Dear editor and referees,
thank you very much for the review of our manuscript. Please find the enclosed detailed response to your comments and suggestions as well as the marked-up version of the manuscript. Please note that our responses are marked by blue colour and indentation.

Anonymous Referee #1

Received and published: 23 April 2018

1. General comments

   The paper presents results for the implementation of a simple sea ice thermodynamic model in a numerical weather prediction forecasting system. This is the first time such a sea ice model has been implemented in this particular NWP system, and a paper on the implementation and results should thus in principle merit publication.

   However, I have serious concerns about the use of a constant ice thickness, about the limited nature of the results used for validation of the model (covering only a short period in March-April 2013), and about some of the methods used for analysis. Some of the language used is also quite cumbersome and difficult to follow. I can therefore recommend publication only after major revisions, which I outline below.

2. Specific comments

   As mentioned above, I thought some of the language used was quite cumbersome and difficult to follow. I would strongly advise the authors to ask a native English speaker to proof-read the paper before resubmitting.

   > We have reworked some parts of our manuscript to make it more readable.

   The authors refer in the introduction to the Met Office Unified Model (page 2, lines 19-20), and state that "...to our knowledge there are no publications about the details of coupling between the advanced sea ice model and the atmospheric model in this system". This is incorrect. The model setup for coupled NWP is described by Lea et al. (2015), while the coupling is described in detail by Hewitt et al. (2011). The authors should cite both of these papers.

   > The paper by Lea et al. (2015) was cited in our manuscript couple of sentences earlier (page 2, line 18 of the first version of the manuscript). However in this paper there is no description of operational system, only of some experimental configuration. That it why we cite it to support the sentence about using coupled systems for research purposes. For the operational configuration of UK Met Office NWP system, from this paper it is clear only that SST is taken from OSTIA and kept constant during the forecast, but not clear what is happening with the ice surface temperature. The paper by Hewitt et al. (2011) provides many interesting details of coupling, but it describes the climate simulation system. In the new version of the manuscript, we cite the reference papers by Walters et al. (2017) and Rae et al. (2015) and correct the text to be more accurate in details: "In operational NWP they are applied in the global NWP systems to provide medium-range weather forecasts, e.g., in UK Met Office Unified Model (Walters et al., 2017; Rae et al., 2015) or IFS-ENS (Integrated Forecasting System–ENSemble prediction system, ECMWF, 2017). However, there are a number of reasons that advanced sea ice models are not widely used for short range operational NWP."

   On page 2, lines 15-18, the authors state that advanced sea ice models are "applied ... in coupled ocean-ice-atmosphere systems for research purposes and seasonal forecasting". I would suggest that they also mention that such sea ice models are used in coupled climate models, such as HadGEM3 (Hewitt et al., 2011).

Corrected, as described in the response to the previous comment.

The description of the model appears to be split between Sections 1, 2 and 3. How- ever, I would prefer to see one single model description section. This could be done in an expanded Section 2. In the current version of the paper, there is some description of the model setup in lines 5-27 of page 3 in the introduction. The authors then describe the sea ice parametrization scheme itself in Section 2. Then, within Section 3, Section 3.2 describes the experimental configuration. I appreciate that some level of discussion of the model is needed in the introduction, but I think the authors probably go into too much detail here. The introduction should set the scene, describe briefly the scientific background and work done by previous authors, and how the present paper builds on that. Much of the discussion of the model would more properly belong in a model description section. Similarly, Section 3 is primarily a results section. I think the description of the experimental configuration in Section 3.2 would again be better placed in an expanded Section 2.

Corrected. Sentences related to model description are moved from Sections 1 and 3 to Section 2.

The model is run with a constant ice thickness of 0.75m. However, the use of a constant thickness is likely to lead to be a source of considerable uncertainty, as ice surface temperature will be extremely sensitive to thickness. Indeed, the authors state in Section 3.1 that "when the ice thicknesses [in SICE and HIGHTSI] are very different, the ice surface temperature values may differ by more than 5°C", and in the conclusions, they say "the sensitivity of the results to the prescribed value of the ice thickness was noticed". Later, they also say that "the simplest way to go forward is to replace the prescribed constant value of the ice thickness by the climatology, to reproduce its seasonal and horizontal large scale variations". This begs the question of why the authors didn't use such a climatology in the current paper. In a resubmitted version of the paper, I would like to see results from a configuration of the model in which the local thickness in each gridbox is prescribed from climatology.

We have run an additional experiment where the ice thickness from the model climatology derived from TOPAZ4 ice reanalysis is used and extended our manuscript accordingly.

I am confused by Section 3.1. The authors state that they only compared SICE with HIGHTSI where the ice thickness in HIGHTSI was "approximately equal" to the constant thickness used in SICE (0.75m). However, they then state that they consider "small" ice thickness differences to be less than 0.4m, which is more than 50% of 0.75m; I would not say that such thicknesses are "approximately equal" to each other. The authors then state "When the ice thicknesses are very different, the ice surface temperature value may differ by more than 5°C", which suggests that, contrary to their previous statement, they have in fact analysed the results for larger differences in ice thickness between the two models. They also do not state what they mean by "very different" - do they mean the difference is greater than 0.4m?

We guess that your confusion was arisen from quite an unclear wording in that paragraph. We did not intend to state that difference of 0.4 m in the ice ice thickness of HIGHTSI and SICE is small or large. We wanted to say that when ice thickness differ by more than 0.4 metres the ice surface temperatures computed by two models become considerably different.

We have reworded the problematic paragraph to make it more clear and avoid confusing assumptions about "small" or "large" differences in the ice thickness.

In Section 3.2 (page 10, line 16), the authors mention that the model is "started from the snow-free state and allowed to accumulate snow from precipitation during the modelling period". What impact will this have on the forecasts? How long will SICE2D-S take to "spin up" to a realistic representation of the snow cover? It will surely be much longer than the 48 hours of the forecasts in the current paper. For this reason, I am not sure how much weight we can give to the SICE2D-S results. The authors should at least comment on this in the paper, and should preferably present results for SICE2D-S forecasts that have been started from a spun-up snow state.

Indeed, the SICE2D-S experiment started from a snow-free state, but it does not mean that each forecast started from a snow-free sea ice surface. HARMONIE-AROME experiments in this study have been run by using so-called cycling. Each cycle consists of data assimilation procedure and model forecast, and cycling period was 3 hours in this study. So, when HARMONIE-AROME starts a new cycle, it uses 3 hourly forecast results from the previous cycle as a background fields for data assimilation. But for sea ice variables there is no data assimilation, so sea-ice variables (e.g., the ice temperatures, snow water equivalent, snow-density, etc.) from the previous forecast are used, without any modifications, as an initial state for the current forecast. Therefore, in SICE2D-S experiment evolution of snow cover does not take just 48 hours, but goes throughout the whole experiment, that is 2 months.

To avoid further confusion we have added explanation to section 3.2 to show how snow on sea ice is initialized from cycle to cycle.

In Sections 3.3 and 3.4, the authors analyse forecasts only for a short period (March- April 2013). However, the performance of the model is likely to vary during the year, and the results may well be different in different seasons. I note that the authors state in the conclusions that in the future the model will be evaluated for more regions and more seasons, but I am of the opinion that results for other seasons must be included in the present paper for it to be worthy of publication. Given that Arctic sea ice exhibits a clear annual cycle, it is important that the impact of this on the performance of the model is analysed.

Yes we completely agree that selected study period is somewhat short and does not provide information about the performance of the ice scheme throughout the year. But when we test a high-resolution atmospheric model combined with a data assimilation system it becomes extremely costly in terms of CPU time, run time, storage, as well as in terms of cost of computing time on HPC to run experiments that are longer than just a couple of months.

We have no resources and funding that would allow us to run all experiments that were discussed in our study as a one year long experiment. However it is possible to use operational archive data to show the performance of the ice scheme during different seasons, selecting the ice surface temperature as a verification variable. In this case additional experiments are not needed.

To show the performance of the new sea ice scheme throughout the year we have compared ice surface temperatures from the operational AROME Arctic and near real time L2 ice surface temperature values derived from MODIS and VIIRS sensors that cover the time period from December 1st 2015 to December 1st 2016. To study the performance of HARMONIE-AROME without SICE scheme throughout the year, the ice surface temperature fields interpolated from the host model IFS-HRES were used (because they would be used in the forecast without SICE scheme). We have extended Section 3.4 accordingly to show the results of this new comparisons.

In Figures 2, 3 and 6, the authors present results for the mean error in the forecast MSLP, 2-metre temperature, and 10-metre wind speed, where the error is defined as the difference between the modelled and observed quantities, and the mean is taken over several observing sites. However, the standard deviations plotted in Figures 2, 3 and 6 are often much larger than the differences between the means. The authors discuss the differences between the mean errors for different experiments at length in Section 3.3, and consider possible reasons for them, but the fact that the standard deviations are so large compared to the differences suggests that the differences may not be significant. If this is indeed the case, then it suggests that the SICE scheme, and the related snow and form drag schemes, may not have a significant impact on the model-obs errors. It would be interesting to see if this conclusion changes when the authors use RMS error rather than simple mean error, and when they look at forecasts for different times of year.

This question indeed contains three aspects: 1) statistical significance,

2) relation between bias, error standard deviation and root-mean-square error, and
3) forecast errors for different times of the year. Aspects (2) and (3) are also mentioned in your other comments, so we discuss them in the answers to these comments. Aspect (1) we discuss here. We compare forecast errors between different numerical experiments for the forecasts of different length. We consider the forecast error as a random value. So, our sample for calculating statistics has the following volume: $(number\ of\ forecasts\ per\ day) * (number\ ofdays) * (number\ of\ stations)$. For example, for Figure 2 in the previous version of the manuscript, the sample volume is $(1 forecast\ per\ day) * (60 days) * (7 stations) = 420$ cases. This is quite large sample volume, even if we consider only 7 stations – because we have 30 days to make statistics. With samples of these volume, even with large standard deviations, the results are usually statistically significant. In this example, for the 39-hour forecasts, the difference in the mean errors (biases) between experiments REF and SICE2D-NS is statistically significant for 2-metre temperature and MSLP with the level of 0.1%, and for 10-metre wind with the level of 5%. Usually in NWP the sample volumes are large or very large, sometimes thousands of cases. That is why the mathematically strict statistical significance estimates are often not displayed on verification plots in NWP. In our case, the sample volume may be sometimes lower then hundreds of cases. Thus, in the new version of the manuscript, although we don't estimate the statistical significance strictly, but we discuss the sample volume and make circumspect conclusions in cases when it is small.

Another point relating to Figures 2, 3 and 6 is that the use of a simple mean error will potentially lead to positive and negative errors cancelling each other out. This will hide potentially-relevant results if some stations have a very large positive bias and others have a large negative bias. For this reason, I think the root-mean-square error would be a more useful quantity to assess, and I would like to see a plot of this, either instead of or in addition to the simple mean error that the authors have plotted here.

In fact, all three errors, namely the mean error (BIAS), the error standard deviation (ESTD) and the root-mean-square error (RMSE) are connected with each other. Indeed, if FE is the forecast error, then BIAS is an average FE, ESTD is a SQRT of an average of (FE-BIAS)^2, RMSE is a SQRT of an average of FE^2, then

RMSE^2=BIAS^2+ESTD^2

That is why, we may use either RMSE or ESTD to estimate the random error of the forecast. With BIAS, we estimate a systematic error, so BIAS is also important. RMSE, in turn, contains both: BIAS and ESTD. Positive and negative errors of different forecasts, although cancelling each other in BIAS, are reflected in ESTD, not only in RMSE. But RMSE is very traditional, so we add it. Also, estimates of statistical significance of the difference of RMSE between two experiments is problematic, because in this case the random value is FE^2, which has non-gaussian distribution. Then, Student's criterion is not applicable, and more complicated methods are necessary.

It would also be interesting to see the contributions of the different observing stations to the mean (or RMS) error. This could be done using maps of a similar form to Figure 3 of Bellouin et al. (2011), where the observations are shown with boxes superimposed on a map showing the fields output by the model.

Thank you for such a good idea! We have added a map showing the relative change in 2 metre air temperature RMSE of the SICE2D-NS experiment compared to the REF experiment (See Figure 2 in the updated manuscript).

This figure also shows that stations with the most noticeable difference in RMSE between the SICE2D-NS and REF experiment are located in coastal areas close to the sea ice and for the inland stations there is almost no differences in RMSE between the SICE2D-NS and REF experiments.

The authors mention in the text (page 10, line 21) that they used 12 Svalbard stations, and indeed 12 are shown on the map in Figure 1. However, in the captions of Figures 2 and 6, they mention "7 Svalbard stations", and list the 7 stations. I presume this is because of the issues described in Section 3.3 (page 10, lines 23-26) whereby some Svalbard stations were excluded because they were in fjords. But were the other 5 stations used at all in this analysis? If not, then it is incorrect to state at the beginning of Section 3.3 that measurements from 12 Svalbard stations were used (as in fact only 7 were used). The authors should re-word this paragraph (page 10, lines 21-26) to make this clearer.

Yes, you are right. From the 12 selected stations only 7 were surrounded by sea ice and the rest of them was close to the open sea or located in fjord areas where interpolation issues were noticed.

To avoid confusion, we have reworded description of selected Svalbard stations to clearly show that some of stations were removed. We have also updated the Figure 1 to show only the stations that were actually used for the comparisons.

In Section 3.3 (page 11, lines 14-18), the authors discuss the relative sizes of the standard deviations of the errors in REF and SICE2D-NS, without any mention of the implications or relevance of these results. Presumably a smaller standard deviation implies that there is a smaller range of errors between stations. Is this relevant, and if so why? Is there any indication what might be causing it? Does the implementation of the sea ice scheme affect the 2-metre temperature at some stations more than others? Is there an obvious reason for this?

Indeed, the smaller error standard deviation indicates that the range of 2 metre temperature (T2m) errors is smaller in SICE experiments than in REF. In the other words, random component of the modelling errors has been reduced. This is a desirable property for an operational NWP system because systematic errors of a forecast could be corrected later during the post-processing stage.

When using SICE, sea ice cover is defined by the ice concentration field and ice surface temperature is computed by the ice scheme. But in REF ice cover is taken from the surface temperature forecast by IFS-HRES and kept constant during the forecast. As result, in REF, constant ice surface temperature filters out variations in T2m and leads to considerable range of errors for different SYNOP stations.

Of course, the effect of SICE on the T2m forecast varies from stations to station. This is because the T2m forecast could be strongly dependent on the local conditions, such as characteristic wind direction or station elevation, notwithstanding the fact that stations surrounded by closed ice would show more clear response than stations that have only traces of ice in their vicinity.

We have extended the section 3.3 to make the text more clear and indicate that errors in SICE experiments are less random than in REF.

At the end of Section 3.3 (page 12, lines 31-34), the authors state "... with observations from coastal stations only, we lack understanding of the ice temperature behaviour for larger scales". This is a very good point to make, and I would recommend that when the authors resubmit the paper they include results for a wider range of stations within the forecast domain, including non-coastal (i.e. inland) stations. Does the implementation of the sea ice scheme affect the results only at stations that are physically close to the sea ice, or are there larger-scale effects?

When stating "... with observations from coastal stations only, we lack understanding of the ice temperature behaviour for larger scales" we meant the characteristics of the ice field itself rather than performance of the inland stations. Inland stations that are located hundreds and thousands kilometres away from the sea ice show almost no response to changes in the sea ice parameterization, as can be seen from the provided map in the updated version of manuscript. And to study the larger-scale performance over the sea ice covered areas we have compared results of numerical experiments and extended our manuscript accordingly.

We have provided scatter plots of relative change in RMSE induced by SICE to show the impact over all available weather stations.

Figure 7 shows surface temperature derived from MODIS data, and forecast by the model. However, it is quite difficult to get an idea of the differences between the temperature fields in the plots. It would be much more useful if the authors could present maps showing the difference between these (i.e., model minus MODIS). This would help the reader to understand better the results discussed in Section 3.4.

Figure 7 was meant to illustrate the similar patterns in the ice surface temperature fields from SICE within the operational NWP system and MODIS ice surface temperature product. We have replaced it by the maps of root means square errors of the ice surface temperature. We provide such figures for short-term experiments, discussed in the Section 3.3 and for evaluation of performance of the operational NWP system which uses SICE for parameterization of sea ice, discussed in the Section 3.4.

The authors mention in Section 3.4 (page 13, lines 10-12) that most MODIS swaths were in the daytime. However, the model results shown in Figure 7 are whole-day averages. How will this affect the comparison of the two? I imagine there may be a warm bias in the MODIS observations as a result of the fact that they are generally restricted to daytime. The authors should comment on this, and its implications for the results, in the paper.

Thank you for this comment. We agree that daily aggregated ice surface temperature product should not be compared to time-averaged model results. To make a more clear comparison in the new version of the manuscript we use near real rime ice surface temperature products instead of daily aggregated ones. To show the difference between the forecasted ice surface temperature and data provided by satellite products we have replaced Figure 7 by maps of the root mean square error for different experiments (see our reply to the previous comment).

3. Technical corrections

   – Page 2, lines 2-3: "Over areas with a mixture of floes and polynyas, the form drag appears, which affects the turbulent fluxes". I would suggest re-wording this, so that it reads: "Over areas with a mixture of floes and polynyas, the turbulent fluxes are affected by form drag".

     Reworded according to the suggested variant

   – Page 2, lines 19 and 34: I don't like the use of "To our knowledge…", as it seems unscientific to me. I have already mentioned above that the statement made on lines 19-20 is in fact incorrect. I would also suggest an alternative wording for the sentence on lines 34-35. Indeed, if one doesn't know whether a particular statement is true or not, it is often best not to include it at all, rather than preceding it with "To our knowledge…".

     We have reworded those sentences to avoid unscientific language.

   – Page 4, line 6: "…it is designed so that it can be naturally coupled with a snow scheme…": I don't know what the authors mean by "naturally coupled". I think that "…so that it can be coupled to a snow scheme…" would suffice.

     Reworded according to the suggested variant

   – Page 7, line 25: "It is important to mention that…": This is unnecessary. If it's important to mention it, then mention it – there is no need to say that it's important to do so.

