 (134)	–	0.1 (67)	0.2 (376)	0.7 (910)	0.8 (671)	0.8 (29071)	0.4 (101407)
MERRA-Ocean	0.4 (64)	0.3 (328)	0.5 (379)	0.3 (26)	0.8 (540)	0.8 (1001)	1.9 (149)	–	0.2 (133)	0.3 (36)	0.1 (66)	0.3 (857)	0.5 (666)	0.4 (202)	0.7 (28937)	0.4 (45491)
MOVE-CORE	0.1 (41)	0.2 (167)	0.4 (379)	0.2 (26)	0.4 (843)	0.5 (1001)	1.9 (149)	0.2 (27)	0.2 (129)	–	0.2 (39)	0.2 (412)	–	–	0.7 (24966)	–
MOVE-G2	0.7 (64)	0.4 (348)	0.5 (342)	0.1 (26)	0.4 (216)	0.4 (638)	–	–	0.3 (100)	–	0.2 (48)	–	0.7 (651)	0.5 (199)	0.8 (28413)	0.5 (52074)
ORAP5	0.3 (64)	0.2 (340)	0.4 (339)	0.1 (26)	0.8 (646)	0.7 (1000)	1.3 (149)	–	0.2 (112)	0.3 (35)	0.2 (50)	0.3 (1840)	0.6 (683)	0.5 (204)	0.7 (29366)	0.3 (53293)
PIOMAS	0.5 (64)	0.1 (371)	0.4 (382)	0.1 (26)	0.4 (646)	0.5 (1001)	2.0 (149)	–	0.3 (134)	0.3 (36)	0.2 (67)	0.3 (2789)	0.8 (933)	0.6 (860)	0.7 (29452)	0.3 (127270)
TOPAZ4	0.2 (64)	0.2 (371)	0.4 (345)	0.2 (26)	0.2 (216)	0.3 (638)	–	–	0.3 (115)	0.3 (36)	0.1 (67)	0.1 (121)	0.7 (914)	0.6 (442)	0.6 (28882)	0.3 (76046)
UR025-4	0.5 (64)	0.1 (273)	0.4 (345)	0.1 (26)	0.3 (216)	0.3 (637)	–	–	0.3 (115)	–	0.1 (67)	0.2 (241)	0.4 (258)	–	0.7 (29071)	0.2 (5922)

Table B3. Correlation coefficient (R) estimated between the SIT from reanalysis and observations, either draft or SIT. Draft observational datasets are distinguished from the SIT datasets by the (*).

