

Response to reviewer 1 (Jean-Francois Lemieux)

2nd review of Presentation and evaluation of the Arctic sea ice forecasting system neXtSIM-F by Williams et al.

Thank you, Jean-Francois, for your review of our paper and your useful suggestions, that helped improve the paper a lot.

The authors have addressed my main concerns. I recognize the huge amount of work that was put in order to evaluate the performances of nextSIM-F and to improve the paper. I just have a few additional minor comments to improve the clarity of the text. I recommend that the manuscript could be published once these minor comments will have been addressed.

1. Minor comments

1) You often refer to viscoplastic models. To be consistent with the literature, I suggest you use the expression viscous-plastic.

We have changed this term.

2) p.1 line 12: ...greatly improveS...

Corrected.

3) p.3 line 5: Parenthesis are missing for the references.

Corrected.

4) p.2 line 22: OSISAF is not assimilated in RIPS. In the paper (Lemieux et al. 2016) it is written:

"Retrievals of sea ice concentration from passive microwave (Special Sensor Microwave Imager, SSM/I; Special Sensor Microwave Imager/Sounder, SSMIS) and advanced scatterometer data, and manually produced sea ice charts from the Canadian Ice Service (CIS) are assimilated by the 3D-Var system."

Corrected this description of RIPS.

5) p.2 footnotes: The "S" in RIPS and RIOPS stands for System...not Service.

Corrected.

6) p.3 line 12: ...enterED into operations in...

Corrected, since this date is now in the past.

7) p.3 line 23: I would be surprised that mariners currently plan their operations based on forecasts of leads. The only thing I heard of is that the US navy is usingNRL's lead forecasts for their submarines...

Changed "More pertinently in a forecast context, they are also highly relevant for navigation." to "While the precise forecast of individual leads is very challenging (and probably requiring assimilation of quite specific data like SAR-derived deformation, Korosov and Rampal, 2017), reliable information of this sort would be very useful for icebreakers that wish to reduce fuel consumption or submarines wishing to surface." (Thanks for the idea about submarines.)

8) p.5 line 1: Please rephrase.

Changed "In order to combine the advantages of these products we generated a blended product that was used both for assimilation during the forecasts." to "In order to combine the advantages of these products we generated a blended product that was used for assimilation during the forecasts."

9) p.5 line 28: Please rephrase (I don't understand "at the ice velocity").

Clarified this procedure.

10) p.6 line 23: Rampal updated the paper of Rampal...this sounds weird. Something like: "Following the work of Rampal 2016, Rampal 2019 evaluated..."

Actually deleted this sentence.

11) p.6 line 24: "Addition" is repeated twice...please rephrase.

Rephrased.

12) p.7 line 2: k_1 should be unitless while k_2 should be in Nm^{-3} . Here is an important comment: The suggested (and optimized value) of k_1 is 8 not 10. This clearly explains why nextSIM overestimates landfast ice in the Laptev and East Siberian Seas. As nextSIM tends to simulate ice a bit too thick you could even use $k_1 = 7$.

Thanks - we have noted this comment. In the latest round of tuning $k_1=5$ improved things a lot, and we will try $k_1=7$ also.

13) p.8 section 3.2: It should be clearer that this is done only at the beginning of the free run and before the first forecast. Some people might be confused that this is done to initialize all the forecasts.

We have clarified this.

14) I am confused with all the different variables that are used for concentration (c_t , c_U , c_y , c_F , c_B , etc.). I am sure you can simplify this. For example, if I am right, $c_B = c_O$ and $c_t = c_F$.

We have endeavoured to simplify these variables this time.

15) p.9 line 19 (and at other places): It is not clear what you mean by ice mask.

We have added the definition the first time it occurs.

16) p.10 line 26: You mean "lower" not "greater"?

Yes - changed this sentence

17) p.11 line 4: Replace "reflected" by "reflecting"

Changed to 'reflecting'

18) Fig. 3: there is no shaded area as mentioned in the caption.