     We have removed "It is important to mention that…" as suggested.

– Page 9, line 25: "The background for the data assimilation are fields of prognostic variables. . . ": I think there is a word missing here, and that this should read " The background fields for the data assimilation. . . ".

Fixed

– Page 10, line 27: Figure 6 is mentioned before Figures 4 and 5. The figures should be re-ordered to avoid this.

We have re-ordered figures to avoid this situation

– Page 11, lines 6-7: ". . . the underestimation of night-time 2 metre temperatures over land is a characteristic feature of the model known from operational verification (not shown)". If this is known from operational verification, is there a reference that the authors can cite?

That sentence was supposed to be removed from the submitted manuscript and was left there by a mistake. We have removed that sentence and added explanation why there is a noticeable diurnal cycle of the 2 metre temperature bias in the REF experiment.

– Page 11, lines 14-15: "The error standard deviation for the 2 metre temperature forecasts. . . ". This should read "The standard deviation of the errors in the 2 metre temperature forecasts. . . ".

Reworded according to suggested variant

– Page 11, line 22: "mean sea level pressure error standard deviation" sounds clumsy. I would suggest "standard deviation of the error in mean sea level pressure". The authors could also abbreviate "mean sea level pressure" to "MSLP", if they define the abbreviation the first time they use it.

That phrase has been reworded according to suggestions

– Page 11, line 31: " over the part of the grid cell related to the sea with ice". I'm not sure what this means. Does it mean "over the ice-covered part of the grid cell", or something else? I would suggest re-wording this to make it clearer.

This sentence supposed to mean "over the part of grid cell that contains both open sea and sea ice". We have reworded that sentence to make it less confusing.

– Page 12, lines 4 and 18: I think the authors mean "in agreement with" rather than "in accordance with".

Fixed

– Page 12, line 20: ". . . makes the surface temperature drop more and more". This language ("more and more") is not very scientific. Please consider re-wording.

We have reworded that sentence to use more scientific language

– Page 14, line 18: ". . . the sensitivity of the results to the prescribed value of the ice thickness was noticed". I think the authors mean "noted" rather than "noticed".

Fixed

– In the caption of Figure 3, the authors mention 7 stations in the Gulf of Bothnia, and 7 are shown in the map in Figure 1, but in the text (page 10, line 21) they state that they used 6 stations in the Gulf of Bothnia.

Information about the number of used stations in the Gulf of Bothnia has been corrected in the text.

- The authors state in the text that the modelled ice surface temperature shown in Figure 7 is for the configuration which doesn't include the snow scheme (i.e., SICE2D-NS), but it would be helpful to the reader if they also re-stated this in the figure caption.

  Figure 7 has been removed from the updated version of the manuscript, but we have extended the corresponding captions for the new figures to explicitly state that these results were obtained from the snow-free configuration of SICE.

4. References

Anonymous Referee #2

Received and published: 23 April 2018

General Comments:

This paper describes the impact of including a simplified ice scheme into the ALADIN- HIRLAM numerical weather prediction system version 38h1 for a short 2-month analysis period in March/April 2013 for 2 regions near Svarlbard and the Gulf of Bothnia with horizontal resolutions of 2.5 km. In the AROME Arctic domain, the SICE2D-NS (no snow) model performs best with the lowest mean error as a function of lead forecast time for mean sea level pressure, easily outperforming the reference run which does not include the SICE model. Some improvement is seen for that region with the 2-m air temperature mean error, where up to a 0.5°C improvement is made over the reference run where both have a negative bias. The SICE experiment without the snow model shows a fairly consistent 1° positive bias for the 3-45 hour lead forecast times. However, when examining the wind speed mean error, the reference run without the SICE model consistently showed the lowest bias. When examining the results for the Gulf of Bothnia domain, there was no discernible difference for any of the forecasts when examining sea level pressure. Overall, the SICE-NS experiment performed best for mean error and standard deviation for 2-m air temperature. The SICE-2D-S performed best for wind speed for both mean error and standard deviation. Experiments performed for March 2013 with a form drag parameterization (SICE-AD) could not outperform the reference run which did not include SICE. Qualitative figures presented for 2 days in March 2017 for the model versus MODIS ice surface temperature, bring little additional insight to the model performance.

The paper is filled with acronyms for numerous modeling systems (e.g., HARMONIE-AROME, ALADIN-HIRLAM etc.) which are never defined. The analysis period is short (March 1 – April 30, 2013), with limited data available for model-data comparisons.

For future reference, while coupling to an ice model such as CICE may provide the best overall improvement for Arctic NWP, the authors are encouraged to investigate the CICE Consortium's column physics package Icepack v1.0, which was released in February 2018. It is worth considering for future applications and is freely available to the public. (See https://github.com/CICE-Consortium/Icepack).

This paper is well written, but the study period is short. I recommend publication when the following issues are addressed in a revised version.

Thank you for suggesting the Icepack package. We will definitely consider it for our applications.

Specific Comments:

Although properly referenced, spell out all acronyms for the following:

HARMONIE-AROME, ALADIN-HIRLAM, CICE, GELATO, HIGHTSI, DWD, SURFEX

We have spelled out acronyms for the various models and NWP systems that are mentioned throughout manuscript.

Section 3.3: The experiments occur during a short period of time, ~2 months. Can these be extended for a longer period to better assess the model's performance? How does the model perform during the summer melt season? Instead of initializing with a constant ice thickness value of 0.75 m, consider testing with 28-day averaged near real-time, or seasonal values from CryoSat-2 (CPOM, see http://www.cpom.ucl.ac.uk/csopr/seaice.html). Seasonal (Spring and Autumn) derived ice thickness data is available going back to 2011). Data is available on a 1 and 5 km grid.

Thank you for suggesting the CryoSat-2 data as a source of ice thickness climatology. We have checked this dataset and found it somewhat noisy and a bit problematic to prepare for using in our system. Instead we use model climatology based on the sea ice reanalysis from the TOPAZ4 system.

He have run an additional experiment with SICE where we use the ice thickness from model climatology. This experiment has been run for the same time period as SICE2D-NS and SICE2D-S experiments. We have extended our manuscript accordingly.

Page 13: Sec 3.4: I would like to see actual comparisons between MODIS and the model experiments (e.g., tabular statistics). On page 13 lines 7-8, you state "Statistical assessments require application of special methods, which were out of the scope of this study". Why? This would add value to your paper and possibly complement the results already shown. For the examples shown for March 2017, please add figures that show the HARMONIE-AROME run without SICE. In addition, check the availability of VIIRS ice surface temperature from NSIDC: https://nsidc.org/the-drift/data-update/viirs-sea-ice-surface-temperature-swath-data-now-available. If VIIRS is available, can you examine the difference between the modeled ice surface temperature with the SICE experiments versus the VIIRS product?

From the statistical point of view, model errors obtained from the comparison with satellite observations are in fact time series of random 2D fields. Thus, the verification statistics will be dependent on the methods of sampling (in time, or in space, or both), of aggregation the information from one grid to another, etc. At least, this is a huge amount of data to process. Usually these kind of studies deserve special attention and special publications. But we agree that adding comparisons between satellite products and model experiments would add value to our study and we have extended Section 3.4 to provide this information. Please see our answer to the similar question, raised by the first Referee.

Technical Corrections:

Page 3 line 10: change to "The scheme that is developed"

This sentence has been removed from the updated version of manuscript

Page 6 line 16: define ISBA

Added

Page 6 line 25; insert a comma after "In this case"

Fixed

Page 7 line 25: what do you mean by "screen level"?

The term "screen level" corresponds to the mount height of sensors, situated inside the thermometer screen, which should be within the range of 1.25 to 2 metres according to Guide to Meteorological Instruments and Methods of Observation (WMO, 2008). To avoid further confusion we have rearranged text to avoid mixing screen level parameters and other meteorological parameters such as snow depth or 10 metre wind speed.

Page 11, 12, 13 (twice), 15: replace “happens” with “occurs”

Fixed

Page 15 line 25: should read “which is not available to the general public”

Changed according to recommendation

Page 20: Müller references should be listed as 2017a and 2017b; correct the text as necessary.

We have corrected the text to list Müller references as 2017a and 2017b.

Page 20: check spelling for Posey references, several surnames spelled wrong.

Spelling has been corrected

Page 22: Is there a range for the number of snow layers? If yes, please state it.

We have added the technically valid range of the number of snow layers. The number of snow layers for ISBA ES in SICE can not be changed through the configuration file, but should be done directly in the source code.

Page 23: Table 2: Define the “ice scheme” in the caption.

Caption of Table 2 has been extended to provide definition of the “ice scheme” column.

References

ECMWF: PART V: ENSEMBLE PREDICTION SYSTEM, IFS Documentation, ECMWF, 2017.

Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geoscientific Model Development, 4, 223–253, https://doi.org/10.5194/gmd-4-223-2011, https://www.geosci-model-dev.net/4/223/2011/, 2011.

Lea, D. J., Mirouze, I., Martin, M. G., King, R. R., Hines, A., Walters, D., and Thurlow, M.: Assessing a New Coupled Data Assimilation System Based on the Met Office Coupled Atmosphere-Land-Ocean-Sea Ice Model, Monthly Weather Review, 143, 4678–4694, https://doi.org/10.1175/MWR-D-15-0174.1, 2015.

Rae, J. G. L., Hewitt, H. T., Keen, A. B., Ridley, J. K., West, A. E., Harris, C. M., Hunke, E. C., and Walters, D. N.: Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model, Geoscientific Model Development, 8, 2221–2230, https://doi.org/10.5194/gmd-8-2221-2015, https://www.geosci-model-dev.net/8/2221/2015/, 2015.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geoscientific Model Development, 10, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017, https://www.geosci-model-dev.net/10/1487/2017/, 2017.

WMO: Guide to Meteorological Instruments and Methods of Observation, WMO-No. 8, World Meteorological Organization, Geneva, Switzerland, 2008.

Implementation of a simple thermodynamic sea ice scheme, SICE version 1.0-38h1, within the ALADIN-HIRLAM numerical weather prediction system version 38h1

Yurii Batrak[1], Ekaterina Kourzeneva[2], and Mariken Homleid[1]

[1]Development Centre for Weather Forecasting, Norwegian Meteorological Institute, Oslo, Norway
[2]Finnish Meteorological Institute, Helsinki, Finland

Correspondence to: Yurii Batrak (yurii.batrak@met.no)

Abstract. Sea ice is an important factor affecting weather regimes, especially in polar regions. A lack of its representation in numerical weather prediction (NWP) systems leads to large errors. For example, in the HARMONIE-AROME model configuration of the ALADIN-HIRLAM NWP system, the mean absolute error in 2 metre temperature reaches 1.5 °C after 15 forecast hours for Svalbard. A possible reason for that is that the sea ice properties are not reproduced correctly (there is no prognostic sea ice temperature in the model). Here, we develop a new SImple sea iCE scheme (SICE) and implement it ~~into~~ in the ALADIN-HIRLAM NWP system in order to improve the ~~quality of its forecasts~~ forecast quality in areas influenced by sea ice. General evaluation of the new parameterization is performed within HARMONIE-AROME by experiments covering the Svalbard and Gulf of Bothnia areas for a selected period in March – April 2013. It is found that using the SICE scheme improves the forecast, decreasing the value of the 2 metre temperature mean absolute error on average by 0.5 °C in areas that are influenced by sea ice. The new scheme is sensitive to the representation of the form drag~~: it may increase the~~. The 10 metre wind speed bias ~~on average~~ increases, on average, by 0.4 $\mathrm{m\,s^{-1}}$ when the form drag is not taken into account. Also, the modelling results are compared with the sea ice surface temperature ~~observations from MODIS~~ products from MODIS and VIIRS. The warm bias (of approximately 5 °C) of the new scheme is indicated for the areas of thick ice in the Arctic. Impacts of the SICE scheme on the modelling results and possibilities for future improvement of sea ice representation in the ALADIN-HIRLAM NWP system are discussed.

1 Introduction

Sea ice, permanent or seasonal, covers large areas of the ocean, especially in polar regions. Sea ice is a complex system with many important processes occurring. Being an interface between the atmosphere and the underlying medium, the sea ice surface temperature (in contrast to the sea surface temperature) has a noticeable diurnal cycle. Snow ~~is accumulated~~ accumulation on the ice ~~and~~ is accompanied by specific processes of snow-ice formation during the cold season, and by snow melt and the appearance of melt ponds during the warm season. Freezing of saline water results in brine droplets becoming trapped in the ice. This affects not only the ice thermal properties but also the ice structure, due to the slow movement of the trapped droplets towards the ice bottom and the formation of channels. Finally, the ice covered area is not a solid shield but a mixture of floes and polynyas that drift, being forced by the wind and ocean currents. Large scale ice covered areas strongly affect the properties of the atmospheric surface boundary layer over them. Night cooling of the ice surface may lead to a very stable boundary layer and limited turbulent exchange between the surface and the atmosphere. Over areas with a mixture of floes and polynyas, the form drag appears, which affects the turbulent fluxes turbulent fluxes are affected by form drag. Thus, it is very important to reproduce the these processes over the sea ice correctly in numerical weather prediction models.

Traditionally in NWP applications, simple parameterization schemes for sea ice are used. Information about the presence of sea ice cover is taken from observations (the analysis), and the sea ice thickness and sea ice temperature are modelled by a parameterization scheme. For parameterization of the sea ice in NWP, two main approaches currently exist: sea ice schemes based on the solution of the heat diffusion equation with several ice layers, but constant ice thickness(ECMWF, HIRLAM) (ECMWF, 2017a; Unden et al., 2002), e.g. IFS-HRES (Integrated Forecasting System–High RESolution) by ECMWF (European Centre for Medium-range Weather Forecasts) prior to version cy45r1 (ECMWF, 2017a) or HIRLAM (HIgh Resolution Limited Area Model) (Unden et al., 2002); and bulk sea ice models with prognostic ice depth thickness and an assumed linear or polynomial shape of the temperature profile in the ice(DWD, e.g. COSMO (Consortium for Small-scale Modeling) by DWD (Deutscher Wetterdienst) (Mironov et al., 2012; Mironov and Ritter, 2004; Mironow and Ritter, 2003). Snow on ice in these schemes is either not represented (ECMWFe.g. IFS-HRES), or represented parametrically via changing the albedo from ice to snow during the melting period (DWD) (Mironov and Ritter, 2004)(e.g. COSMO, Mironov and Ritter, 2004). Simple sea ice schemes are used for operational forecasting. However, their performance has mainly been studied in general, with minor validations against observations and without comparisons with more advanced ice models.

For ice forecasting applications and research purposes, more advanced ice models have been developed, for example CICE (Hunke et al., 2015), GELATO (Mélia, 2002) and HIGHTSI (Cheng and Launiainen, 1998)(Community Ice CodE, Hunke et al., 2015), GELATO (Global Experimental Leads and ice for ATmosphere and Ocean, Mélia, 2002) and HIGHTSI (HIGH resolution Thermodynamic Snow and Ice model, Cheng and Launiainen, 1998). They are applied in ocean modelling (e.g., Blockley et al., 2014; Dupont et al., 2015)(e.g. Blockley et al., 2014; Dupont et al., 2015), and in coupled ocean-ice-atmosphere systems for research purposes, climate simulations and seasonal forecasting (Brassington et al., 2015; Lea et al., 2015; MacLachlan et al., 2 However, they are rarely used in operational NWP . One exception is the (Brassington et al., 2015; Lea et al., 2015; MacLachlan et al., 2015; Hewitt et al., 2011; Pellerin et al., 2004). In operational NWP they are applied in the global NWP systems to provide medium-range weather forecasts, e.g., in UK Met Office Unified Model , but to our knowledge there are no publications about the details of coupling between the advanced sea ice model and the atmospheric model in this system. There are (Walters et al., 2017; Rae et al., 2015) or IFS-ENS (Integrated Forecasting System–ENSemble prediction system, ECMWF, 2017b). However, there are a number of reasons that advanced sea ice models are not widely used for short range operational NWP. Firstly, the advanced ice models are computationally expensive, representing in detail many processes that are important for the evolution of the sea ice itself, but of secondary importance for the description of ice-atmosphere interactions. Secondly, their robustness and numerical stability during coupling with atmospheric models needs more studies within a framework of short range operational NWP systems.

Thirdly, they may require advanced methods of data assimilation for their initialization. In NWP research, in addition to coupled systems, advanced ice models may be used for the performance assessment of simple schemes.

Observations of the sea ice properties that are currently used in NWP are very limited ~~. They mainly~~ and often only indicate only the presence of ice. One example is the sea ice concentration product provided by Ocean and Sea Ice Satellite Application Facility (OSI SAF) (Andersen et al., 2012; Breivik et al., 2001), which uses observations from passive microwave sensors. This product is in turn included in the OSTIA (Operational Sea surface Temperature and sea Ice Analysis~~product (OSTIA)~~) product (Stark et al., 2007; Donlon et al., 2012). For using the sea ice concentration data in an NWP system, they need to be projected onto an atmospheric model grid, which is usually done during the analysis step. Subsequently, the sea ice concentration ~~is used to govern~~ governs the sea surface schemes: for the ice-covered part of a grid box, a simple ice model runs, for the ice-free part, the sea surface temperature is kept constant. ~~To our knowledge, no other~~ Other observations of ice properties are rarely used for assimilation in short range NWP systems, and very few of them are used for validation. In ocean modelling systems, sea ice concentration and sea ice drift data from different sources, including remote sensing from passive microwave and visible channels are used (see, for example, Posey et al., 2015; Sakov et al., 2012). Acquisition of the sea ice depth data from active remote sensing is being developed (Tilling et al., 2016), but the latency of these data is not yet acceptable for operational use in short range forecasting. However, these data also could serve as a source of information about sea ice properties for an NWP system.

In the ALADIN-HIRLAM (Aire Limitée Adaptation dynamique Développement InterNational–HIgh Resolution Limited Area Model) NWP system both the sea surface temperature and the sea ice surface temperature ~~are kept~~ remain constant during the whole forecast. ~~They~~ Each forecast cycle they are initialised from an external source (for example, from the global ECMWF model IFS)~~each forecast cycle. Over ice-covered and in ice-surrounded areas, this~~ . This causes noticeable errors in near surface air temperature forecasts over ice-covered and in ice-surrounded areas, especially for forecasts longer than 24 hours.