Reanalyses	NPEO*	BGEP*	IOS-EBS*	IOS-CHK*	SSUB-AN*	USSUB-DG*	UKSUB-AN*	UKSUB-DG*	AWI-GS*	BIO-LS*	Davis-St*	CanCoast*	IceBridge-V2	IceBridge-QL	ICESat1-G	CryoSat-AWI
C-GLORS05	0.73 (64)	0.83 (309)	0.49 (340)	0.95 (26)	0.92 (216)	0.74 (766)	0.35 (40)	–	0.41 (112)	–	0.80 (54)	-0.01 (14)	0.47 (390)	–	0.49 (28413)	0.73 (29058)
ECCO-V4	0.65 (64)	0.83 (265)	0.54 (340)	0.96 (26)	0.87 (216)	0.65 (701)	–	–	0.55 (110)	-0.47 (36)	0.82 (48)	0.28 (195)	0.31 (264)	–	0.33 (29147)	0.68 (5934)
ECDA	0.68 (64)	0.86 (369)	0.52 (382)	0.90 (26)	0.81 (794)	0.73 (1000)	0.46 (149)	-0.65 (27)	0.51 (131)	0.27 (36)	0.65 (53)	0.62 (2664)	0.20 (926)	0.43 (439)	0.29 (29102)	0.48 (76909)
G2V3	0.35 (64)	0.73 (310)	0.65 (324)	0.89 (26)	0.44 (216)	0.56 (637)	–	–	0.61 (108)	–	0.80 (54)	–	0.34 (390)	–	0.52 (28413)	0.62 (29058)
GECCO2	0.40 (64)	0.69 (308)	0.34 (357)	0.79 (23)	0.36 (840)	0.39 (1001)	0.56 (149)	-0.28 (27)	0.46 (115)	-0.39 (36)	0.48 (46)	0.17 (2864)	0.45 (399)	–	0.58 (27560)	0.73 (30891)
GloSea5	0.32 (64)	0.75 (348)	0.61 (382)	0.95 (26)	0.55 (253)	0.59 (816)	0.69 (40)	–	0.56 (134)	–	0.87 (67)	0.27 (399)	0.40 (670)	0.14 (201)	0.28 (29071)	0.58 (53005)
GloSea5-GO5	0.54 (64)	0.81 (371)	0.63 (382)	0.95 (26)	0.63 (253)	0.59 (777)	0.70 (40)	–	0.59 (134)	–	0.88 (67)	0.21 (376)	0.53 (910)	0.53 (671)	0.30 (29071)	0.58 (101407)
MERRA-Ocean	0.63 (64)	0.65 (328)	0.56 (379)	0.79 (26)	0.52 (540)	0.43 (1001)	0.60 (149)	–	0.65 (133)	0.66 (36)	0.86 (66)	0.22 (857)	0.56 (666)	0.57 (202)	0.50 (28937)	0.54 (45491)
MOVE-CORE	0.80 (41)	0.75 (167)	0.65 (379)	0.83 (26)	0.78 (843)	0.73 (1001)	0.61 (149)	0.33 (27)	0.65 (129)	–	0.71 (39)	0.52 (412)	–	–	0.47 (24966)	–
MOVE-G2	-0.16 (64)	0.48 (348)	0.52 (342)	0.93 (26)	0.68 (216)	0.63 (638)	–	–	0.29 (100)	–	0.76 (48)	–	-0.27 (651)	-0.41 (199)	0.16 (28413)	0.38 (52074)
ORAP5	0.79 (64)	0.81 (340)	0.60 (339)	0.90 (26)	0.50 (646)	0.52 (1000)	0.75 (149)	–	0.52 (112)	0.64 (35)	0.77 (50)	0.33 (1840)	0.49 (683)	0.43 (204)	0.49 (29366)	0.63 (53293)
PIOMAS	0.48 (64)	0.87 (371)	0.70 (382)	0.94 (26)	0.79 (646)	0.71 (1001)	0.58 (149)	–	0.60 (134)	0.57 (36)	0.74 (67)	0.11 (2789)	0.29 (933)	0.68 (860)	0.48 (29452)	0.71 (127270)
TOPAZ4	0.81 (64)	0.81 (371)	0.67 (345)	0.82 (26)	0.86 (216)	0.66 (638)	–	–	0.52 (115)	0.53 (36)	0.91 (67)	0.50 (121)	0.39 (914)	0.54 (442)	0.57 (28882)	0.63 (76046)
UR025-4	0.58 (64)	0.84 (273)	0.62 (345)	0.95 (26)	0.82 (216)	0.75 (637)	–	–	0.61 (115)	–	0.86 (67)	0.41 (241)	0.41 (258)	–	0.40 (29071)	0.74 (5922)

Competing interests. No competing interests are present

Acknowledgements. The work presented in this paper has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 727862: APPLICATE project (Advanced prediction in Polar regions and beyond). David Docquier is funded by the EU Horizon 2020 PRIMAVERA project, grant agreement no. 641727. We thank two anonymous reviewers for their constructive suggestions and criticism. We thank Axel Schweiger for making the observational dataset available by means of the Unified Sea Ice Thickness Climate Data Record (http://psc.apl.uw.edu/sea_ice_cdr) and also for kindly make clear some aspects of the data. The BGEP data used in Fig. 2 were collected and made available by the Beaufort Gyre Exploration Project based at the Woods Hole Oceanographic Institution (http://www.whoi.edu/beaufortgyre). The python wavelet software is provided by Evgeniya Predybaylo based on Torrence and Compo (1998) and is available at

URL: http://atoc.colorado.edu/research/wavelets/

References

Anisimov, O. A., Vaughan, D. G., Callaghan, T. V., Furgal, C., Marchant, H., Prowse, T. D., Vilhjálmsson, H., and Walsh, J. E.: Polar regions (Arctic and Antarctic). Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 2007.

Balmaseda, M. A., Hernandez, F., Storto, A., Palmer, M. D., Alves, O., Shi, L., Smith, G. C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y.-S., Chepurin, G. A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Meyssignac, B., Mogensen, K., Parent, L., Peterson, K. A., Tang, Y. M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J.-F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., Wilmer-Becker, K., and Gaillard, F.: The Ocean Reanalyses Intercomparison Project (ORA-IP), J. Oper. Oceanogr., 8, s80–s97, doi:10.1080/1755876X.2015.1022329, 2015.