Corrected this caption

219) p.14 line 1: You refer to Jan-Feb in Figure 4 but it does not exist. Do you want to write "not shown"? Same idea for Nov-Dec.

This is correct - changed this sentence

20) Fig. 5: the text is very small. Please improve this.

Increased the font size for this figure

21) p.22 line 13...: You use too much "we" in this paragraph. For example: "...we are systematically lower in concentration...". Replace by: "The forecasts are systematically...". Same idea for "...we score..."

Changed to your suggestion

Congratulations for your paper!

Jean-François Lemieux

Thanks!

Response to reviewer 2 (anonymous)

Thank you, Reviewer 2, for your extremely thorough review of our paper and your useful suggestions, that helped improve the paper a lot.

Overview

I am pleased to see that the authors have addressed one of my main concerns from the last review by extending the hindcast runs, and associated assessment, out to more than 1 year. Given that things are worse in the second winter than the first, I feel it would be better to have run the system for even longer, but 20 months is much better than what we had before. Glad you feel this is better - we are constrained in going back further by the version of the atmospheric forcing product which began in the summer of 2018 (we initialise our forecast early in November of that year, as soon as the CS2-SMOS settles down).

However, I find myself still rather frustrated with the presentation of this manuscript. The wording used is still very informal/unscientific and the arguments are not laid out in a manner that is conducive to transfer information to the reader – i.e., detailing exactly what has been done, why certain decisions were made, etc.. Moreover, the figures are quite difficult to understand. Several of the newer figures use very small text and/or very small coloured dots, that make them almost illegible. Furthermore, nearly all of the figures have captions that are either wrong or don't contain the required detail, and all figures suffer from a lack of annotation/labelling or panel subtitling, that would allow the reader to quickly ascertain exactly what is plotted where.

We have endeavoured to improve the language, figures (particularly by increasing font size and adding some extra annotations) and captions.

Finally, after reading this through again, I'm afraid to say that I now find myself even more confused about exactly what is done in the neXtSIM-F operational forecasting system (and why!). I don't find Figure 1 very informative in this regard without the appropriate explanation. For example, I am not sure what the difference between the "initialization" and "assimilation" boxes – when most operational centres would actually class the assimilation as the initialisation (i.e., its only job is to get the forecast starting from the best possible initial conditions!). I am also left wondering why the assimilation seems to be being applied during the forecasts (P19 L29-31) and why it seems to make little difference either way. Given this last point, I am further left wondering what the analysis is about (i.e., what is analysed, how is it used, how does it differ from the 1-day forecasts that contain assimilation?).

I think this issue can be fairly easily fixed - the authors need to take some time to carefully explain the top-level structure of the forecasting system. I would recommend doing this right at the start of Section 3 where Fig 1 is introduced (i.e., before the model is discussed in Section 3.1).

We have endeavoured to improve the description and have also simplified Fig 1 to show only what happens every day.

Also I find that things are further confused by the dual discussion of MEB and BBM versions of neXtSIM but only the latter is evaluated here. In some places (e.g. Appendix A1) this feels like a CMEMS report about the operational system (in particular given that December 2020

is in the past)! Given that in this paper you are solely showing results from the BBM version of neXtSIM, much of the MEB stuff is not relevant - except to show the differences in the new (current?) system you are documenting here. So I feel it would make sense to tone the MEB descriptions down and just to focus on the new/current operational system.

This discussion/description is now much reduced.

General comments

You need to be careful when talking about other systems all using the VP rheology because many readers will likely ask "what about EVP?". Many will not know how similar these are so it might be best to elaborate a little and say "based upon the viscous-plastic (E)VP formulation" or "the VP family of rheologies" or something. Also some might ask "what about the EAP" rheology? I think that RASM are using EAP for their coupled seasonal forecasts now(?). Maybe you could avoid that potential conflict by better quantifying the area you are talking about by stating "short-range analysis and forecasting system" instead of just "forecasting system"?