~~The main purpose of this study is~~ This study presents the development of a simple ~~parameterization scheme for sea ice~~ sea ice parameterization scheme for the ALADIN-HIRLAM NWP system. The scheme ~~that was developed~~ solves the heat diffusion equation on a vertical grid within a sea ice slab of constant thickness. This level of simplification was chosen as a first step. ~~Prognostic sea ice thickness, which needs special attention during the initialization and analysis step, is considered for future developments. Technically, the new sea ice scheme is developed in the framework of the land and ocean surface modelling platform SURFEX (Masson et al., 2013), which is incorporated in the ALADIN-HIRLAM NWP system. In the sea ice~~ In the scheme, provision is made to couple it with the snow scheme ~~from SURFEX~~ after Boone and Etchevers (2001). Also, in the case of fractional ice cover, the form drag caused by ice floating over the water surface is taken into account following Lüpkes et al. (2012).

The simple sea ice model (parameterization scheme) is checked for sanity and its performance is assessed through a comparison with the off-line sea ice model HIGHTSI (Cheng and Launiainen, 1998). ~~HIGHTSI, although not containing the ice dynamics (unlike CICE and GELATO), reproduces the temperature profiles in the ice with a sufficient level of accuracy (Cheng et al., 2008) and needs a minimal amount of forcing data.~~ The overall performance of the HARMONIE-AROME

(HIRLAM ALADIN Research On Meso-scale Operational NWP in Euromed–Application of Research to Operations at MEsoscale) configuration of the ALADIN-HIRLAM NWP system with the new sea ice parameterization scheme is evaluated against temperature and wind measurements from coastal meteorological (SYNOP) stations and ice surface temperature ~~observations from the~~ products derived from measurements by the MODIS (Moderate Resolution Imaging Spectroradiometer~~(MODIS)~~) and VIIRS (Visible Infrared Imaging Radiometer Suite) sensors. The scheme ~~results in an improvement of~~ improves the forecast verification scores of the ALADIN-HIRLAM NWP system in coastal areas that are influenced by sea ice ~~, but it~~ but overestimates the sea ice surface temperature in the Arctic, where the prescribed constant value of the ice thickness is too small. The experiments and results described in this paper ~~, and the experience gained, will enable a~~ enable better understanding of the forecast errors and uncertainties and provide an advancement in the description of the interactions between sea ice and the atmosphere in NWP.

The paper is organized as follows. In Section 2 the scheme description ~~is given, which includes~~ and an overview of the physical equations are given. Numerical methods to solve the scheme equation are described in Appendix A. Section 3 ~~addresses~~ evaluates the performance of the new scheme ~~evaluated~~ by comparison with the thermodynamic sea ice model HIGHTSI, measurements from SYNOP stations and observations from MODIS and VIIRS. In the final section a short summary of the obtained results is given and the perspectives for further developments are discussed. Fortran source code of the SICE scheme version 1.0-38h1 is provided in the Supplement.

2 Description of the sea ice parameterization scheme

The ~~purpose of the~~ SImple sea iCE scheme (SICE, pronounced "ess ice") ~~, which~~ is developed for the parameterization of sea ice in NWP ~~, is the prediction of~~ to predict the surface temperature of a thick layer of sea ice. The ice thickness is prescribed. No ice melting or ice formation processes are included and the heat flux from water to ice is neglected. Processes of snow-ice formation, which are discussed e.g. by Saloranta (2000), are not represented. The scheme describes only the processes in the ice slab ~~, but it is designed so that it can be naturally coupled with~~ but it can be coupled to a snow scheme that can provide the value of the heat flux on the lower boundary of the snow layer. The prognostic temperature profile in the ice is obtained from the solution of the heat diffusion equation using the heat balance equation and the temperature of water freezing as upper and lower boundary conditions respectively:

$$\begin{cases} C\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\lambda\frac{\partial T}{\partial z} - \frac{\partial Q}{\partial z} \\ 0 = F + \lambda \left.\frac{\partial T}{\partial z}\right|_{z=0} & \text{if } z = 0, \\ T = T_{frz} & \text{if } z = H, \end{cases} \tag{1}$$

where t is time (s); z is depth (m); C is the volumetric heat capacity of ice $(\mathrm{W \cdot s\, m^{-3}K^{-1}})$; λ is the ice thermal conductivity $(\mathrm{W\, m^{-1}K^{-1}})$; Q is the solar radiation flux penetrating through the ice $(\mathrm{W\, m^{-2}})$; T is the ice temperature (K); T_{frz} is the freezing point of sea water (K); and H is the prescribed ice thickness (m). The term F in the second row of Eq. (1) represents the balance of incoming downward and upward heat fluxes:

$$F = \delta_{H_{snow}} \left[LW\!\downarrow - \varepsilon\sigma T_s^4 - \rho_a c_p c_{\mathrm{H}} V_{\mathrm{N}} \left(T_s - T_{\mathrm{N}} \frac{T_s}{\Pi_s} - \frac{T_{\mathrm{N}}}{\Pi_{\mathrm{N}}} \right) - L \rho_a c_{\mathrm{H}} V_{\mathrm{N}} \left(q_{sat}(T_s) - q_{\mathrm{N}} \right) \right] + (1 - \delta_{H_{snow}}) G_{snow} \tag{2}$$

where $\delta_{H_{snow}}$ is the Kronecker delta:

$$\delta_{H_{snow}} = \begin{cases} 1 & \text{if } H_{snow} = 0, \\ 0 & \text{if } H_{snow} \neq 0; \end{cases} \tag{3}$$

and T_s is the ice surface temperature (K) ($T_s \equiv T|_{z=0}$); $LW\!\downarrow$ is the downward longwave radiation flux ($\mathrm{W\,m^{-2}}$); ε is the surface emissivity; σ is the Stefan-Boltzmann constant ($\mathrm{W\,m^{-2}K^{-4}}$); ρ_a is the air density ($\mathrm{kg\,m^{-3}}$); c_p is the air heat capacity with constant pressure ($\mathrm{W \cdot s\,kg^{-1}K^{-1}}$); c_{H} is the drag coefficient for heat; V_{N} is the wind speed ($\mathrm{m\,s^{-1}}$); $\Pi_{\{s,\mathrm{N}\}}$ is the value of the Exner function on the corresponding level; T_{N} is the air temperature (K); L is the latent heat of sublimation ($\mathrm{W \cdot s\,kg^{-1}}$); $q_{sat}(T_s)$ is the saturation specific humidity near the ice surface ($\mathrm{kg\,kg^{-1}}$); q_{N} is the specific humidity of air ($\mathrm{kg\,kg^{-1}}$); G_{snow} is the heat flux from snow to ice ($\mathrm{W\,m^{-2}}$); and H_{snow} is the snow thickness (m). Index N denotes a variable at some level in the atmosphere (the lowest atmospheric model level if the scheme is included in an atmospheric model). The right hand side of Eq. (2) is the sum of the longwave part of the radiative balance $LW\!\downarrow - \varepsilon\sigma T_s^4$ and the turbulent fluxes of sensible ~~$H = \rho_a c_p c_{\mathrm{H}} V_{\mathrm{N}} (T_s - T_{\mathrm{N}})$~~ $H = \rho_a c_p c_{\mathrm{H}} V_{\mathrm{N}} (T_s/\Pi_s - T_{\mathrm{N}}/\Pi_{\mathrm{N}})$ and latent $LE = L\rho_a c_{\mathrm{H}} V_{\mathrm{N}} (q_{sat}(T_s) - q_{\mathrm{N}})$ heat in the case of bare ice, or the heat flux from snow to ice in the case when snow is present.

The term Q in the first row of Eq. (1) describes the heat flux from solar radiation penetrating into the ice pack. ~~For calculation of this heat flux ,~~ This heat flux is calculated by using the Bouguer-Lambert law~~is used~~, with an approximation of radiation absorption in the thin layer of the ice following Grenfell and Maykut (1977):

$$Q(z) = \delta_{H_{snow}} (1 - \alpha) SW\!\downarrow i_0 \cdot e^{-k \cdot z} \tag{4}$$

where α is the ice albedo; $SW\!\downarrow$ is the downward solar radiation flux ($\mathrm{W\,m^{-2}}$); i_0 is the fraction of radiation penetrating through the thin layer of the ice and k is the extinction coefficient for the ice ($\mathrm{m^{-1}}$), which is parameterized according to values suggested by Grenfell and Maykut (1977). The value i_0 parameterizes the vertical inhomogeneity of the ice transparency ~~. It~~ and is dependent on depth~~: it~~. It is equal to 1 in the uppermost 0.1 m layer of ice, and equal to 0.18 in the lower layers (Grenfell and Maykut, 1977). Note that in the case of snow on ice, the remaining solar radiation that was not absorbed during penetration through the snow pack is assumed to be completely absorbed by the underlying ice surface.

The main prognostic variable of the SICE scheme is the temperature of the ice. Other parameters are either physical constants or should be taken from the external forcing. For calculation of the ice thermal conductivity and heat capacity we used the following formulations, which represent their dependency on the ice temperature and salinity (Schwerdtfecer, 1963; Feltham et al., 2006; Sakatume and Seki, 1978):

$$C = C_0 - \frac{T_{mlt}(S) - T_{mlt}(0)}{\theta^2} L \tag{5}$$

$$\lambda = \lambda_{bi} - (\lambda_{bi} - \lambda_b) \frac{T_{mlt}(S) - T_{mlt}(0)}{\theta} \tag{6}$$

where, following Bailey et al. (2010):

$$T_{mlt}(S) = 273.15 - 0.0592S - 9.37 \cdot 10^{-6}S^2 - 5.33 \cdot 10^{-7}S^3$$

$$\lambda_{bi} = \frac{2\lambda_i + \lambda_a - 2V_a(\lambda_i - \lambda_a)}{2\lambda_i + \lambda_a + 2V_a(\lambda_i - \lambda_a)}\lambda_i$$

$$\lambda_i = 1.162\left(1.905 - 8.66 \cdot 10^{-3}\theta + 2.97 \cdot 10^{-5}\theta^2\right)$$

$$\lambda_b = 1.162\left(0.45 + 1.08 \cdot 10^{-2}\theta + 5.04 \cdot 10^{-5}\theta^2\right)$$

$$\lambda_a = 0.03 \qquad V_a = 0.025$$

and C_0 is the volumetric heat capacity of the fresh ice ($\mathrm{W \cdot s\ m^{-3}K^{-1}}$); $T_{mlt}(S)$ is a function of the melting point of the saline ice depending on the salinity (K); S is the ice salinity, parts per thousand; $\lambda_{\{i,b,bi,a\}}$ is the heat conductivity of fresh ice, brine, bubbly ice and air respectively ($\mathrm{W\ m^{-1}}$); V_a is the fractional volume of air in the sea ice; and θ is the ice temperature in °C.

In the case of bare ice (no snow), information about the ice albedo is needed to calculate the surface energy balance from Eq. (2). The ice albedo strongly affects the temperature regime of the ice pack. The effects of some processes taking place on the ice surface, such as the effect of melt ponds, may be parameterized through the ice albedo even without their real physical description. In the SICE scheme, we propose a choice between several options based on simple parameterizations of ice albedo from (Perovich, 1996; Parkinson and Washington, 1979; Roeckner et al., 1992). In these parameterizations, albedo is defined as a constant value or as a function of the ice surface temperature. Numerical methods to solve Eq. (1) and Eq. (2) are presented in Appendix A.

The assumption of bare ice is the simplest possible approximation and may give reasonable results. However, such a simple parameterization ~~would describe~~ describes processes on the ice surface covered by snow in a very approximate way. Snow upon the ice serves as an insulating layer with higher albedo and lower thermal conductivity than the underlying ice~~, and for~~. For more physically correct simulations the ice scheme should reproduce the processes related to the evolution of snow on the ice surface. The form of the upper boundary condition presented by Eq. (1), which contains the heat flux from the snow layer to the ice layer F as a term, allows easy coupling with an external snow model to represent snow on ice. In our study, we used the snow module ISBA ~~Explicit Snow (ISBA-ES) (Boone and Etchevers, 2001; Boone, 2000)~~ (Interactions Surface Biosphere Atmosphere) Explicit Snow (ISBA ES, Boone, 2000; Boone and Etchevers, 2001) to represent processes in snow. In the current version of SICE, when snow pack exists, it always covers the ice part of the grid cell as a layer of uniform thickness.

ISBA ES is a multi-layer snow scheme with prognostic snow water equivalent, snow heat content and snow density. The number of layers may be defined by the user ~~with a~~ (the default value of ~~3.~~ 3). The uppermost snow layer is always less than or equal to 0.05 m. The scheme ~~describes explicitly~~ explicitly describes the following processes: snow accumulation due to precipitation, heat redistribution, melting processes and snow pack compaction. It also represents processes related to the melt water within a snow layer. The heat diffusion and surface energy balance equations are solved numerically with implicit schemes. The snow module needs information about the atmospheric forcing and the temperature, heat conductivity and thickness of the topmost layer of ice. It predicts the snow variables and also provides the flux from the snow pack to the underlying medium. Thus, the coupling between snow and ice schemes is explicit. The snow surface albedo in ISBA ES is calculated through a simple aging scheme, which covers dry- and wet-snow albedo degradation formulations. In this aging scheme, the snow albedo may decrease during the degradation process from its maximum value of 0.85 to a minimum value of 0.5. When applying this scheme over sea ice, a snow albedo minimum value of 0.75 is used following ~~(Perovich, 1996; Semmler et al., 2012)~~Perovich (1996); Semmler et al. (2012). The ISBA ES scheme ~~was developed to parameterize~~ parameterizes snow over land surfaces ~~, thus it~~ and contains no parameterizations of specific snow-over-ice processes, such as snow-ice formation or evolution of melt ponds.

~~For better representation of the surface processes, an~~ An atmospheric model may apply a tiling approach for better representation of the surface processes. This means that a model grid cell may be covered with a mixture of both sea water and ice. When a grid box contains a certain amount of ice and open water, the ice and open water calculations are performed independently and the output flux to the atmosphere is represented by a weighted average of the fluxes from the water and ice parts. In this case, information about the ice concentration may be utilized to obtain the weighting coefficients. Ice concentration is estimated from satellite observations using the analysis procedure. This procedure contains a consistency check between the sea surface temperature and sea ice concentration fields (Stark et al., 2007; Donlon et al., 2012).

Turbulent exchange between the sea ice and the atmosphere is a complex process that is influenced by the morphological features of the ice pack such as the presence of melt ponds, ice topography, ridges and deformations. In the current version of the SICE scheme these complex features are not represented and the ice part of the grid cell is assumed to be a flat surface with uniform characteristics. When the sea ice concentration is less than 100 % (which means a mixture of open water and sea ice), one more factor influences the turbulent exchange with the atmosphere. This is the form drag, which is caused by the floes floating on the water with their upper edge higher than the water surface (ice obstacles). This subtle effect might be important for NWP because the ice concentration from observations is considered in NWP as a percentage of ice in a grid cell. Indeed, the roughness length of water is lower than that of ice, and simple weighted averaging according to the ice concentration values will lead to a decrease of the roughness length (compared to fully ice-covered area), while in nature it should increase. An accurate sea ice scheme should include a parameterization of the drag caused by ice obstacles (the form drag). Such schemes, discussed for example in Lüpkes et al. (2012), usually introduce an additional term in the weighted average, which depends on the ice fraction. The form drag was introduced into the SICE scheme in the following way:

$$C_{d,mean} = \eta C_{d,ice} + (1-\eta) C_{d,sea} + C_{d,f} \tag{7}$$

where $C_{d,\{ice,sea,mean\}}$ is the drag coefficient over ice, sea and the mean drag coefficient over the grid cell respectively, under neutral conditions; η is the fraction of sea ice in the grid cell; and $C_{d,f}$ is the form drag. For calculation of the form drag term a parameterization suggested by Lüpkes et al. (2012) was used:

$$C_{d,f} = 7.68 \cdot 10^{-3} \left[\frac{\ln(0.41/z_{0,w})}{\ln(10/z_{0,w})} \right]^2 (1-\eta)^{\beta} \eta \tag{8}$$

where $z_{0,w}$ is the roughness length of the sea water surface and β is the tuning constant. Parameters of the SICE scheme are summarized in Table 1.

Technically, the SICE scheme was developed as a part of the externalized land and ocean modelling platform SURFEX (Masson et al., 2013). The externalized surface modelling platform SURFEX is a set of models used for the description of different types of surfaces: sea and inland water bodies, soil/vegetation and urban environments. It assumes a tiling approach, distinguishing different surface types ~~in one~~ within one grid box of an atmospheric model~~grid~~. Each atmospheric model grid box contains some fraction of 4 different surface types (tiles): nature, urban, inland water and sea. Fractions of these tiles are ~~actually model parameters , they are permanent and~~ permanent model parameters known from land-use maps. For land-use mapping (physiography) SURFEX incorporates the 1 km resolution database ECOCLIMAPII (Faroux et al., 2013). Over a sea tile, in turn, some fraction of sea ice may exist. This fraction is constant during the forecast run, but it changes at the moment of analysis (model initialization) according to the sea ice concentration estimated from observations. Thus, sea ice may be considered as sub-tile (or patch). The functionality of using main tiles is provided by SURFEX, but the possibility to use information about the fractional sea ice was introduced into SURFEX while implementing the SICE scheme. SICE utilizes the standard heat diffusion equation solver from the SURFEX suite. SICE, which will in future contain a more advanced description of the sea ice, provides technical compatibility with the developing versions of SURFEX. ~~It is important to mention that SURFEX provides the possibility to diagnose screen level meteorological values (2 metre~~

SURFEX provides diagnostic screen level temperature and specific humidity and 10 metre wind speed ~~)~~ from the predicted surface state and the atmospheric values (provided that the forcing is given at some upper level or at the lowest level of the host model) using the interpolation-like procedure of Businger et al. (1971).