Blanchard-Wrigglesworth, E. and Bitz, C.: Characteristics of Arctic Sea-Ice Thickness Variability in GCMs, J. Clim., 27, 8244–8258, doi:10.1175/JCLI-D-14-00345.1, 2014.

Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C., and DeWeaver, E.: Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations, J. Clim., 24, 231–250, doi:10.1175/2010JCLI3775.1, 2011.

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A., and Storkey, D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, Geosci. Model Dev., 7, 2613–2638, doi:10.5194/gmd-7-2613-2014, 2014.

Budyko, M. I.: The effect of solar radiation variations on the climate of the Earth, Tellus, 21, 611–619, doi:10.1111/j.2153-3490.1969.tb00466.x, 1969.

Chang, Y. S., Zhang, S., Rosati, A., Delworth, T. L., and Stern, W. F.: An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation, Clim. Dyn., 40, 775–803, doi:10.1007/s00382-012-1412-2, 2013.

Chevallier, M. and Salas-Mélia, D.: The role of sea ice thickness distribution in the Arctic sea ice potential predictability: a diagnostic approach with a coupled GCM, J. Clim., 25, 3025–3038, doi:10.1175/JCLI-D-11-00209.1, 2012.

Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Zuo, H., Balmaseda, M., Chang, Y.-S., Ferry, N., Garric, G., Haines, K., Keeley, S., Kovach, R. M., Kuragano, T., Masina, S., Tang, Y., Tsujino, H., and Wang, X.: Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORAIP project, Clim. Dyn., 19, 1107–1136, doi:110.1007/s00382-016-2985-y, 2017.

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Boning, C., Bozec, A., Canuto, V., Cassou, C., Chassignet, E., Coward, A. C., Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., Fogli, P. G., Forget, G., Fujii, Y., Griffies, S. M., Gusev, A., Heimbach, P., Howard, A., Jung, T., Kelley, M., Large, W. G., Leboissetier, A., Lu, L., Madec, G., Marsland, S. J., , Masina, S., Navarra, A., Nurser, A. J. G., Pirani, A., Salas-Melia, D., Samuels, B. L., Scheinert, M., Sidorenko, D., Treguier, A. M., Tsujino, H., Uotila, P., Valcke, S., Voldoire, A., and Wang, Q.: North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states, Ocean Model., 73, 76–107, doi:10.1016/j.ocemod.2013.10.005, 2014.

Day, J. J., Tietsche, S., and Hawkins, E.: Pan-Arctic and Regional Sea Ice Predictability: Initialization Month Dependence, J. Clim., 27, 4371–4390, doi:10.1175/JCLI-D-13-00614.1, 2014.

Drange, H. and Simonsen, K.: Formulation of air-sea fluxes in the ESOP2 version of MICOM, Technical Report No. 125, Tech. rep., Nansen Environmental and Remote Sensing Center, 1996.

Drijfhout, S.: Competition between global warming and an abrupt collapse of the AMOC in Earth's energy imbalance, Sci. Rep., 5, 1–12, doi:10.1038/srep14877, 2015.

Drucker, R., Martin, S., and Moritz, R.: Observations of ice thickness and frazil ice in the St. Lawrence Island polynya from satellite imagery, upward looking sonar, and salinity/temperature moorings, J. Geophys. Res., 108, 18–1–18–18, doi:10.1029/2001JC001213, 2003.

Ferry, N., Parent, L., Garric, G., Barnier, B., and Jourdain, N. C.: Mercator global Eddy permitting ocean reanalysis GLORYS1V1: description and results, Mercator-Ocean Q. Newslett., 36, 15–27, 2010.

Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C.: ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation, Geosci. Model Dev., 8, 3071–3104, doi:10.5194/gmd-8-3071-2015, 2015.

Gerdes, R. and Köberle, C.: Comparison of Arctic sea ice thickness variability in IPCC Climate of the 20th Century experiments and in ocean–sea ice hindcasts, J. Geophys. Res., 112, C04S13, doi:10.1029/2006JC003616, 2007.

Gleick, P. H.: The implications of global climatic changes for international security, Climatic Change, 15, 309–325, doi:10.1007/BF00138857, 1989.

Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., Fučkar, N. S., Germe, A., Hawkins, E., Keeley, S., Koenigk, T., Salas y Mélia, D., and Tietsche, S.: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales, Q. J. R. Meteorol. Soc., 142, 546–561, doi:10.1002/qj.2401, 2016.

Handorf, U.: Tourism booms as the Arctic melts. A critical approach of polar tourism, GRIN Verlag, Munich, 2011.

Hansen, J., Sato, M., Hearty, P., Ruedy, R., Kelley, M., Masson-Delmotte, V., Russell, G., Tselioudis, G., Cao, J., Rignot, E., Velicogna, I., Tormey, B., Donovan, B., Kandiano, E., von Schuckmann, K., Kharecha, P., Legrande, A. N., Bauer, M., and Lo, K. W.: Ice melt, sea level rise and superstorms: evidence from paleoclimate data, climate modeling, and modern observations that 2°C global warming could be dangerous, Atmos. Chem. Phys., 16, 3761–3812, doi:10.5194/acp-16-3761-2016, 2016.

Harms, S., Fahrbach, E., and Strass, V.: Sea ice transports in the Weddell Sea, J. Geophys. Res., 106, 9057–9073, doi:10.1029/1999JC000027, 2001.

Hibler, W. D.: A dynamic thermodynamic sea ice model, J. Phys. Oceanogr., 9, 815–846, doi:10.1175/1520-0485(1979)009<0815:ADTSIM>2.0.CO;2, 1979.

Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in coupled models, Clim. Dyn., 21, 221–232, doi:10.1007/s00382-003-0332-6, 2003.

Holland, M. M., Bailey, D. A., and Vavrus, S.: Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3, Clim. Dyn., 36, 1239–1253, doi:10.1007/s00382-010-0792-4, 2011.

Hunke, E. C. and Dukowicz, J. K.: An elastic-viscous-plastic model for sea ice dynamics, J. Phys. Oceanogr., 27, 1849–1867, doi:10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2, 1997.

Jung, T., Gordon, N. D., Bauer, P., Bromwich, D. H., Chevallier, M., Day, J. J., Dawson, J., Doblas-Reyes, F., Fairall, C., amd M. Holland, H. F. G., Inoue, J., Iversen, T., Klebe, S., Lemke, P., Losch, M., Makshtas, A., Mills, B., Nurmi, P., Perovich, D., Reid, P., Renfrew, I. A., Smith, G., Svensson, G., Tolstykh, M., and Yang, Q.: Advancing polar prediction capabilities on daily to seasonal time scales, Bull. Am. Meteorol. Soc., 97, 1631–1647, doi:10.1175/BAMS-D-14-00246.1, 2016.

Krishfield, R. A., Proshutinsky, A., Tateyama, K., Williams, W. J., Carmack, E. C., McLaughlin, F. A., and Timmermans, M. L.: Deterioration of perennial sea ice in the Beaufort Gyre from 2003 to 2012 and its impact on the oceanic freshwater cycle, J. Geophys. Res.: Oceans, 119, 1271–1305, doi:10.1002/2013JC008999, 2013.

Krupnik, I. and Jolly, D.: Earth is Faster Now: Indigenous Observations of Arctic Environmental Change, Arctic Research Consortium of the United States, Fairbanks, Alaska, 2002.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data, The Cryosphere, 7, 1035–1056, doi:10.5194/tc-7-1035-2013, 2013.

Köhl, A.: Evaluation of the GECCO2 Ocean Synthesis: Transports of Volume, Heat and Freshwater in the Atlantic, Q. J. R. Met. Soc., 141, 166–181, doi:10.1002/qj.2347, 2015.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, doi:10.5194/tc-9-269-2015, 2015.

Lindsay, R. W.: A new sea ice thickness climate data record, EOS, 91, 405–406, doi:10.1029/2010EO440001, 2010.

Lindsay, R. W. and Zhang, J.: Arctic Ocean Ice Thickness: Modes of Variability and the Best Locations from Which to Monitor Them, J. Phys. Oceanogr., 36, 496–506, doi:10.1175/JPO2861.1, 2006.

Lindsay, R. W., Zhang, J., Schweiger, A. J., and Steele, M. A.: Seasonal predictions of ice extent in the Arctic Ocean, J. Geophys. Res., 113, C02 023, doi:10.1029/2007JC004259, 2008.