We have mentioned the EVP now as a method of solving the VP equations

NB. page and line numbers used below correspond to those used in the tracked-changes version of the document. I now notice that these are somewhat odd with some pages having some line repetition

Specific comments

P2 L15: "while Hunke et al. (2020, Table 1) give a comprehensive list of modelling systems that include sea ice": I would be a little careful here because the Hunke et al. paper is actually about the use of climate model sea ice modelling formulations for operational forecasting. Therefore, it does not aim to give an overview of all operational forecasting systems (certainly not all "modelling systems" as you claim!), instead only those using the classic AIDJEX continuum model formulation (CICE, LIM/SI3, MITgcm).

While this paper was not explicitly aiming to give an overview of forecasting systems, it did nevertheless give a long list of them in table 1 which made it a convenient paper to refer to. We changed this sentence to:

"Tonani et al. (2015) give a good overview of the 2015 status of operational forecasting (here we take "operational forecasts" to refer to those with forecast horizons of about a week), while Hunke et al. (2020, Table 1) give many examples of modelling systems that include sea ice, most of which are used operationally in national forecasting capacities."

P2 L18: "[models/systems] do not vary in their numerical framework": I would change the wording here because many would argue that there are lots of differences between models/systems - with some models having complicated thermodynamics schemes and others not having sea ice thermodynamics at all (zero-layer). If you mean the dynamics and the overall structural formulation (i.e., AIDJEX continuum dynamic-thermodynamic models) then you should say so.

Good point - we have clarified that we are referring to the sea ice dynamics. The sentence is now: "We note however that their sea-ice dynamics schemes are all based on Eulerian advection schemes and on variants of the viscous-plastic (VP) rheology (although some solve the rheological equations directly while others solve modified equations as is done with the elasto-viscous plastic method (EVP))."

P2 L24: I don't quite get the RIPS->RIOPS paragraph here (aside from J-F requesting it!) because it almost invalidates your argument for running standalone neXtSIM. You need more discussion here, I think. For instance, why did ECCO move away from the basic/standalone RIPS forecast towards RIOPS? What physics/skill does neXtSIM-F likely miss out on by sticking with the "RIPS" style approach (i.e., standalone) rather than the full "RIOPS"?

Good point. It seems like one of the main motivations for RIOPS was having forecasts of currents and tides for search and rescue reasons. There are some drawbacks to the stand-alone approach of course (which is our only possible approach currently), and we added some text, giving the ice edge location as an example.

P3 L8-10: again with the first forecasting system not to use VP you should be careful of RASM using EAP and to ensure that EVP is captured (as in General comments)

We have now clarified VP/EVP and that we are not dealing with seasonal forecasting earlier in the paper.

P4 L20: ocean forcing – please specify the frequency of forcing field update – daily? Hourly? Daily - changed the text.

P4 L20: "TOPAZ near-surface (30m) velocity": Is this exactly the velocity at 30m or the integrated velocity over the top 30m? You need to be more specific. Either way can you say you use this approach?

This was actually a mistake - we currently use the surface (0-3m) currents. Changed the text.

P5 L10: atmos forcing – please specify the frequency of forcing field update – daily? Hourly? 6-hourly

P5 L25: "(European western time)": do you mean UTC? If so that would be easier to understand.

Actually it is Central European Time (CET) since our server runs on Bergen time. Fixed this in the text.

P6 L12-3: "(It is therefore an independent validation dataset for our forecasts.)": I find this odd because the previous sentence – i.e., used for evaluation of free run and forecasts – does not make the data independent! Needs rewording.

Changed to: "It is not assimilated."

P6 L19: "As part of our evaluation we sometimes apply a filter on the uncertainty...": I don't understand why we have sometimes here! Did you do this filtering here or not? If so remove the sometimes; if not why mention it?

Rephrased this section (we always do it but sometimes vary the error threshold for retaining an observation or not).

P7 L11: "In order to compare neXtSIM drift...": You don't need to do this to compare drift though do you? You could use the model velocity fields? The point here is that you are trying to make a better comparison and "compare apples with apples" so why not say that?