The modelling platform SURFEX is incorporated into the ALADIN-HIRLAM NWP system to parameterize the underlying surface processes. The ALADIN-HIRLAM NWP system includes the model configuration HARMONIE-AROME (Bengtsson et al., 2017), which is a version of the non-hydrostatic limited area atmospheric model AROME (Seity et al., 2011). A variety of sub-grid scale physical processes are taken into account by the model parametrization schemes. In the ALADIN-HIRLAM NWP system, boundary conditions and some initial conditions are taken from larger scale models, such as IFS ~~(ECMWF)~~ or HIRLAM. The ALADIN-HIRLAM NWP system contains a data assimilation system that uses the three-dimensional variational analysis (3DVAR) method for upper air. Data assimilation of the surface variables uses the optimal interpolation method for snow depth and screen level temperature ~~, relative humidityand snow depth~~and relative humidity. In the configuration of the system used in this study, variables in the soil are initialized according to the optimal interpolation method described in Mahfouf et al. (2009).

To produce forecasts, HARMONIE-AROME performs short-term cycles, each cycle contains the data assimilation procedure and the model forecast. The background fields for the data assimilation are fields of prognostic variables at the end of the previous model forecast. In the configuration of HARMONIE-AROME used in this study, the length of the cycle was 3 hours. Starting from 0000 UTC and 1200 UTC analysis times, longer forecasts (up to 48 hours) are performed. Each cycle, the sea water surface temperature and the fraction of sea ice are are kept constant during the forecast being interpolated bilinearly from the host model IFS-HRES, with extrapolation by the nearest neighbour method in specific areas such as fjords. The same is done with the initial value of the sea ice surface temperature. In turn, IFS-HRES uses OSTIA data (Donlon et al., 2012) for the sea surface temperature and ice fraction. For the ice surface temperature, it runs its own simple ice model.

Prior to implementation of the SICE scheme into the ALADIN-HIRLAM NWP system (through SURFEX), sea ice was accounted for in a very crude way in the HARMONIE-AROME configuration. The sea ice surface temperature was initialised by values modelled by IFS-HRES. The ice surface temperature was simply equal to its initial value through the whole forecasting period (similar to the sea surface temperature); this introduced large errors, mainly connected to the absence of a diurnal cycle over the ice surface. When the SICE scheme is used, its prognostic variables are updated during each cycle as described in the following. If in the grid cell in question the ice cover exists in the background field, the prognostic variables of SICE are kept unchanged (SICE runs freely). Otherwise, in the situation when the new ice is observed according to OSTIA, the initial (analysed) values of the prognostic SICE variables are obtained via extrapolation from the nearest grid cells of the background field, where the ice exists.

3 Performance of the sea ice parameterization scheme

The main objective of the ~~simple sea ice parameterization~~ SICE scheme is to reproduce the evolution of ice surface temperature ~~, since~~ because this variable provides an interface between the atmosphere and the underlying surface. Observations of sea ice surface temperature in the area of interest, which may be used to evaluate the performance of the model/parameterization scheme, are limited. ~~For testing the SICE scheme as a part of the ALADIN-HIRLAM NWP system, we compared the modelling results with the screen-level atmospheric observations. Visual comparisons with sea ice surface temperature observations from MODIS were also performed. Before verification against observations~~First, we compared modelling results from SICE with the results of the well tested sea ice model HIGHTSI (Cheng and Launiainen, 1998) ~~. The purpose of this comparison was the~~ as an overall technical sanity check of SICE and ~~better understanding of~~ to better understand its limitations and weaknesses. Then we compared the modelling results with the screen level temperature and 10 metre wind speed observations from SYNOP stations to test the SICE scheme as a part of the ALADIN-HIRLAM NWP system. Comparisons against sea ice surface temperature products from MODIS and VIIRS were also performed.

3.1 Preliminary experiments comparing SICE and HIGHTSI results

HIGHTSI is a one dimensional thermodynamic sea ice model, which was developed for research purposes and climate studies. Although HIGHTSI does not contain the ice dynamics (unlike CICE and GELATO), it reproduces the temperature profiles in the ice with a sufficient level of accuracy (Cheng et al., 2008) and needs a minimal amount of forcing data. The model describes the evolution of ice mass and energy balance and is based on the heat conduction equation, which is solved with an implicit finite difference numerical scheme (Launiainen and Cheng, 1998). Parameterization of snow in HIGHTSI includes processes of snow accumulation from the forcing precipitation, snow melting and refreezing, and snow-ice formation (which in our experiments was switched off). Comparison with HIGHTSI was carried out in off-line (stand-alone) mode, since HIGHTSI is not coupled with an atmospheric model.

For the atmospheric forcing we used ~~data~~ the following variables from HIRLAM (Unden et al., 2002) operational forecasts (with a horizontal spatial resolution of 8 km)~~, namely~~: lowest model level air temperature, wind speed and specific humidity;

surface pressure; global downward shortwave and downward longwave radiation fluxes at the surface; rainfall and snowfall rates. Stand-alone experiments were performed for the 12 selected synoptic stations in the Svalbard coastal area~~, see Fig. 1~~. The period of off-line experiments was from August 2011 to June 2012, with a temporal resolution of the forcing data of one hour. For each model, ~~2~~ two experiments were performed: a snow-free experiment, and an experiment considering the evolution of snow. For ice albedo, a simple parameterization based on Roeckner et al. (1992) was used. The ice salinity was set to a uniform value of 3 ppt.

In the SICE scheme, the prescribed ice thickness was given a value of 0.75 m, with 4 layers in the ice slab and 3 layers within the snow. HIGHTSI was configured in the default way with 20 layers within the ice slab and 10 layers within the snow pack. The first month of the simulations was considered as a spin-up.

~~We analyse here the results of the experiments only for the period when the ice thickness modelled by HIGHTSI was approximately equal to~~ Stand-alone experiments showed that ice surface temperatures modelled by SICE and HIGHTSI may differ by more than 5 °C when difference of the ice thickness ~~prescribed in SICE. In this way we see~~ in the two models is greater than 0.4 m. Thus, to analyse the difference in reproducing the thermal regime in the ice between the two schemes ~~. Here we consider “small” ice thickness difference between the two models to be~~ we consider only the period when difference of ice thickness in SICE and HIGHTSI is less than 0.4 m. In “no-snow” experiments, ~~the period of small ice thickness difference~~ this period lasted approximately for 3 months, from mid-September to mid-December. ~~“No-snow~~And for this period “no-snow” experiments showed that SICE and HIGHTSI tend to produce similar results~~for the period of the small ice thickness difference. When the ice thicknesses are very different, the ice surface temperature values may differ by more than 5 °C~~. In the experiments with the snow schemes included, the evolution of the snow thickness was quite similar in HIGHTSI and SICE. Due to the presence of snow in these experiments, the ice thickness in HIGHTSI was lower~~, so that~~ . This led to the period of ~~small ice thickness difference lasted~~ the thickness difference less than 0.4 m lasting from mid-September to the end of June. When the ice surface is insulated by snow and only the thin snow layer reacts to the atmospheric forcing, the oscillations of the snow surface temperature are very large. Due to this high variability, the snow surface temperature was sometimes very different between the HIGHTSI and SICE experiments, with a maximum difference of 3–5 °C. ~~In general, the~~ The mean value of the surface temperature difference between SICE and HIGHTSI for all 12 points in the “no-snow” experiments was 0.71 °C (SICE gave higher values than HIGHTSI), with a difference standard deviation of 1.04 °C. For “snow” experiments these values are -0.46 °C (SICE gave lower values than HIGHTSI) ~~, with a standard deviation of~~ and 1.99 °C respectively.

~~From these preliminary experiments with~~ These results of stand-alone ~~runs, we conclude~~ experiments show that the SICE scheme is able to adequately reproduce the evolution of the ice surface temperature, ~~but~~ however the result is sensitive to the value of the prescribed ice thickness. Thus, the ice thickness may be important even if the main focus of the simulations is the ice surface temperature. ~~However this~~ Although the approach with the prescribed ice thickness ~~, although being~~ is very simplified, it may reproduce the ice surface temperature oscillations of different time scales and serve as a first approximation for the description of the sea ice cover behaviour.

3.2 ~~Experimental configuration~~ Design of experiments with the ALADIN-HIRLAM NWP system

For coupled experiments, the HARMONIE-AROME model configuration (Bengtsson et al., 2017) of the ALADIN-HIRLAM NWP system ~~was used. Prior to implementation of the SICE scheme into the ALADIN-HIRLAM NWP system (through SURFEX), sea ice was accounted for in a very crude way. The ice surface temperature was simply equal to its initial value through the whole forecasting period (similar to the sea surface temperature); this introduced large errors, mainly connected to the absence of a diurnal cycle over the ice surface. To produce forecasts, HARMONIE-AROME performs short-term cycles, each cycle contains the data assimilation procedure and the model forecast. The background for the data assimilation are fields of prognostic variables at~~ described in the end of ~~the previous model forecast. In the configuration of HARMONIE-AROME used in this study, the length of the cycle was 3 hours. Starting from 0000 UTC and 1200 UTC analysis times, longer forecasts (up to 48 hours) are performed. Each cycle, the sea water surface temperature and the fraction of sea ice are interpolated bilinearly from the host model IFS, with extrapolation by the nearest neighbour method in specific areas such as fjords. The same is done with the initial value of the sea ice surface temperature. In turn, IFS uses OSTIA data (Donlon et al., 2012) for the sea surface temperature and ice fraction. For the ice surface temperature, it runs its own simple ice model. Thus, without SICE, HARMONIE-AROME actually usedthe sea ice surface temperature values modelled by IFS.~~

~~If the SICE scheme is used, its prognostic variables are updated during each cycle as described in the following. If in the grid cell in question the ice cover exists in the background field, the prognostic variables of SICE are kept unchanged (SICE runs freely). Otherwise, in the situation when the new ice is observed according to OSTIA, the initial (analysed) values of the prognostic SICE variables are obtained via extrapolation from the nearest grid cells of the background field, where the ice exits. The fraction of sea ice is also interpolated (or extrapolated) from OSTIA.~~

Section 2 was used. For this study, HARMONIE-AROME experiments were performed over two operational domains (see Fig. 1): (A) the AROME Arctic domain, which includes large ice-covered areas in the Arctic ocean, and (B) the MetCoOp domain, where the ALADIN-HIRLAM NWP system with HARMONIE-AROME is run operationally in a cooperation between the Norwegian Meteorological Institute and the Swedish Meteorological and Hydrological Institute[1] (Müller et al., 2017b), and which covers the Scandinavian peninsula and the Baltic sea. Grids over both domains have a horizontal spatial resolution of 2.5 km. Experiments cover the time period from March to April of 2013. ~~Four~~ Five experiments defined for this part of the study are summarized in Table 2. These are: the reference experiment (REF) without the SICE scheme, SICE experiments without and with the ISBA ES snow module (SICE2D-NS and SICE2D-S respectively), ~~and a~~ SICE experiment with the form drag parameterization included (SICE2D-AD), and the SICE2D-NS-CLIM experiment which uses the model climatology of the ice thickness provided by TOPAZ4 reanalysis (Sakov et al., 2012; Xie et al., 2017). The following SICE configuration was used in the experiments: ~~the prescribed thickness of the ice pack was given the value 0.75 with~~ 4 layers in the ice, the ice albedo was calculated based on Roeckner et al. (1992). For the SICE2D-AD experiment the coefficient β in Eq. (8) was set to a value of 1. The experiments SICE2D-S~~and~~ , SICE2D-AD and SICE2D-NS-CLIM were only run over the Arctic domain. In SICE2D-S, the default 3 layer configuration of the snow scheme was chosen. The first cycle of the SICE2D-S experiment

[1]The Finnish Meteorological Institute joined the MetCoOp collaboration in September 2017 and MetCoOp domain was extended towards the east.

started from the snow-free state and accumulated snow for the next cycles the initial snow fields were taken from the previous cycle's 3 hour forecast. Snow was accumulated from the precipitation during the whole modelling period. The ice fraction was taken into account in all SICE experiments. The sea ice fraction was the only sea ice variable that was influenced by observations in the analysis procedure.

3.3 Experiments with the SICE scheme included in the ALADIN-HIRLAM NWP system: validation against meteorological observations

For the evaluation of the The relative impact of the SICE scheme in terms of the root mean square errors (RMSE) of the mean sea level pressure, 2 metre temperature and 10 metre wind speed forecasts starting from 0000 UTC for the time period from 1 March 2013 to 30 April 2013 for all Norwegian national weather stations and SYNOP weather stations within the AROME-Arctic domain is summarized on the Fig. 2 and Figs. S1-S3. Number of cases is approximately 960 for each station. Experiments with SICE compared to the REF experiment show no considerable differences except coastal stations surrounded by sea ice.

To evaluate the model performance, measurements from 12 Svalbard stations and 6 we selected stations in the area of Svalbard archipelago and stations situated in the coastal area of the Gulf of Bothniawere used. These stations are strongly affected by the sea and may show the impact of the improved ice representation. Some . According to Fig. 2 and Figs. S1-S3, modelling results from these stations show the largest relative changes in RMSE when using SICE. To emphasize the effects of using SICE we excluded from the comparison coastal stations that were always surrounded by open sea in March and April 2013. Some of selected Svalbard stations are located in fjord areas where the forecast is strongly dependent on the quality of the ice fraction field. Due to the low resolution of the original ice fraction data and a too crude extrapolation procedure, for some ice-covered fjords only open water existed in the model runs. Stations located in such fjords were also excluded from the comparison. The final set of SYNOP stations considered for the comparison consists of 7 stations in Svalbard archipelago and 7 stations in the Gulf of Bothnia. Locations of the selected stations are shown on the Fig. 1.

Figures 3, 4, 5 and 6 show the impact of the new sea ice scheme, including the representation of snow on ice and form drag. These figures show the statistics of the forecast errors obtained by sampling the forecasts starting from 0000 UTC during the period of the experiments for the selected groups of points in the Svalbard and Bothnian areas. For the statistics, the mean forecast error (bias)and the forecast error standard deviation , the root mean square error (RMSE) and the standard deviation of errors (ESTD) as a function of the forecast lead time for the mean sea level pressure, 2 metre temperature and 10 metre wind speed were calculated for various experiments. Note that statistics for REF and SICE2D-NS in Fig. 5 and Fig. 6 are different because they cover different time periods: the period of March-April of 2013 for Fig. 5 and the period of only March 2013 for Fig. 6.

The main impact of the SICE scheme is seen in the scores for the 2 metre temperature. Figure 43b shows that in REF, over the Svalbard stations the 2 metre temperature forecasts have a negative bias increasing in absolute value with the forecast length from 0.5 °C up to 2 °C. This evolution is caused by the influence of the surface temperature over the sea (both the open water and ice cases), which remains constant during the whole forecast period in this experiment. For the Bothnian stations (see

Fig. ~~5band Fig. 5d~~3b, right panel) in REF, the 2 metre temperature mean error has a diurnal cycle. This is because ~~the Bothnian stations are much more affected by the land than the Svalbard stations, andthe underestimation of night-time~~ in the REF experiment for a cycle starting at 0000 UTC the ice surface temperature is initialized from IFS-HRES forecast and represent the cold night time ice surface. This temperature is in a good agreement with reality and, for the night time, bias in the 2 metre ~~temperatures over land is a characteristic feature of the model known from operational verification (not shown)~~temperature is relatively small. After 12 hours of forecast, during the day time, sea ice grid cells still hold these very low temperatures and that leads to considerable negative bias in the 2 metre air temperature. The situation is illustrated by Fig. 7, which represents the observed values of air temperature for Kemi I lighthouse (WMO No. 02863, station position: 65°25' N; 24°08' E) and the forecast time series of different length and starting time. It shows that for REF, the air temperature can be more than 5 °C lower in the model forecast than in reality.

The sea ice scheme allows the ice surface temperature to evolve in time and improves the 2 metre temperature forecasts. According to Fig. ~~4~~3b, over Svalbard stations the 2 metre temperature bias for SICE2D-NS and for SICE2D-NS-CLIM is smaller than for REF, it is now positive and has an almost constant value of 1 °C. For the Bothnian stations (see Fig. ~~5band Fig. 5d~~3b) the bias in SICE2D-NS and SICE2D-NS-CLIM still has a diurnal cycle, but now the night-time errors are much smaller, only 1 °C in absolute value. The ~~error standard deviation for~~ RMSE and ESTD of the 2 metre temperature forecasts ~~is~~ are also considerably smaller in SICE2D-NS and SICE2D-NS-CLIM compared to REF, especially for forecasts longer than 24 hours. For the Svalbard stations ~~it is~~ (see Fig. 4b) they are more than 4 °C in REF but only 3 °C in SICE2D-NS and SICE2D-NS-CLIM, and for the Bothnian stations these values (see Fig. 5b) are 3 °C and 2 °C (note that for the Bothnian stations the ~~forecast error standard deviation~~ standard deviation of forecast errors also shows a diurnal cycle). Smaller ESTD and RMSE means that in case of SICE experiments not only systematic, but also the random component of the forecasting error was reduced comparing to REF. Experiment SICE2D-NS-CLIM comparing to SICE2D-NS gives slightly better results in terms of the 2 metre temperature forecast biases for Svalbard stations and slightly worse for Bothnian stations (see Fig. 3b). The difference in biases between these two experiments is approximately 0.2 °C and might be not statistically significant. This is because the default sea ice thickness of 0.75 m in the SICE2D-NS is very close to the climatology in the coastal areas.

Although the mean sea level pressure is usually controlled by the large scale rather than local processes, a local positive impact for this field is also visible, especially for the Svalbard stations. Figure ~~4~~3a shows that the ~~mean sea level pressure~~ positive bias of up to 0.5 hPa of mean sea level pressure forecasts in REF is removed in SICE2D-NS ~~. This happens~~ and SICE2D-NS-CLIM. This occurs due to warmer (in general) temperatures in ~~SICE2D-NS~~these experiments. In terms of ESTD of the mean sea level pressure ~~error standard deviation~~ forecasts (and since the bias is small, also RMSE), there is no considerable difference between REF~~and~~ , SICE2D-NS and SICE2D-NS-CLIM for the Svalbard stations (see Fig. 4a). For the Bothnian stations, there is no considerable difference between REF~~and~~ , SICE2D-NS and SICE2D-NS-CLIM experiments for both the mean error ~~and error standard deviation~~ (Fig. 3a), ESTD and RMSE (Fig. 5a) of the mean sea level pressure forecasts.