Lindstad, H., Bright, R. M., and Strømmanb, A. H.: Economic savings linked to future Arctic shipping trade are at odds with climate change mitigation, Transp. Policy, 45, 24–34, doi:10.1016/j.tranpol.2015.09.002, 2016.

Manabe, S. and Stouffer, R. J.: Sensitivity of a global climate model to an increase of CO_2 in the atmosphere, J. Geophys. Res., 85, 5529–5554, doi:10.1029/JC085iC10p05529, 1980a.

Manabe, S. and Stouffer, R. J.: Sensitivity of a global climate model to an increase of CO_2 concentration in the atmosphere, J. Geophys. Res., 85, 5529–5554, doi:10.1029/JC085iC10p05529, 1980b.

Massonnet, F., Vancoppenolle, M., Goosse, H., Docquier, D., Fichefet, T., and Blanchard-Wrigglesworth, E.: Arctic sea-ice change tied to its mean state through thermodynamic processes, Nat. Clim. Change, 8, 599–603, doi:10.1038/s41558-018-0204-z, 2018.

Maykut, G. A.: Large-scale heat exchange and ice production in the central Arctic, J. Geophys. Res., 87, 7971–7984, doi:10.1029/JC087iC10p07971, 1982.

Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., Hyder, P., Siddorn, J., and Sinha, B.: GO5.0: the joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications, Geosci. Model Dev., 7, 1069–2014, doi:10.5194/gmd-7-1069-2014, 2014.

Mellor, G. L. and Kantha, L.: An ice-ocean coupled model, J. Geophys. Res., 94 (C8), 10 937–10 954, 1989.

Nelson, F. E., Anisimov, O. A., and Shiklomanov, N. I.: Climate Change and Hazard Zonation in the Circum-Arctic Permafrost Regions, Natural Hazards, 26, 203–225, doi:10.1023/A:1015612918401, 2002.

Nuttall, M., Berkes, F., Forbes, B., Kofinas, G., Vlassova, T., and Wenzel, G.: Arctic Climate Impact Assessment, Cambridge University Press, Cambridge, 2005.

Pettipas, R., Hamilton, J., and Prinsenberg, S.: Moored current meter and CTD observations from Barrow Strait, 2003-2004, Can. Data Rep. Hydrogr. Ocean Sci., 173, 134 p, 2008.

Prinsenberg, S. and Pettipas, R.: Ice and ocean mooring data statistics from Barrow Strait, the central section of the NW Passage in the Canadian Arctic Archipelago, Int. J. Offshore Pol. Eng., 18, 277–281, 2008.

Prinsenberg, S., Hamilton, J., Peterson, I., and Pettipas, R.: Influence of climate change on the changing Arctic and Sub-Arctic conditions, J. Nihoul and A. Kostianoy, eds, Springer, Dordrecht, 2009.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M.: Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, The Cryosphere, 8, 1607–1622, doi:10.5194/tc-8-1607-2014, 2014.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Clim., 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Rothrock, D. A.: The energetics of the plastic deformation of pack ice by ridging, J. Geophys. Res., 80, 4514–4519, doi:10.1029/JC080i033p04514, 1975.

Rothrock, D. A. and Wensnahan, M.: Global atmospheric forcing data for Arctic iceocean modeling, J. Geophys. Res., 112, C04S14, doi:10.1029/2006JC003640, 2007a.

Rothrock, D. A. and Wensnahan, M.: The Accuracy of Sea Ice Drafts Measured from U.S. Navy Submarines, J. Atmos. Oceanic Technol., 24, 1936–1949, doi:10.1175/JTECH2097.1, 2007b.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic, Ocean Sci., 8, 633–656, doi:10.5194/os-8-633-2012, 2012.

Serreze, M. C., Barrett, A. P., Stroeve, J. C., Kindig, D. N., and Holland, M. M.: The emergence of surface-based Arctic amplification, The Cryosphere, 3, 11–19, doi:10.5194/tc-3-11-2009, 2009.

Storto, A., Masina, S., and Dobricic, S.: Estimation and Impact of Nonuniform Horizontal Correlation Length Scales for Global Ocean Physical Analyses, J. Atmos. Ocean. Technol., 31, 2330–2349, doi:10.1175/JTECH-D-14-00042.1, 2014.

Stroeve, J. C., Hamilton, L., Blitz, C. M., and Blanchard-Wrigglesworth, E.: Predicting September sea ice: Ensemble skill of the SEARCH sea ice outlook 2008–2013, Geophys. Res. Lett., 41, 2411–2418, doi:10.1002/2014GL059388, 2014.