Clarified this procedure, to explain better that in this way the drift can be updated every model time step as opposed to every hour (time resolution of the CMEMS product)

P7 L13: "...place Lagrangian drifters...": I would be more explicit and say that you "...seed synthetic Lagrangian drifters into the model...".

We changed this phrase to your suggestion.

P7 L15: Sentence starts with "We use this product" but you've not defined/identified a product yet! Unless you include the "CS2-SMOS" in the subtitle (which I don't). Needs rewording.

Fixed

P7 L27-28: “A delay of one week for thickness would probably be acceptable for assimilation in a real time forecast in the future.” You need to be careful here because many others (including me) would be of the opposite opinion. I guess you mean that you can pretend that the thickness from 7 days ago is the thickness from today and assimilate it? Even so that doesn't sound ideal. If however you're talking about weekly thick ice from CS2 but still using daily SMOS data for thinner ice then perhaps that is ok. However you don't say these things! Either way I would reword to make things clearer. Many of us are pushing hard to get satellite SIT available for near-real-time assimilation and so any statements along the line of what you say here might derail that. So you need to be clear and careful.

Good point - we are certainly thinking about the thicker ice and would have a fairly slow nudging (which also introduces an additional time lag) in order to stop the thicker ice drifting too much from observations, and also recognising there are a lot of uncertainties in these observations. We just deleted this sentence.

P11 L4: “The mesh is generated with Unref (a component of Gmsh...” needs more information to explain what these are. Are they numerical packages or publicly available tools or something?

Changed to “Unref (a component of the open-source mesh-generation library GMSH: Geuzaine and Remacle, 2009)”

P11 L26-29: regarding satellite observations being interpolated to model grid - This is the opposite way around that most people do assimilation, and model-observations comparisons generally, where the model is always translated to the observational locations. Why do you do this the other way around? What impact might this have?

It is done this way mainly to keep the assimilation as simple as possible, especially given the relatively unsophisticated system (e.g. without an ensemble to determine model covariances). However, we agree that this is not the usual way of comparing model to observations. There could be some benefit to interpolating to the observation grid (as you suggest), smoothing the model to allow for a larger satellite footprint and then using this comparison to do the update. It is hard to know exactly how much difference this would make without implementing it, but it is not a priority.

P12 L17: “(this is a kind of assimilation of extent)” given that you don't use the concentration in the pack why do you even need to do the assimilation given that the ice edge would come through in the SMOS SIT initialisation anyhow?

We have now tried to distinguish better in the paper between initialisation of the forecast (done once only, using CS2-SMOS) and the daily assimilation (only uses OSISAF concentration). Thus we don't assimilate SMOS SIT daily. There could be some benefit to adding assimilation of SMOS in winter to constrain the thickness of new ice. However we note the ice edge from SMOS is very low resolution.

P12 L3 (bottom of page!): “the heat flux out of the ocean increases and the ice freezes up again very fast”. Surely this is not always the case? Only if the atmospheric forcing thinks there should be ice? i.e., if there is ice in the ECMWF model used to provide the then the atmospheric fields will be cold and conducive to ice regrowth but if, however, there is no ice in the ECMWF model then the near-surface atmosphere will be warmer?

We found the most dramatic refreezing of removed ice to be due to periods of very cold atmosphere. Rephrased this explanation.

P15 L4: Re Figs 2&3 - what is “mean concentration”? Is this the mean over the whole domain, only ocean points, only sea ice point, or what?

Mean over ocean points.

P17 L12-14: both “Jan-Feb” and “Nov-Dec” are discussed here in relation to Fig 4 but neither panels are present in Fig 4?!

Rephrased this discussion since they are not shown.

P24 L4 (bottom page): “...with 1-day forecasts being launched in between so assimilation still performed daily.” Are these not the analyses?

Yes.

P25 L11: (equation 7): why, when you say IIEE is analogous to RMSE, is the RMSE ratio squared for concentration (eq7a) but the IIEE ratio is not squared for extent (eq7b)?

Rephrased this sentence - we meant that it is like the RMSE in that it is also a positive-definite error metric.