For the 10 metre wind speed~~the experiment~~ , in REF bias is positive, with values between 0.1 and 0.5 $\mathrm{m\,s^{-1}}$. The experiments SICE2D-NS ~~has~~ and SICE2D-NS-CLIM have an approximately 0.5 $\mathrm{m\,s^{-1}}$ higher mean error than REF for all forecast lengths ~~. In REF, the 10 metre wind speed bias is positive, with a value between 0.1 and 0.5 , but in SICE2D-NS it is even larger~~ (see

Fig. ~~4c~~3c, left panel). The source of the ~~increased wind speed~~ larger wind speed in SICE2D-NS and SICE2D-NS-CLIM is the absence of the form drag over fractional ice ~~in SICE2D-NS~~these experiments. In REF the sea-related part of the grid cell may have only two states: either covered by open water or by ice. As a result, in this experiment all of the Svalbard stations are affected by the surrounding compact ice areas and the simulated wind speed at these points depends on the ice roughness length. In SICE2D-NS and SICE2D-NS-CLIM, the stations are surrounded by a mixture of ice and open water. In this case the average drag coefficient for momentum over ~~the part of the grid cell related to the sea with~~ a grid cell that contains both open water and sea ice is smaller than in REF, since the roughness of a water surface is much lower than that of ice. This leads to ~~the higher wind speed values~~higher wind speeds. Thus, ~~this increased~~ the large positive bias in SICE2D-NS and SICE2D-NS-CLIM is the effect of averaging the drag coefficients for the sea-related part of the grid cell between the drag coefficients over open water and ice. They are averaged with weights according to the ice fraction value and the form drag is not taken into account.

In SICE2D-AD, the form drag is taken into account in the SICE scheme. In this experiment, we add the form drag only when calculating the momentum flux. The effect of the form drag term is shown in Fig. 8, displaying the difference between the drag coefficients calculated in an ordinary way and with the additional term. The impact of the form drag is mostly noticeable in areas near the ice edge, where the ice fraction field has values of around 60 %. This is in ~~accordance~~ agreement with Elvidge et al. (2016); Tsamados et al. (2014); Lüpkes et al. (2012). In SICE2D-AD the wind speed bias is smaller than in SICE2D-NS and is just slightly larger than in REF, as shown in Fig. 6. This improvement is seen both for the Svalbard and Bothnian stations~~; for the Svalbard stations~~ , although it is more pronounced ~~, mainly~~ for the Svalbard stations due to the differences in the ice concentration fields around Svalbard and in the Baltic sea. However, in the SICE2D-AD experiment the sample volume (number of forecasts) might be not large enough to make the statistically significant conclusions, especially for the group of Bothnian stations, due to the short experiment length. The error statistics for the other fields are not deteriorated in SICE2D-AD compared to SICE2D-NS (not shown).

In the discussion above, we compared ~~only~~ the snow-free experiments, where the physical processes over ice are ~~still~~ represented very roughly. A more advanced modelling system should also simulate the snow layer on top of the ice pack. In SICE2D-S, the explicit snow scheme ISBA ES is used to represent the snow over ice. ~~In terms of the~~ The 2 metre temperature ~~errors, the results of SICE2D-S are worse~~ forecast errors are larger for SICE2D-S than for SICE2D-NS ~~, the~~ and SICE2D-NS-CLIM. The bias in SICE2D-S for the Svalbard stations is almost the same as in REF (see Fig. ~~4~~3b). For the Bothnian stations, the bias in SICE2D-S is smaller than in REF, but still larger than in SICE2D-NS ~~(see Fig. 5)~~and SICE2D-NS-CLIM. Also, a shift in the bias diurnal cycle in SICE2D-S compared to SICE2D-NS and SICE2D-NS can be seen. This shift is caused by the difference in the thermal resistances of the snow ~~and ice~~ in the SICE2D-S ~~and~~ experiment and ice in SICE2D-NS and SICE2D-NS-CLIM experiments. The ~~error standard deviation~~ ESTD and RMSE of the 2 metre temperature forecasts in SICE2D-S is also larger than in SICE2D-NS and SICE2D-NS-CLIM, but it is still smaller than in REF, both for the Svalbard and Bothnian stations, especially for longer lead times (see Fig. 4b and Fig. 5b). These results are in ~~accordance~~ agreement with the off-line experiments. These cold 2 metre temperatures in the SICE2D-S experiment may be caused by different reasons. When conditions in the atmospheric boundary layer are stable, the cold surface becomes decoupled from the atmosphere, and a positive feedback appears, which ~~makes~~ induces further drop of the surface temperature~~drop more and more~~. This situation is very difficult to reproduce in modelling. Moreover, model errors may be both positive or negative. This may depend on errors in boundary layer parameterization, radiation fluxes, snow density or precipitation. This complex problem is well explained e.g. in ~~(Slater et al., 2001)~~Slater et al. (2001). In atmospheric modelling it is usually called "the stable boundary layer problem", because it appears during the periods of low shortwave radiation, cooling surface and near-surface inversions. In Atlaskin and Vihma (2012) it is shown how this problem appears in different NWP systems. Also, errors in the amounts of snow accumulated by the model may affect the quality of the screen level temperature forecast. For example, a caveat of the current scheme is the absence of the snow-ice formation representation, which could be important in the case of a thick snow layer covering relatively thin ice. Parameterization of these effects would require description of the ice mass balance, which is not implemented in the current version of SICE. In addition, errors in the snow depth and snow water equivalent over the ice are not corrected by the snow data assimilation procedure, as ~~happens~~ occurs over land.

Validation against coastal SYNOP observations allows the impact of the sea ice temperature evolution to be understood on the local scale, which is the main concern of regional NWP models. However with observations from coastal stations only, we lack understanding of the ice temperature behaviour ~~for larger scales~~over large sea ice covered areas.

3.4 Comparisons with observations from MODIS and VIIRS

~~The purpose of the comparison with observations from MODIS was to assess~~ Although comparison of different model experiments with data from SYNOP stations could give us an indirect estimate of the performance of ~~SICE over large ice-covered areas, especially in the Arctic region~~the new ice scheme using the forecast scores, it does not provide much information about the actual quality of representation of the sea ice cover. Data from ice mass balance buoys or manned drifting ice stations is a valuable source of in-situ measurements of the sea ice, but these data represent local conditions of the sea ice field and only very few of them are located within the AROME Arctic domain. Therefore, the large scale performance of the sea ice scheme could be better assessed by using remote sensing data.

In the current study the ice surface temperature products from MODIS (Terra and Aqua satellites) and VIIRS (Suomi NPP satellite) sensors were used to verify the performance of the SICE scheme. The satellite observations of sea ice surface temperature ~~(Hall and Riggs, 2015; ?)~~ (Hall and Riggs, 2015; Tschudi et al., 2017) were retrieved from the archives of the NASA Land Processes Distributed Active Archive Center (LP DAAC)~~for the thermal remote sensing sensor MODIS of the Terra satellite~~. The resolution of the data is approximately 1 km for ~~swathe data and 4 for daily fields. We used 5-minute images of different satellite swathes, attempting to consider the diurnal cycle of the sea~~ MODIS swathes and 750 m for VIIRS swathes.

MODIS and VIIRS ice surface temperature products contain gaps due to cloudiness, which decrease the number of valid data points from a single swathe considerably. To reduce uncertainties caused by mismatch between the cloud masks used to generate MODIS and VIIRS products and cloud cover predicted by HARMONIE-AROME, only cloud-free grid cells (both from the point of view of the model and remote sensing data) were considered for comparison with satellite products. Pixels of MODIS and VIIRS products that were reported as having quality other than 'best' or 'good' by the quality assessment procedure were excluded from the comparison. Study area for each specific date was selected according to the 'closed ice' map provided by OSI SAF ice edge product (Aaboe et al., 2017) to exclude the marginal ice zone, open sea and coastal regions.

This ice edge product uses a threshold value of 0.7 for the ice fraction product of OSI SAF to separate 'closed ice'. In our study, modelling results and ice surface temperature . Since the main focus was the large scale, we used a visual method for comparing maps. Statistical assessments require application of special methods, which were out of the scope of this study. Since MODIS observations use the optical band, they contain gaps due to cloudiness, which decrease products were compared with each other only within this 'closed ice' zone.

Usually for validation of the model against observations, the model data are interpolated to the observational points. But in our case this is impossible, because the resolution of the remote sensing data is finer than of the model grid. Moreover, the locations of the pixels of the remote sensing products are different for a two different swathes. That is why in our comparisons we first aggregated the remote sensing data on the atmospheric model grid, and then referred the model errors to the locations of model grid boxes. Also, we used all available swathes with a time stamp within the one hour window for a given forecast lead time. All this makes the estimates of statistical significance to be complicated, because the number of useful maps considerablycases varies a lot depending on availability of swathes, their spatial location and the cloud-covered area. Therefore, here we provide only general statistics, leaving the details for the future studies.

First, we compared the results of model experiments described in Sec. 3.3 with MODIS data. During the period of these experiments (which is March-April 2013), the majority of satellite swathes within the area of interest providing and VIIRS data. Figure 9 shows the mean bias, RMSE and ESTD of the ice surface temperature data were day-time. Thus, from these data it was impossible to estimate the amplitude of the diurnal cycle forecasts starting from 0000 UTC for REF, SICE2D-NS, SICE2D-S and SICE2D-NS-CLIM experiments averaged over the whole study area calculated from MODIS and VIIRS products depending on the forecast lead time. The spatial distribution of the mean error and ESTD of the ice surface temperature . In the model experiments, this amplitude in the polar Arctic regions was also not large during this period, approximately 8 °C for both the after 24 hours of forecast started at 0000 UTC in REF, SICE2D-NSand , SICE2D-S experiments. In the REF experiment , the diurnal cycle of the sea and SICE2D-NS-CLIM experiments calculated over the experiment period by using MODIS product is shown on the Figs. 10 and 11 . Figures S4 and S5 provide the same maps but using VIIRS ice surface temperature was not reproduced within a single forecast because the ice temperature was prescribed. However, in general , product for verification. It can be seen from Fig. 9a that biases of the ice surface temperature forecasts in general are highest for the SICE2D-NS experiment and lowest for SICE2D-S experiment. Results of the REF experiment in comparison with lie between the extremes rendered by SICE2D-NS showed ice surface temperature values closer to those observed by MODIS for the first 24 hours of forecast.In and SICE2D-S. From the Fig. 9a, the experiment SICE2D-NS-CLIM gives better results than SICE2D-NS, but still worse than REF. However from the Fig. 10 we may see that in the polar Arctic, the maximum area averaged ice surface temperature simulated by SICE2D-NS was approximately 10bias is approximately 6 °C higher than observed by MODISin SICE2D-NS, while in REF is was only 5SICE2D-NS-CLIM and REF the biases are smaller, approximately 4 °Chigher. This happens because the ice model of IFS uses higher values for . From this map, biases in SICE2D-NS-CLIM experiment are much closer to that in REF. Large biases in the SICE2D-NS experiment, when compared to REF, occur because in this experiment SICE uses low value of the prescribed ice thickness, namely 1.50.75 m (ECMWF, 2017a) against 0.75against 1.5 min SICE (ECMWF, 2017a), which is used in the ice model of IFS-HRES. Note that IFS uses the parameter values the ice thickness in IFS-HRES is set to reproduce the large scale processes rather than local ones. Also, since the MODIS and VIIRS ice surface temperature derived from MODIS data corresponds to products provide information only for the clear-sky conditions, it tends they tend to have cold bias in average scores comparing to in-situ measurements (see, e.g., Hall et al., 2004). For the

Standard deviations of forecast errors as a function of forecast lead time on Fig. 9c shows that errors in SICE2D-NS and SICE2D-NS-CLIM are less random than in REF and SICE2D-S. This indicates that the ice surface temperature in SICE2D-NS and SICE2D-NS-CLIM follows the observed evolution patterns found in the MODIS and VIIRS products better than REF and SICE2D-Sexperiment, the results were closer to MODIS observations and to REF , which is in accordance while in SICE2D-NS is surface temperature is generally higher. Spatial distribution of the standard deviation of forecast errors is represented on Fig. 11. From this figure, ESTD in the experiments SICE2D-NS and SICE2D-NS-CLIM is in general smaller than in SICE2D-S and REF experiments by approximately 2 °C. In terms of ESTD SICE2D-NS and SICE2D-NS-CLIM show the best scores.

The experiment SICE2D-S shows the smallest forecast bias almost without diurnal variation (Fig. 9a and Fig. 10) for MODIS and VIIRS, but the high values of the forecast ESTD (Fig. 9c and Fig. 11). This is in agreement with the point comparisons of Sec. 3.3. However, this may be the result of compensating errors

All performed HARMONIE-AROME experiments show smaller forecast bias for the inner part of the ice field and large errors over the ice edge in Barents sea (Fig. 10). Such pattern could indicate inconsistency between the ice concentration field in the model and real structure of the sea ice field. Another possible source of these errors is inability of the model to represent characteristics of the sea ice in those areas. This situation can be illustrated by Fig. 10d where the high forecast bias values in the South-Western part of the domain are caused by underestimated ice thickness according to reanalysis climatology in SICE2D-NS-CLIM. Spatial distributions of the forecast bias and standard deviation of errors are in correspondence with area aggregated statistics.

Then, To check the performance of the SICE scheme throughout the year we compared MODIS observations and VIIRS ice surface temperature products with results of operational runs. We considered the time period from 1 December 2015 to 1 December 2016.

SICE has been running operationally within the ALADIN-HIRLAM NWP system version 38h1.2 by Norwegian Meteorological Institute for the AROME-Arctic domain (Müller et al., 2017a) since the end of October 2015 (in June 2017 the operational system was updated to the ALADIN-HIRLAM NWP system version 40h1.1). In the operational runs, the snow block in SICE is not active, and the prescribed value of the ice thickness is equal to 0.75 m, with 4 layers in the ice slab. Examples of the maps for 1 March 2017 and 9 March 2017 are given in Fig. ??. Here for the comparison we used Data from the operational ALADIN-HIRLAM NWP system archive of Norwegian Meteorological Institute for the aggregated product of the sea AROME-Arctic domain are referenced as AA-OPER throughout the text. The default operational configuration of the ALADIN-HIRLAM NWP system uses the ice surface temperature data from IFS-HRES keeping it constant throughout the forecast. To imitate the “reference” experiment we used these data from the IFS-HRES operational archive. We refer to this dataset as AA-PRESCRIBED throughout the text. AA-OPER and AA-PRESCRIBED datasets have been validated against MODIS and VIIRS ice surface temperature products.

Figure 12 shows monthly series of RMSE of the ice surface temperature ~~observations from all MODIS swathes per day (?) and daily average of the atmospheric model underlying surface temperature . From this figure, the simulated~~ as a function of the forecast lead time. These series are calculated over the AROME-Arctic domain for the forecasts initialized at 0000 UTC for AA-PRESCRIBED and AA-OPER using the MODIS product. The same plots for biases are provided by the Fig. S6, and the statistics using the VIIRS product by Figs. S7 and S8. Figures 13 and 14 show the monthly spatial distribution of RMSE of the ice surface temperature after 66 hours of forecast for AA-PRESCRIBED and AA-OPER forecasts initialized at 0000 UTC, calculated using MODIS product.

It can be seen from Fig. 12 and Figs. S6–S8 that the quality of representation of the sea ice surface temperature is ~~generally overestimated in comparison with the MODIS data for approximately 5~~varying considerably throughout the year for both AA-PRESCRIBED and AA-OPER. Averaged over the whole territory monthly biases of the ice surface temperature forecasts are positive, small during end-spring and summer time, and large during autumn and winter time in both datasets. The corresponding RMSE are also small during end-spring and summer and large during autumn and winter. Variations in RMSE between a two different months can reach approximately 15 °C~~. This happens because of the too low value of the prescribed icethickness in SICE. However, the surface temperature field patterns are reproduced well, see for example the narrow stripe of colder ice surface temperatures to~~ , with 1 °C in July and 15 °C in November (see Fig. 12). Small forecast errors during the summer time occur due to the warm (in general) state of the sea ice during this season and are constrained by the melting temperature of the ice. Large errors in the ice surface temperature forecasts during the autumn and winter time are caused by deficiencies in the parameterization of the sea ice cover.

AA-PRESCRIBED tends to show smaller RMSE than AA-OPER for the short lead times, because IFS-HRES parametrizes the sea ice as a layer with 1.5 m thickness, which is better approximation in the inner Arctic than 0.75 m used in the SICE operational configuration, especially for winter months. But for forecasts longer than 12 hours AA-PRESCRIBED show same or even worse results than AA-OPER. Using the SICE scheme in AA-OPER constrains the growth of RMSE, while in AA-PRESCRIBED, where the ice surface temperature remains constant, RMSE grows with a forecast lead time. These results indicate that for the ~~North-West of Svalbard on 01 March 2017. Overestimation~~ short forecasts (shorter than 6 hours) the ALADIN-HIRLAM NWP system better represents the sea ice cover when using the initial state provided by IFS-HRES rather than the SICE scheme. But for the forecasts longer than 12 hours, using the ice surface temperature provided by SICE leads to considerably better results.

Spatial distribution of RMSE of the ice surface temperature forecasts from AA-PRESCRIBED and AA-OPER, which is shown on Fig. 13 and 14 supports the conclusions made from the analysis of the area averaged statistics. In case of AA-PRESCRIBED, the root mean square error has larger values and is less uniform than for AA-OPER during winter, early spring and autumn months. For the summer months errors are similar.

The reason for choosing of the prescribed sea ice thickness value of 0.75 m in the operational runs of SICE was to provide the best forecast scores in the coastal areas. From the verification results, we see that with this uniform value of the ice thickness the SICE scheme overestimates the sea ice surface temperature over large territories the large ice covered areas in Arctic. This may deteriorate the large scale dynamic simulations in the operational forecast. However, in regional modelling the large scale dynamics are mainly governed by the boundary conditions from a global model thus less influenced by inaccuracies in the representation of the ice surface temperature. For this reason, we keep the prescribed sea ice thickness value in SICE optimized from comparisons with SYNOP observations (0.75) and do not increase it.