Sévellec, F., Fedorov, A. V., and Liu, W.: Arctic sea-ice decline weakens the Atlantic Meridional Overturning Circulation, Nat. Clim. Change, 7, 604–610, doi:10.1038/nclimate3353, 2017.

Tandon, N. F., Kushner, P. J., Docquier, D., Wettstein, J. J., and Li, C.: Reassessing Sea Ice Drift and its Relationship to LongTerm Arctic Sea Ice Loss in Coupled Climate Models, J. Geophys. Res., doi:10.1029/2017JC013697, 2018.

Tietsche, S., Balmaseda, M. A., Zuo, H., and Mogensen, K.: Arctic sea ice in the global eddy-permitting ocean reanalysis ORAP5, ECMWF technical memorandum 736, 49, 775–789, doi:10.1007/s00382-015-2673-3, 2017.

Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79, 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2, 1998.

Toyoda, T., Fujii, Y., Yasuda, T., Usui, N., Iwao, T., Kuragano, T., and Kamachi, M.: Improved Analysis of Seasonal-Interannual Fields Using a Global Ocean Data Assimilation System, Theor. Appl. Mech. Jpn., 61, 31–48, doi:10.11345/nctam.61.31, 2013.

Tucker III, W. B., Weatherly, J. W., Eppler, D. T., Farmer, D., and Bentley, D. L.: Evidence for the rapid thinning of sea ice in the western Arctic Ocean at the end of the 1980s, Geophys. Res. Let., 28, 2851–2854, doi:10.1029/2001GL012967, 2001.

Ungermann, M., Tremblay, L. B., Martin, T., and Losch, M.: Impact of the ice strength formulation on the performance of a sea ice thickness distribution model in the Arctic, J. Geophys. Res., 122, 2090–2107, doi:10.1002/2016JC012128, 2017.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V. S., Marnela, M., Massonnet, F., Mignac, D., Peterson, K. A., Sadikni, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J., and Zhang, Z.: An assessment of ten ocean reanalyses in the polar regions, Clim. Dyn., pp. 1–38, doi:10.1007/s00382-018-4242-z, 2018.

Valdivieso, M., Haines, K., Zuo, H., and Lea, D.: Freshwater and heat transports from global ocean synthesis, J. Geophys. Res., 119, 394–409, doi:10.1002/2013JC009357, 2014.

Wadhams, P.: Arctic sea ice morphology and its measurement. Arctic Technology and Policy. I. Dyer and C. Chryssostomidis, eds., Hemisphere Publishing Corp., Washington, D.C., 1984.

Wadhams, P. and Horne, R. J.: An analysis of ice profiles obtained by submarine in the Beaufort Sea, J. Glaciol., 25, 401–424, doi:10.3189/S0022143000015264, 1980.

Walsh, J. E.: Intensified warming of the Arctic: Causes and impacts on middle latitudes, Global Planet. Change, 117, 52–63, doi:10.1016/j.gloplacha.2014.03.003, 2014.

Wensnahan, M. and Rothrock, D. A.: Sea-ice draft from submarine-based sonar: Establishing a consistent record from analog and digitally recorded data, Geophys. Res. Lett., 32(L11502), doi:10:1029/2005GL022507, 2005.

Xie, J., Bertino, L., Counillon, F., Lisæter, K., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991-2013, Ocean Sci., 13(1), 123–144, doi:10.5194/os-13-123-2017, 2017.

Zhang, J. L. and Rothrock, D. A.: Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, Mon. Weather Rev., 131, 845–861, 2003.

Zhang, S., Harrison, M. J., Rosati, A., and Wittenberg, A.: System Design and Evaluation of Coupled Ensemble Data Assimilation for Global Oceanic Climate Studies, Mon. Weather Rev., 135, 3541–3564, doi:0.1175/MWR3466.1, 2013.

Zuo, H., Balmaseda, M. A., and Mogensen, K.: The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice reanalysis ORAP5. Part 1: implementation, ECMWF technical memorandum 736, 2015.

Zwally, H. J., Yi, D., Kwok, R., and Zhao, Y.: ICESat Measurements of Sea Ice Freeboard and Estimates of Sea Ice Thickness in the Weddell Sea, J. Geophys. Res., 113 (C02515), doi:10.1029/2007JC004284, 2008.