P25 L12-15 (bottom page): this needs rewording. I have no idea what “straw man” means here – are you seeking feedback on a basic idea from others? or trying to scare birds away from your crops? Neither seem appropriate. Also “We have rough benchmarks from some other models however” doesn’t read well. If this relates to the next sentences then you might as well delete it. If not, what are these benchmarks?

A straw man is an expression for an easy-to-beat opponent, but this is probably too colloquial, so it was changed. We rephrased these sentences. The “benchmark” sentence now reads “However, we can make some rough comparisons to the drift errors of other products, even though they use different observations for their evaluation so we can't make a direct comparison. We note that the drift from the TOPAZ forecast generally has a bias in speed of about 2 km/day and a VRMSE of about 5 - 8 km/day (Melsom et al, 2018), while Metzger et al (2017) report an RMS drift speed error of about 5 - 8 km/day in the Arctic for the GOFS 3.1 system.”

P27 L12-13: “we don’t try to correct this as reducing the concentration in the pack causes serious problem with the drift and thickness”: this might work well for a standalone sea ice model but you could have trouble with heat fluxes if you coupled to the ocean and/or atmosphere. Can you say any more about this – particularly in light of the last line of the paper where you mention plans to couple with ocean and/or waves.

We agree that such a method could cause problems when coupled to an ocean or atmospheric model. The ongoing work on using the ENKF assimilation could possibly help - if the model always has 100% concentration in the pack (at least in the winter) then probably the model variance will be very low there compared to the observation error, and the correction would probably be quite small - possibly small enough not to cause problems for the drift and thickness. Another approach could be to let the model variance drop to zero in a smooth way as distance from the edge increases. Another solution could be to use a different product which had a higher concentration in the pack.

Figures

Figs 2 & 3: the text used for legend & axes is very small and hard to read comfortably

Increased the font size for these figures

Fig 4 (caption): the caption should state at the start that these are “differences” between the free run and OSI-SAF

Fixed figure caption

Fig 5: the text is far too small in this figure and the caption makes no mention of what the left and right panels show (I had to zoom in to 200% to see the dates to work out they are the 2 winters - 18-19 and 19-20)!

Increased the font size for this figure and added titles to the 2 columns (“2018-2019 winter”, “2019-2020”)

Figs 6 & 8: panel titles or annotation would make it much easier to see what is plotted in each column

Added annotations

Figs 9 & 12: The detail on this figure is tiny. Even at 200% zoom I struggle to see what's going on properly - particularly the tiny coloured dots.

Increased the font size for these figures

Fig 9 (caption): the caption is all wrong and relates to 3 rows when there are only 2.

Fixed caption

Figs 10 & 11: what do the columns show? panel titles or annotation would make it much easier to see what is plotted in each column

Added annotations

Typos

P1 L4 (and other locations): VP should be “viscous-plastic” not “viscoplastic”

Fixed

P2 L12: “MOSAIC” should be “MOSAiC”

Fixed

P2 L21; “operation” should be “operational”?

Fixed

P3 L11: CMEMS = “Copernicus Marine Environment Monitoring Service”. Your version with “Marine and Environmental” would be a very different beast rather than just doing the “Marine Environment”!

Fixed

P4 L9: Arctic MFC = “Arctic Monitoring and Forecasting Centre” (see <https://marine.copernicus.eu/about/producers/arctic-mfc>)

Fixed

P4 L12: (as mentioned in the last review) the reference for CICEv4.1 is “Hunke and Lipscomb, (2010)” not “Hunke et al.”. It looks like in the references you have changed the date to 2010 but not changed any other details of the reference - including the author list! “Hunke, E. C. and Lipscomb, W. H.: CICE: the Los Alamos sea ice model. Documentation and software users manual, Version 4.1 (LA-CC-06-012).”

Fixed

P4 L30: You should drop the “both” from “used both for assimilation” now that “and evaluation” is deleted

Fixed

P5 L19: AMSR2 = “Advanced Microwave Scanning Radiometer 2”

Fixed

P6 L17: “Cryosat-2” should be “CryoSat-2”

Fixed