4 Conclusions

A simple thermodynamic sea ice scheme, SICE, to represent sea ice processes in NWP was developed. In this scheme, the temperature profile in the ice is predicted by solving the heat diffusion equation in the slab of ice with a prescribed thickness. The scheme design allows explicit coupling with a snow schemeexplicitly, via the fluxes and temperature at the snow-ice interface. Also, the scheme includes the form drag for the momentum flux in the surface layer due to ice obstacles in the case of the fractional ice cover.

The scheme was preliminarily tested by comparing it with the sea ice model HIGHTSI (Cheng and Launiainen, 1998) in off-line mode. In the off-line experiments, when the snow block was switched off in both schemes, the difference in the simulated ice surface temperature between SICE and HIGHTSI was small (the difference standard deviation is equal to 1.04 °C), when the ice thickness modelled by HIGHTSI was approximately equal to that prescribed in SICE. With the snow block included, due Due to high variability of the snow surface temperature, the difference between the two schemes was larger, with the difference standard deviation of 1.99 °C, when the snow block is included. From this comparison, the conclusion about overall sanity of SICE was made, but the sensitivity of the results to the prescribed value of the ice thickness was noticednoted.

General evaluation of SICE was performed in the coupled modeSICE was evaluated in a coupled framework, within the HARMONIE-AROME configuration of the ALADIN-HIRLAM NWP system , to assess the scheme performance and to study possible errors. The ice fraction field for the SICE experiments was provided by the OSTIA product (Stark et al., 2007) via the lower boundary conditions from the ECMWF model, IFSIFS-HRES. In the reference experiment, the sea ice surface temperature was taken from the IFS IFS-HRES model and remained constant during the forecasting cycle. Coupled experiments were performed for the SICE scheme with and without the snow scheme and the form drag included. A separate experiment where the ice thickness in SICE is initialized from the model climatology was also performed. For validation, data from coastal SYNOP stations in the Svalbard and Gulf of Bothnia regions were used. Also, visual comparisons with observations from MODIS To study the performance of SICE on the large scale, comparisons against ice surface temperature products from MODIS and VIIRS were performed.

In verification against measurements from coastal SYNOP stations, coupled experiments showed that the impact of SICE on the 2 metre temperature scores was positive without snow model, however with the snow model no clear positive impact was seen. For the mean sea level pressure verification scores, a minor positive impact was seen for all SICE experiments. In the SICE experiments without the form drag compared to the reference experiment (which contains no ice fraction representation), an increased positive 10 metre wind speed bias was noted. This bias was reduced due to after accounting for the form drag in

SICE. However, our conclusion about the impact of the form drag is still preliminary~~, operational implementation will need~~ . Operational implementation needs additional testing, including experiments covering different seasons and tuning. Also, the form drag term strongly depends on the ice fraction value, thus ice concentration observations of better quality than low resolution passive-microwave data used in OSTIA are desirable. For example, in Posey et al. (2015) it is shown that using high resolution passive-microwave data from Advanced Microwave Scanning Radiometer 2 (AMSR-2) leads to a substantial decrease in model errors.

Comparisons of the model experiments with ~~observations from MODIS~~ the satellite ice surface temperature products over the Arctic domain showed that in the reference experiment ~~simulated~~ the ice surface temperature ~~field with smaller bias than the SICE experiment~~ forecast has smaller errors than in the experiments with SICE (with no snow scheme included) for the first 12 hours of forecast. This happens because the prescribed ice thickness in the ~~simple ice model of IFS~~ sea ice parameterization used by IFS-HRES is tuned to reproduce the large scale fields rather than local effects, and it is larger than in SICE. However general patterns of the ice surface temperature field are well captured by SICE and after 24 hours of forecast, the predicted ice surface temperature in the experiments with SICE shows smaller root mean square error than in the reference experiment. The best results considering both the forecast bias and standard deviation of errors were obtained from the SICE scheme experiments with the ice thickness prescribed from the model climatology provided by TOPAZ4 reanalysis (Xie et al., 2017).

Our general conclusion from the numerical experiments is that SICE can improve forecasts of the HARMONIE-AROME configuration of the ALADIN-HIRLAM NWP system in ice-surrounded areas, especially for forecasts longer than 24 hours. At the moment we recommend to use it without snow parameterization for a trouble-proof result. The prescribed ice thickness is an important parameter, and since no estimates of the ice thickness from observations or other sources are used in the current version of ALADIN-HIRLAM NWP system, it should be tuned.

~~The simplest way to go forward is to replace the prescribed constant value of the ice thickness by the climatology, to reproduce its seasonal and horizontal large scale variations.~~ Of course, in the future the scheme itself should reproduce the spatial and temporal inhomogeneity of the ice thickness. Further development will be focused on the physical processes that control the evolution of sea ice, such as ice freezing and melting. Additionally, possibilities to improve the parameterization of snow on sea ice will be studied. More tests on the parameterization of the form drag are planned. The performance of the scheme will be more carefully evaluated for more regions and more seasons, for example for a summer period in the Arctic region when ice melting processes occur, which did not get much attention during this study. Melt ponds affect the atmosphere mainly through changing radiation fluxes, but they may also influence the modelling results of a whole NWP system, since they lead to higher uncertainty in the ice concentration observations coming from passive microwave remote sensing. The initialization of the ice parameterization scheme and model error corrections (especially for the snow module) using observations are also of high importance. The possibilities to use more observations and to develop methods to assimilate them, as well as to improve the methods of using existing observations, should be carefully studied.

Code availability. SICE is a part of the ALADIN-HIRLAM NWP system, which is not available to the general public. A copy of the ALADIN-HIRLAM NWP system source code can be obtained, for non-commercial research purposes only, from a member institution of ALADIN or HIRLAM consortium in applicant's country after signing a standardized License Agreement. An extract from the source code of the ALADIN-HIRLAM NWP system version 38h1 that contains only the source code of the SICE scheme version 1.0-38h1 is available in the Supplement.

Appendix A: Numerical solution

To solve equations Eq. (1) and Eq. (2) numerically, the ice slab of thickness H is divided into K layers of equal thickness, except for the topmost layer. For the thickness of the topmost layer, the following formulation is used:

$$z_1 = min\left|z^*, \frac{H - z^*}{K-1}\right| \tag{A1}$$

$$z^* = \begin{cases} 0.05 & \text{if } H \geq 0.2 \\ 0.25 \cdot H & \text{otherwise} \end{cases} \tag{A2}$$

The ice temperature and thermal properties are assumed to be constant within the current layer. Then, according to the implicit Euler numerical scheme, the first row of Eq. (1) may be rewritten for the layer number $j = 1 \dots K$ as follows (subscripts denote the layer number j, superscripts $^-$ and $^+$ denote the variables at the beginning and at the end of the time step Δt, Δz_j is the thickness of layer j)

$$\frac{C_j \Delta z_j}{\Delta t}\left(T_j^+ - T_j^-\right) = \frac{\bar{\lambda}_{j-1}}{\Delta \bar{z}_{j-1}}\left(T_{j-1}^+ - T_j^+\right) - \frac{\bar{\lambda}_j}{\Delta \bar{z}_j}\left(T_j^+ - T_{j+1}^+\right) - Q\big|_{z=z_j}^- + Q\big|_{z=z_j-\Delta z_j}^- \tag{A3}$$

where

$$\Delta \bar{z}_j = \frac{\Delta z_j + \Delta z_{j+1}}{2} \quad \text{and} \quad \bar{\lambda}_j = \frac{\Delta z_j \lambda_j + \Delta z_{j+1} \lambda_{j+1}}{\Delta z_j + \Delta z_{j+1}} \tag{A4}$$

This defines a tridiagonal matrix (see Boone (2000) for a detailed description). The skin temperature of ice could be obtained by integrating the first row of Eq. (1) over the topmost layer assuming that the properties of ice are constant within the selected layer. Thus, combined with the second equation from the system Eq. (1), Eq. (2) and Eq. (4), the equation for the ice temperature within the skin layer can be written as:

$$C_t \frac{\partial T_s}{\partial t} = \delta_{H_{snow}}(R_n - H - LE) + (1 - \delta_{H_{snow}})G_{snow} + \lambda \left.\frac{\partial T}{\partial z}\right|_{z=z_1} \tag{A5}$$

where $C_t \equiv C|_{z=z_1} \cdot \Delta z_1$ is the surface thermal resistance ($\mathrm{W \cdot s\ m^{-2} K^{-1}}$); Δz_1 is the thickness of the upper layer of ice (m); $R_n = (1 - i_0 \cdot e^{-k \cdot z_1})(1-\alpha)SW\!\downarrow + LW\!\downarrow - \varepsilon\sigma T_s^4$ is the radiative balance. The finite differential representation of Eq. (A5) with the implicit Euler scheme gives the upper row of the matrix Eq. (A3). In the case of no snow it reads:

$$\frac{C_t}{\Delta t}\left(T_s^+ - T_s^-\right) = R_n^{\pm} - H^{\pm} - LE^{\pm} - \frac{\lambda}{\Delta \bar{z}_1}\left(T_s^+ - T_1^+\right), \tag{A6}$$

Note that all the fluxes $R_n^{\pm}$, $H^{\pm}$, $LE^{\pm}$ are calculated using the prognostic variables at the end of the time step. For example, in the case of coupling with an atmospheric model, $H^{\pm}$ can be written as

$$H^{\pm} = \rho_a^- c_p^- c_{\mathrm{H}}^- V_{\mathrm{N}}^- \left(T_s^+ - T_{\mathrm{N}}^+\right) \tag{A7}$$

For obtaining the future value of T_{N}^+, a procedure known as "implicit coupling" (Best et al., 2004) is used. According to this procedure, the atmospheric variable X_N^+ from the lowest model level at the end of the time step can be found from

$$X_N^+ = A_{X,N}^- \cdot F_{X,S}^{\pm} + B_{X,N}^- \tag{A8}$$

This procedure uses the coefficients $A_{X,N}^-$ and $B_{X,N}^-$ from the implicit numerical solution of the vertical diffusion scheme from the atmospheric model, and the surface flux $F_{X,S}^{\pm}$ of the variable X. The coupling coefficients in Eq. (A8) are provided by the host model. Term R_n in Eq. (A6) represents the radiative balance and contains the nonlinear term $\varepsilon\sigma {T_s^+}^4$, which defines the thermal radiation flux from the ice surface to the atmosphere at time step $t+\Delta t$. Linearization of this term can be done by use of the Taylor series which results in:

$$\varepsilon\sigma {T_s^+}^4 \approx 4\varepsilon\sigma {T_s^-}^3 T_s^+ - 3\varepsilon\sigma {T_s^-}^4 \tag{A9}$$

Then, Eq. (A8) and Eq. (A9) may be applied to transform Eq. (A6) to the form: $T_s^+ - \mathcal{A}_2 T_1^+ = \mathcal{A}_1$. This form is suitable to be the upper row in the tridiagonal matrix represented by Eq. (A3). For the lower boundary condition, the temperature at the bottom of the ice slab (at the bottom of the layer K) is equal to the freezing point of the sea water, according to the last equation of system Eq. (1). In this case, the lower row of the matrix represented by Eq. (A3) can be written as:

$$-\frac{\bar{\lambda}_{\mathrm{K}-1}}{\Delta\bar{z}_{\mathrm{K}-1}} T_{\mathrm{K}-1}^+ + \left[\frac{C_{\mathrm{K}}\Delta z_{\mathrm{K}}}{\Delta t} + \frac{\bar{\lambda}_{\mathrm{K}-1}}{\Delta\bar{z}_{\mathrm{K}-1}} + \frac{2\bar{\lambda}_{\mathrm{K}}}{\Delta z_{\mathrm{K}}}\right] T_{\mathrm{K}}^+ = \frac{C_{\mathrm{K}}\Delta z_{\mathrm{K}}}{\Delta t} T_{\mathrm{K}}^- + \frac{2\bar{\lambda}_{\mathrm{K}}}{\Delta z_{\mathrm{K}}} T_{frz} - Q\big|_{z=H}^- + Q\big|_{z=H-\Delta z_{\mathrm{K}}}^- \tag{A10}$$

The resulting system of linear equations may be solved through the Thomas algorithm (Thomas, 1949).

Actual implementation can be found in the source file `src/surfex/SURFEX/simple_ice.F90` available in the Supplement.

Competing interests. The authors declare that they have no conflict of interests.

Acknowledgements. The authors would like to thank Bin Cheng, Andrew Singleton, Erin Thomas and Laura Rontu for their constructive comments. We are grateful to Malte Müller for preparing the climatological ice thickness data. Comments and suggestions from the two anonymous referees helped to considerably improve our manuscript.

References

Aaboe, S., Breivik, L.-A., Sørensen, A., Eastwood, S., and Lavergne, T.: Product User Manual for Global Sea Ice Edge and Type, Tech. rep., Norwegian Meteorological Institute, Oslo, Norway, 2017.

Andersen, S., Breivik, L.-A., Eastwood, S., Godøy, Ø., Lavergne, T., Lind, M., Porcires, M., Schyberg, H., and Tonboe, R.: Sea Ice Product Manual, Tech. rep., http://osisaf.met.no/docs/osisaf_ss2_pum_ice-conc-edge-type_v3p8.pdf, 2012.

Atlaskin, E. and Vihma, T.: Evaluation of NWP results for wintertime nocturnal boundary-layer temperatures over Europe and Finland, Quarterly Journal of the Royal Meteorological Society, 138, 1440–1451, https://doi.org/10.1002/qj.1885, http://dx.doi.org/10.1002/qj.1885, 2012.

Bailey, E., Feltham, D. L., and Sammonds, P. R.: A model for the consolidation of rafted sea ice, Journal of Geophysical Research: Oceans, 115, https://doi.org/10.1029/2008JC005103, http://dx.doi.org/10.1029/2008JC005103, c04015, 2010.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V., Yang, X., and Køltzow, M. Ø.: The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System, Monthly Weather Review, 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, http://dx.doi.org/10.1175/MWR-D-16-0417.1, 2017.

Best, M. J., Beljaars, A., Polcher, J., and Viterbo, P.: A Proposed Structure for Coupling Tiled Surfaces with the Planetary Boundary Layer, Journal of Hydrometeorology, 5, 1271–1278, https://doi.org/10.1175/JHM-382.1, http://dx.doi.org/10.1175/JHM-382.1, 2004.

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 developments of the Met Office operational forecasting system: an overview and assessment of the new Global FOAM forecasts, Geoscientific Model Development, 7, 2613–2638, https://doi.org/10.5194/gmd-7-2613-2014, 2014.

Boone, A.: Modeling hydrological processes in the land surface scheme ISBA: inclusion of a hydrological reservoir, ice and a snow model, Ph.D. thesis, Paul Sabatier University, 2000.

Boone, A. and Etchevers, P.: An Intercomparison of Three Snow Schemes of Varying Complexity Coupled to the Same Land Surface Model: Local-Scale Evaluation at an Alpine Site, Journal of Hydrometeorology, 2, 374–394, https://doi.org/10.1175/1525-7541(2001)002⟨0374:AIOTSS⟩2.0.CO;2, http://dx.doi.org/10.1175/1525-7541(2001)002<0374:AIOTSS>2.0.CO;2, 2001.

Brassington, G. B., Martin, M. G., Tolman, H. L., Akella, S., Balmeseda, M., Chambers, C. R. S., Chassignet, E., Cummings, J. A., Drillet, Y., Jansen, P. A. E. M., Laloyaux, P., Lea, D., Mehra, A., Mirouze, I., Ritchie, H., Samson, G., Sandery, P. A., Smith, G. C., Suarez, M., and Todling, R.: Progress and challenges in short- to medium-range coupled prediction, Journal of Operational Oceanography, 8, 239–258, https://doi.org/10.1080/1755876X.2015.1049875, 2015.

Breivik, L.-A., Eastwood, S., Godøy, Ø., Schyberg, H., Andersen, S., and Tonboe, R.: Sea ice products for EUMETSAT satellite application facility, Canadian Journal of Remote Sensing, 27(5), 403–410, https://doi.org/10.1080/07038992.2001.10854883, 2001.

Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F.: Flux-Profile Relationships in the Atmospheric Surface Layer, Journal of the Atmospheric Sciences, 28, 181–189, https://doi.org/10.1175/1520-0469(1971)028⟨0181:FPRITA⟩2.0.CO;2, http://dx.doi.org/10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2, 1971.

Cheng, B. and Launiainen, J.: A one-dimensional thermodynamic air-ice-water model: technical and algorithm description report, MERI – Report Series of the Finnish Institute of Marine Research, 37, 15–36, 1998.

Cheng, B., Zhang, Z., Vihma, T., Johansson, M., Bian, L., Li, Z., and Wu, H.: Model experiments on snow and ice thermodynamics in the Arctic Ocean with CHINARE 2003 data, Journal of Geophysical Research: Oceans, 113, https://doi.org/10.1029/2007JC004654, http://dx.doi.org/10.1029/2007JC004654, c09020, 2008.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sensing of Environment, 116, 140–158, https://doi.org/10.1016/j.rse.2010.10.017, 2012.

Dupont, F., Higginson, S., Bourdallé-Badie, R., Lu, Y., Roy, F., Smith, G. C., Lemieux, J.-F., Garric, G., and Davidson, F.: A high-resolution ocean and sea-ice modelling system for the Arctic and North Atlantic oceans, Geoscientific Model Development, 8, 1577–1594, https://doi.org/10.5194/gmd-8-1577-2015, 2015.

ECMWF: PART IV: PHYSICAL PROCESSES, IFS Documentation, ECMWF, 2017. 2017a.

ECMWF: PART V: ENSEMBLE PREDICTION SYSTEM, IFS Documentation, ECMWF, 2017b.

Elvidge, A. D., Renfrew, I. A., Weiss, A. I., Brooks, I. M., Lachlan-Cope, T. A., and King, J. C.: Observations of surface momentum exchange over the marginal ice zone and recommendations for its parametrisation, Atmospheric Chemistry and Physics, 16, 1545–1563, https://doi.org/10.5194/acp-16-1545-2016, 2016.

Faroux, S., Kaptué Tchuenté, A. T., Roujean, J.-L., Masson, V., Martin, E., and P., L. M.: ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1km resolution based on satellite information for use in land surface, meteorological and climate models, Geoscientific Model Development, 6, 563–582, https://doi.org/10.5194/gmd-6-563-2013, 2013.

Feltham, D. L., Untersteiner, N., Wettlaufer, J. S., and Worster, M. G.: Sea ice is a mushy layer, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026290, http://dx.doi.org/10.1029/2006GL026290, 2006.

Grenfell, T. G. and Maykut, G. A.: The optical properties of ice and snow in the arctic basin, Journal of Glaciology, 18, 445–463, 1977.

Hall, D. K. and Riggs, G.: MODIS/Terra Sea Ice Extent 5-Min L2 Swath 1km, Version 6. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, http://dx.doi.org/10.5067/MODIS/MOD29.006, 2015. Hall, D. K. and Riggs, G.: MODIS/TerraSea Ice Extent and ISTDaily L3Global 4km EASE-Grid Day, Version 6. NASANationalSnowand IceDataCenterDistributedActiveArchiveCenter, http://dx.doi.org/10.5067/MODIS/MOD29.006, 2015. 2015.

Hall, D. K., Key, J. R., Casey, K. A., Riggs, G. A., and Cavalieri, D. J.: Sea ice surface temperature product from MODIS, IEEE Transactions on Geoscience and Remote Sensing, 42, 1076–1087, https://doi.org/10.1109/TGRS.2004.825587, 2004.

Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geoscientific Model Development, 4, 223–253, https://doi.org/10.5194/gmd-4-223-2011, https://www.geosci-model-dev.net/4/223/2011/, 2011.

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., and Elliott, S.: CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual, Tech. rep., Los Alamos National Laboratory, Los Alamos NM 87545, 2015.

Launiainen, J. and Cheng, B.: Modelling of ice thermodynamics in natural water bodies, Cold Regions Science and Technology, 27, 153 – 178, https://doi.org/http://dx.doi.org/10.1016/S0165-232X(98)00009-3, http://www.sciencedirect.com/science/article/pii/S0165232X98000093, 1998.

Lea, D. J., Mirouze, I., Martin, M. G., King, R. R., Hines, A., Walters, D., and Thurlow, M.: Assessing a New Coupled Data Assimilation System Based on the Met Office Coupled Atmosphere-Land-Ocean-Sea Ice Model, Monthly Weather Review, 143, 4678–4694, https://doi.org/10.1175/MWR-D-15-0174.1, 2015.

Lüpkes, C., Gryanik, V. M., Hartmann, J., and Andreas, E. L.: A parametrization, based on sea ice morphology, of the neutral atmospheric drag coefficients for weather prediction and climate models, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2012JD017630, http://dx.doi.org/10.1029/2012JD017630, d13112, 2012.

MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A., Comer, R. E., Camp, J., Xavier, P., and Madec, G.: Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system, Quarterly Journal of the Royal Meteorological Society, 141, 1072–1084, https://doi.org/10.1002/qj.2396, 2015.

Mahfouf, J.-F., Bergaoui, K., Draper, C., Bouyssel, F., Taillefer, F., and Taseva, L.: A comparison of two off-line soil analysis schemes for assimilation of screen level observations, Journal of Geophysic Research, 114, https://doi.org/10.1029/2008JD011077, 2009.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geoscientific Model Development, 6, 929–960, https://doi.org/10.5194/gmd-6-929-2013, http://www.geosci-model-dev.net/6/929/2013/, 2013.

Mélia, D. S.: A global coupled sea ice-ocean model, Ocean Modelling, 4, 137 – 172, https://doi.org/http://dx.doi.org/10.1016/S1463-5003(01)00015-4, http://www.sciencedirect.com/science/article/pii/S1463500301000154, 2002.

Mironov, D. and Ritter, B.: Testing the new ice model for the global NWP system GME of the German Weather Service, Research Activities in Atmospheric and Oceanic Modelling, Rep.34, WMO/TD 1220, 4.21–4.22, 2004.

Mironov, D., Ritter, B., Schulz, J.-P., Buchhold, M., Lange, M., and Machulskaya, E.: Parameterisation of sea and lake ice in numerical weather prediction models of the German Weather Service, Tellus A, 64, 17 330, https://doi.org/10.3402/tellusa.v64i0.17330, 2012.

Mironow, D. and Ritter, B.: First version of the ice model for the global NWP system GME of the German weather service, Research activities in atmospheric and oceanic modelling, Report No. 33, 4.13–4.14, 2003.

Müller, M., Batrak, Y., Kristiansen, J., Køltzow, M. A. O., Noer, G., and Korosov, A.: Characteristics of a Convective-Scale Weather Forecasting System for the European Arctic, Mon. Wea. Rev., 145, 4771–4787, https://doi.org/10.1175/MWR-D-17-0194.1, https://doi.org/10.1175/MWR-D-17-0194.1, 2017.
M2017üaller.

Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M. A. Ø., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M., and Vignes, O.: AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model, Weather and Forecasting, 32, 609–627, https://doi.org/10.1175/WAF-D-16-0099.1, http://dx.doi.org/10.1175/WAF-D-16-0099.1, 2017. 2017b.

Parkinson, C. L. and Washington, W. M.: A large-scale numerical model of sea ice, Journal of Geophysical Research: Oceans, 84, 311–337, https://doi.org/10.1029/JC084iC01p00311, http://dx.doi.org/10.1029/JC084iC01p00311, 1979.

Pellerin, P., Ritchie, H., Saucier, F. J., Roy, F., Desjardins, S., Valin, M., and Lee, V.: Impact of a Two-Way Coupling between an Atmospheric and an Ocean-Ice Model over the Gulf of St.Lawrence, Monthly Weather Review, 132, 1379–1398, 2004.

Perovich, D. K.: The Optical Properties of Sea Ice, Tech. Rep. Monograph 96-1, Cold Regions Research and Engineering Lab (CRREL) Hanover NH, U.S. Army Cold Regions Research and Engineering Laboratory 72 Lyme Road Hanover, New Hampshire 03755-1290, http://www.dtic.mil/docs/citations/ADA310586, 1996.

Posey, P. G., ~~Metzeger~~Metzger, E. J., Wallcraft, A. J., Hebert, D. A., Allard, R. A., ~~Smetstad~~Smedstad, O. M., Phelps, M. W., ~~Fettere~~Fetterer, F., ~~Steward~~Stewart, J. S., Meier, W. N., and Helfrich, S. R.: Improving Arctic sea ice edge forecasts by assimilating high ~~horizonlat~~ horizontal resolution sea ice concentration data into the US Navy's ice forecast systems, The Cryosphere, 9, 1735–1745, https://doi.org/10.5194/tc-9-1735-2015, 2015.

Rae, J. G. L., Hewitt, H. T., Keen, A. B., Ridley, J. K., West, A. E., Harris, C. M., Hunke, E. C., and Walters, D. N.: Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model, Geoscientific Model Development, 8, 2221–2230, https://doi.org/10.5194/gmd-8-2221-2015, https://www.geosci-model-dev.net/8/2221/2015/, 2015.

Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U., and Schulzweida, U.: Simulation of the present-day climate with the ECHAM model: Impact of model physics and resolution, Tech. Rep. MPI-Report No. 93, Max Planck Institute for Meteorology, Bundesstrasse 55 D-2000 Hamburg 13 F.R. Germany, http://www.dtic.mil/docs/citations/ADA310586, 1992.

Sakatume, S. and Seki, N.: On the Thermal Properties of Ice and Snow in a Low Temperature Region, Transactions of the Japan Society of Mechanical Engineers, 44, 2059–2069, https://doi.org/10.1299/kikai1938.44.2059, 1978.

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 Science, 8, 633–656, https://doi.org/10.5194/os-8-633-2012, http://www.ocean-sci.net/8/633/2012/, 2012.

Saloranta, T. M.: Modeling the evolution of snow, snow ice and ice in the Baltic Sea, Tellus A, 52, 93–108, https://doi.org/10.1034/j.1600-0870.2000.520107.x, http://dx.doi.org/10.1034/j.1600-0870.2000.520107.x, 2000.

Schwerdtfecer, P.: The Thermal Properties of Sea Ice, Journal of Glaciology, 4, 789–807, https://doi.org/10.1017/S0022143000028379, 1963.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., and Masson, V.: The AROME-France convective-scale operational model, Monthly Weather Review, 139, 976–991, https://doi.org/10.1175/2010MWR3425.1, 2011.

Semmler, T., Cheng, B., Yang, Y., and Rontu, L.: Snow and ice on Bear Lake (Alaska) – sensitivity experiments with two lake ice models, Tellus A: Dynamic Meteorology and Oceanography, 64, 17 339, https://doi.org/10.3402/tellusa.v64i0.17339, http://dx.doi.org/10.3402/tellusa.v64i0.17339, 2012.

Slater, A. G., Schlosser, C. A., Desborough, C. E., Pitman, A. J., Hendersson-Sellers, A., Robock, A., Vinnikov, K. Y., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P. M., de Rosnay, P., Dickinson, R. E., Dai, Y.-J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y. M., Habets, F., Kim, J., Koren, V., Kowalchyk, E. A., Nasonova, O. N., Noilhan, J., Schaake, S., Shmakin, A. B., Smirnova, T. G., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z.-L., and Zeng, Q.: The Representation of Snow in Land Surface Schemes: Results from PILPS 2(d), Journal of Hydrometeorology, 2, 7–25, 2001.

Stark, J., Donlon, C., Martin, M., and McCulloch, M.: OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system, in: OCEANS 2007 - Europe, pp. 1–4, https://doi.org/10.1109/OCEANSE.2007.4302251, 2007.

Thomas, L. H.: Elliptic problems in linear difference equations over a network, Tech. rep., Columbia University, 1949.

Tilling, R. L., Ridout, A., and Shepherd, A.: Near-real-time Arctic sea ice thickness and volume from CryoSat-2, The Cryosphere, 10, 2003–2012, https://doi.org/10.5194/tc-10-2003-2016, 2016.

Tsamados, M., Feltham, D. L., Schroeder, D., Flocco, D., Farrell, S. L., Kurtz, N., Laxon, S. W., and Backon, S.: Impact of Variable Atmospheric and Oceanic Form Drag on Simulations of Arctic Sea Ice, Journal of Physical Oceanography, 44, 1329–1353, https://doi.org/10.1175/JPO-D-13-0215.1, 2014.

Tschudi, M., Riggs, G., Hall, D. K., and Román, M. O.: VIIRS/NPP Ice Surface Temperature 6-Min L2 Swath 750m, Version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/10.5067/VIIRS/VNP30.001, 2017.

Unden, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., Jones, C., Geert, Lenderlink, G., Mcdonald, A., Mcgrath, R., Navascues, B., Nielsen, N. W., Degaard, V., Rodriguez, E., Rummukainen, M., Sattler, K., Sass, B. H., Savijarvi, H., Schreur, B. W., Sigg, R., and The, H.: HIRLAM-5 Scientific Documentation, 2002.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geoscientific Model Development, 10, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017, https://www.geosci-model-dev.net/10/1487/2017/, 2017.

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013, Ocean Science, 13, 123–144, https://doi.org/10.5194/os-13-123-2017, https://www.ocean-sci.net/13/123/2017/, 2017.

Table 1. Physical parameters of the SICE scheme and parameters for the numerical solution. All the parameters except the ice salinity may be selected by the user, the range and default values are given

Parameter	Value and/or reference
Number of layers in the ice	$[3, 99]$, the default is 4
Number of layers in the snow	~~the~~ $[3, 99]$, the hard-coded default is 3
Ice thickness	0.75 m
Ice thermal properties	after Schwerdtfecer (1963); Feltham et al. (2006) and Sakatume and Seki (1978)
Ice salinity	3 ppt
Ice albedo	after Perovich (1996) or Parkinson and Washington (1979) or Roeckner et al. (1992)
Radiative transfer within ice	Bouguer-Lambert law, coefficients after Grenfell and Maykut (1977)
Freezing point	the default is -1.8 °C

Table 2. Design of experiments: Exp. name – the experiment name, Domain – the experiment domain, Length – the length of the experiment run, Ice cover – "fractional" or "binary" for the ice fraction taken into account or not, respectively, Ice scheme – which sea ice parameterization scheme is used if any, Ice thickness – how the ice thickness is initialized in case of using SICE, Snow scheme – which snow module is used if any, Form drag – whether the parameterization of the form drag used or not.

Exp. name	Domain	Length	Ice cover	Ice scheme	Ice thickness	Snow scheme	Form drag
REF	Arctic, MetCoOp	03-04.2013 03.2013	binary	no		no	no
SICE2D-NS	Arctic, MetCoOp	03-04.2013 03.2013	fractional	SICE	uniform, 0.75 m	no	no
SICE2D-S	Arctic	03-04.2013	fractional	SICE	uniform, 0.75 m	ISBA ES	no
SICE2D-AD	Arctic	03.2013	fractional	SICE	uniform, 0.75 m	no	yes
SICE2D-NS-CLIM	Arctic	03-04.2013	fractional	SICE	climatology, 0.2–2.2 m	no	no

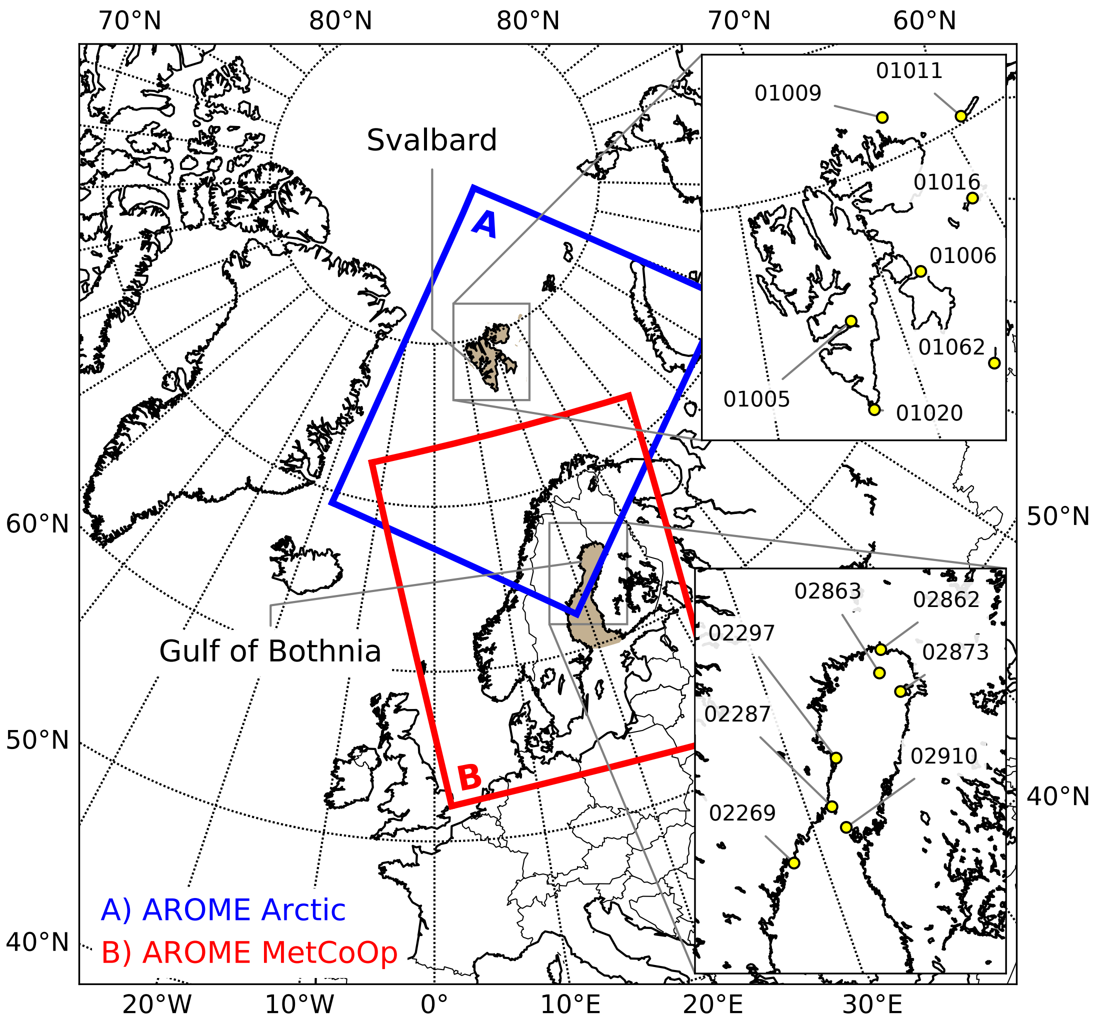

Figure 1. Experiment domains. Insets: locations and WMO numbers of the SYNOP stations at Svalbard and around the Gulf of Bothnia used in this study.

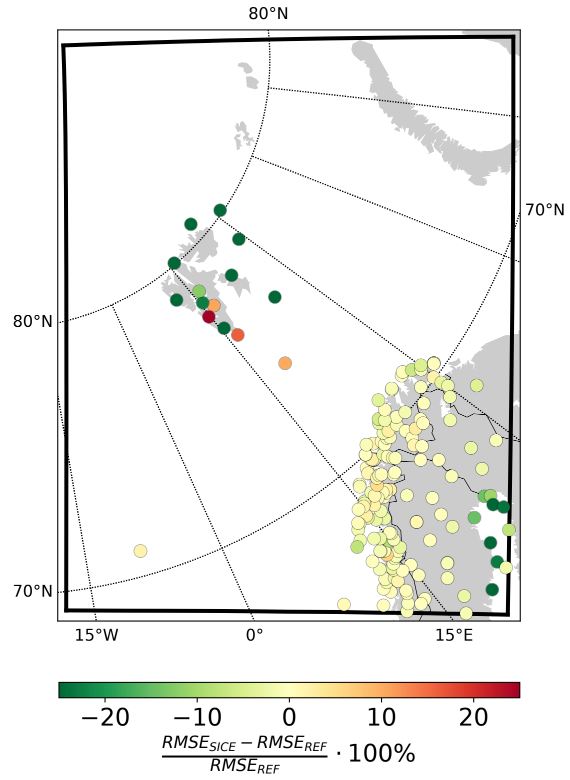

Figure 2. Relative change (in percent) of RMSE of the 2 metre temperature forecasts between SICE2D-NS and REF. Negative values mean that the RMSE is smaller in SICE2D-NS than in REF. Forecasts starting at 0000 UTC within the time period from 1 March 2013 to 30 April 2013 were used for comparison.

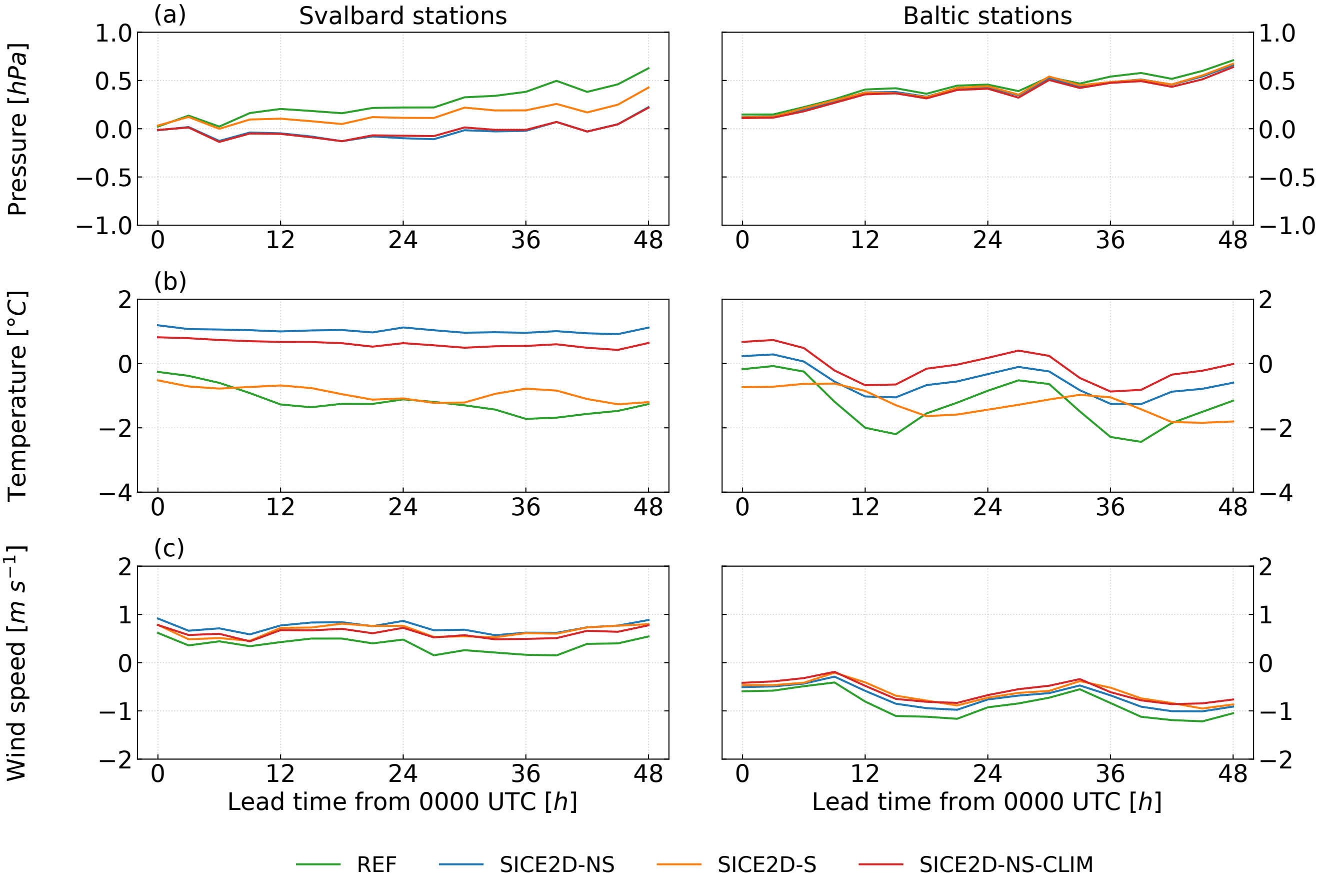

Figure 3. Mean error as a function of lead time for forecasts initialized at 0000 UTC for experiments REF, SICE2D-NS, SICE2D-S and SICE2D-NS-CLIM over the AROME Arctic domain for the period from 1 March 2013 to 30 April 2013. Left panel: for 7 Svalbard stations (WMO Nos. 01005, 01006, 01009, 01011, 01016, 01020, 01062). Right panel: for 7 stations in Gulf of Bothnia (WMO Nos. 02269, 02287, 02297, 02862, 02863, 02873, 02910). The mean error is calculated as the forecasted value minus observed value. a) mean sea level pressure; b) 2 metre temperature; c) 10 metre wind speed.

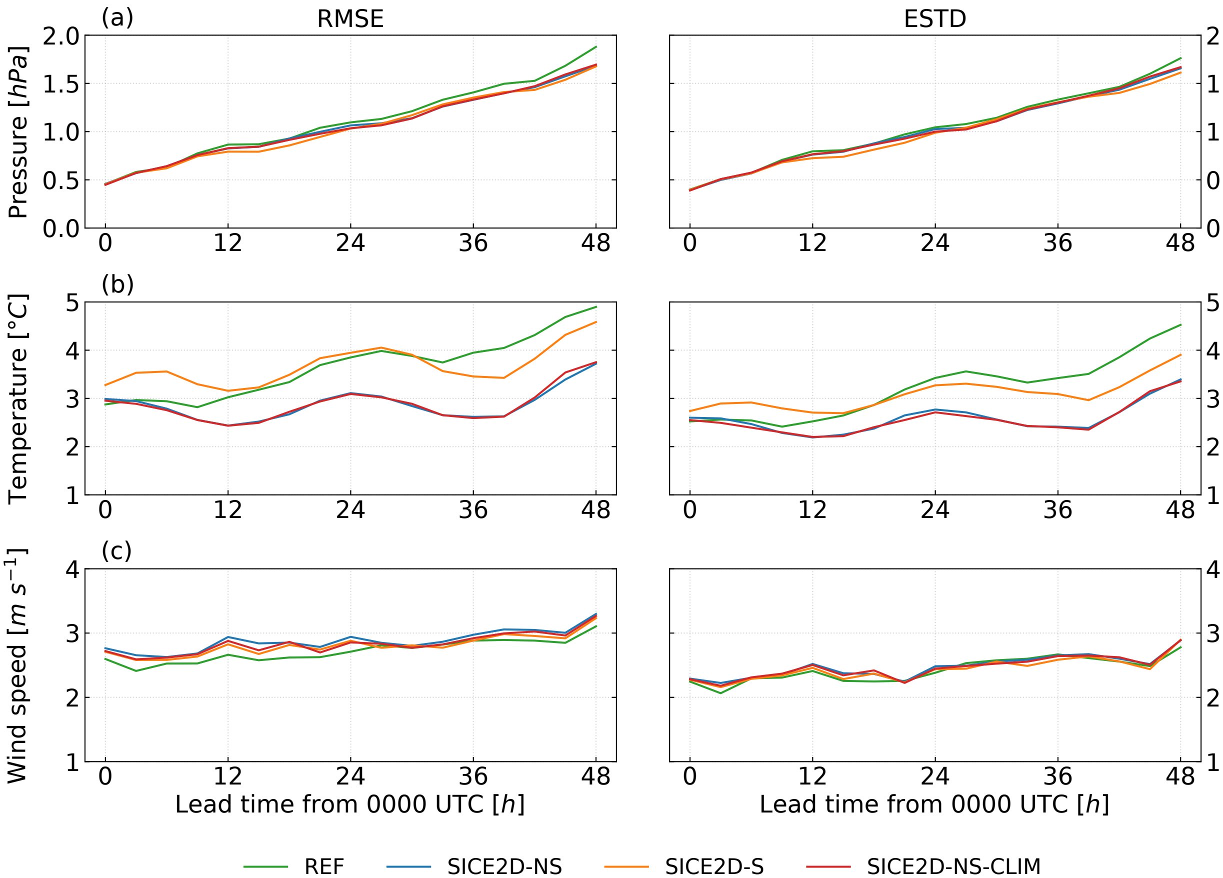

Figure 4. Root mean square error (RMSE, left panel) and standard deviation of errors (ESTD, right panel) as a function of lead time for forecasts initialized at 0000 UTC for experiments REF, SICE2D-NS, SICE2D-S and SICE2D-NS-CLIM over the AROME Arctic domain for the period from 1 March 2013 to 30 April 2013. Series are calculated for 7 Svalbard stations. a) mean sea level pressure; b) 2 metre temperature; c) 10 metre wind speed.

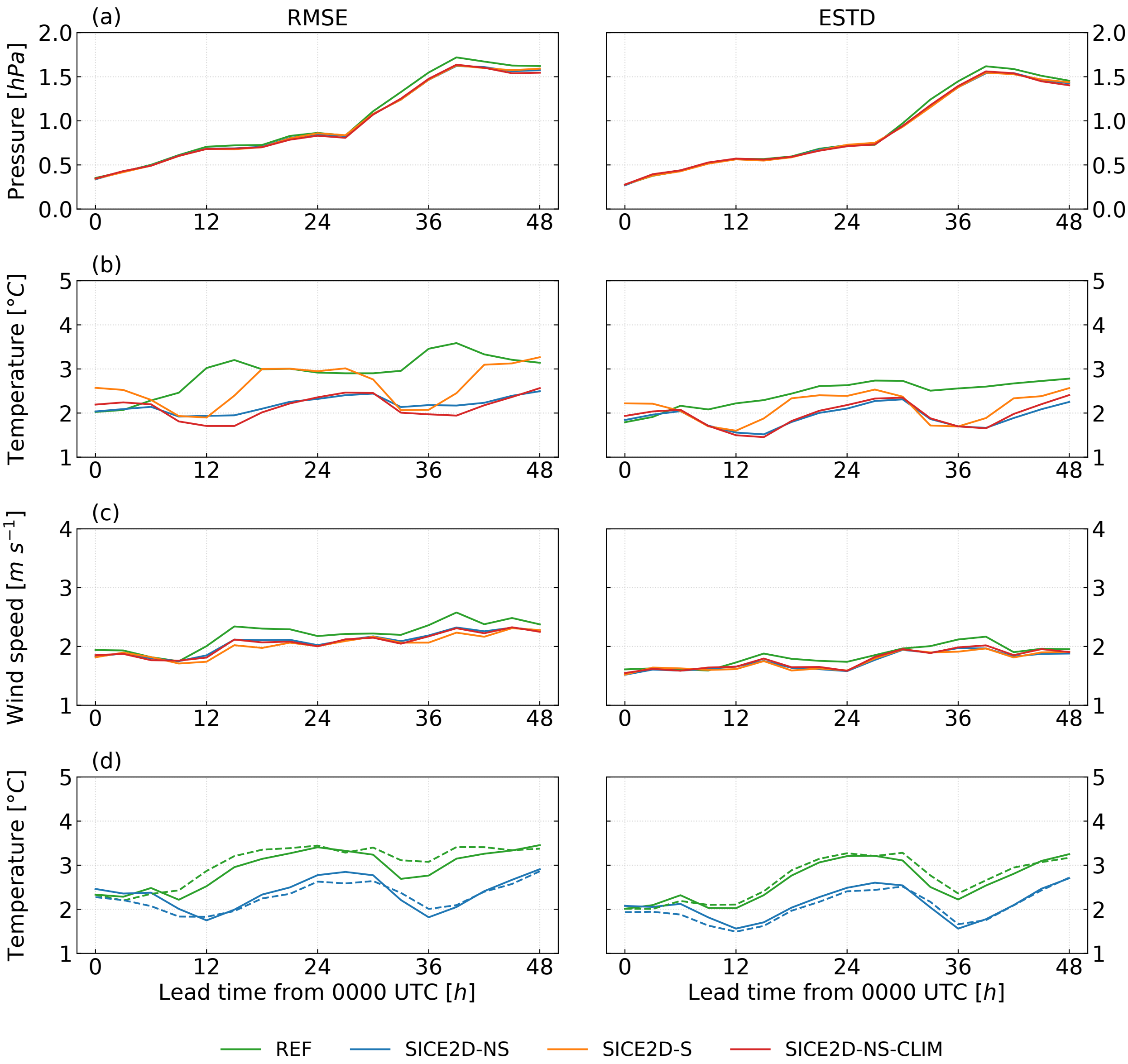

Figure 5. Same as Fig. 4, but for 7 stations located in the coastal area of the Gulf of Bothnia. Panel (d) shows the same error statistics as panel (b) but only for March 2013 for experiments REF and SICE2D-NS over AROME Arctic (solid lines) and MetCoOp (dashed lines) domains.

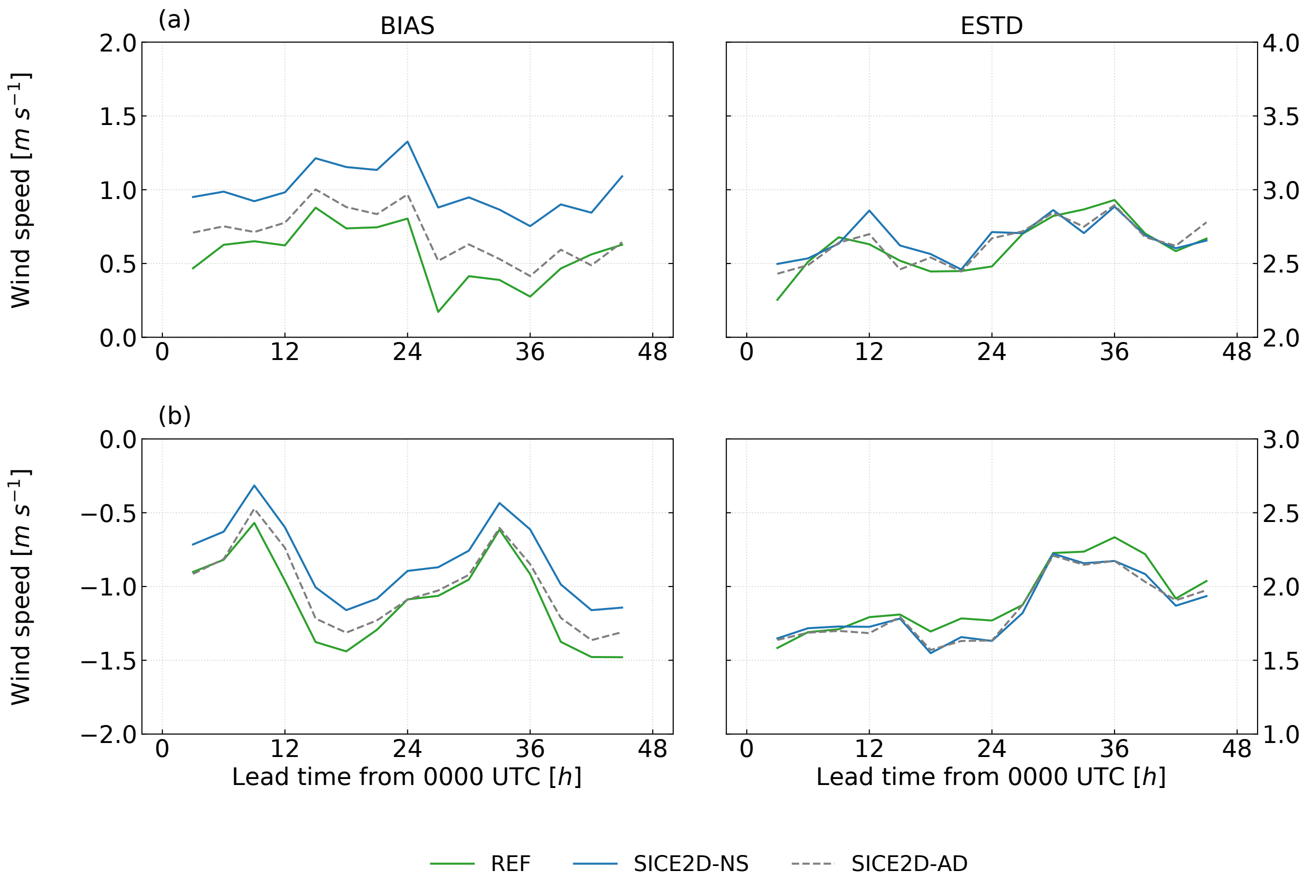

Figure 6. Mean error (BIAS, left panel) and standard deviation of errors (ESTD, right panel) of 10 metre wind speed forecasts initialized at 0000 UTC as a function of lead time for the experiments REF, SICE2D-NS and SICE2D-AD over the AROME Arctic domain for the period from 1 March 2013 to 31 March 2013. a) For 7 Svalbard stations; b) for 7 stations located in the coastal area of the Gulf of Bothnia. The mean error is calculated as the forecasted value minus the observed value.

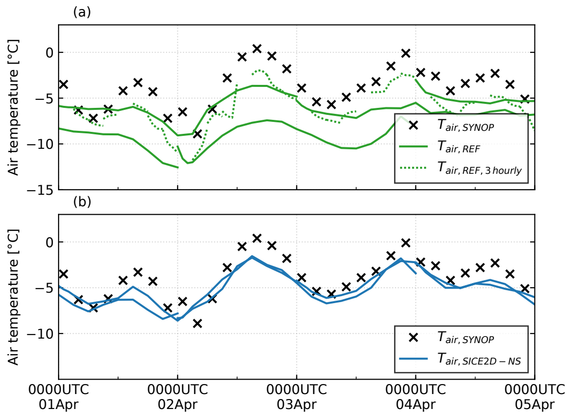

Figure 7. Evolution of the 2 metre temperature for the Kemi I lighthouse station (WMO No. 02863, 65°25' N; 24°08' E), observed (black crosses) and simulated by the 48-hour forecasts initialized at 0000 UTC for experiments: a) REF; b) SICE2D-NS. Also in (a), the simulated 2 metre temperature values for the short 3-hour forecasts initialized every 3 hours (except 1200 UTC, which is not shown) that are needed for the initialization of the long forecasts are shown (dotted lines).

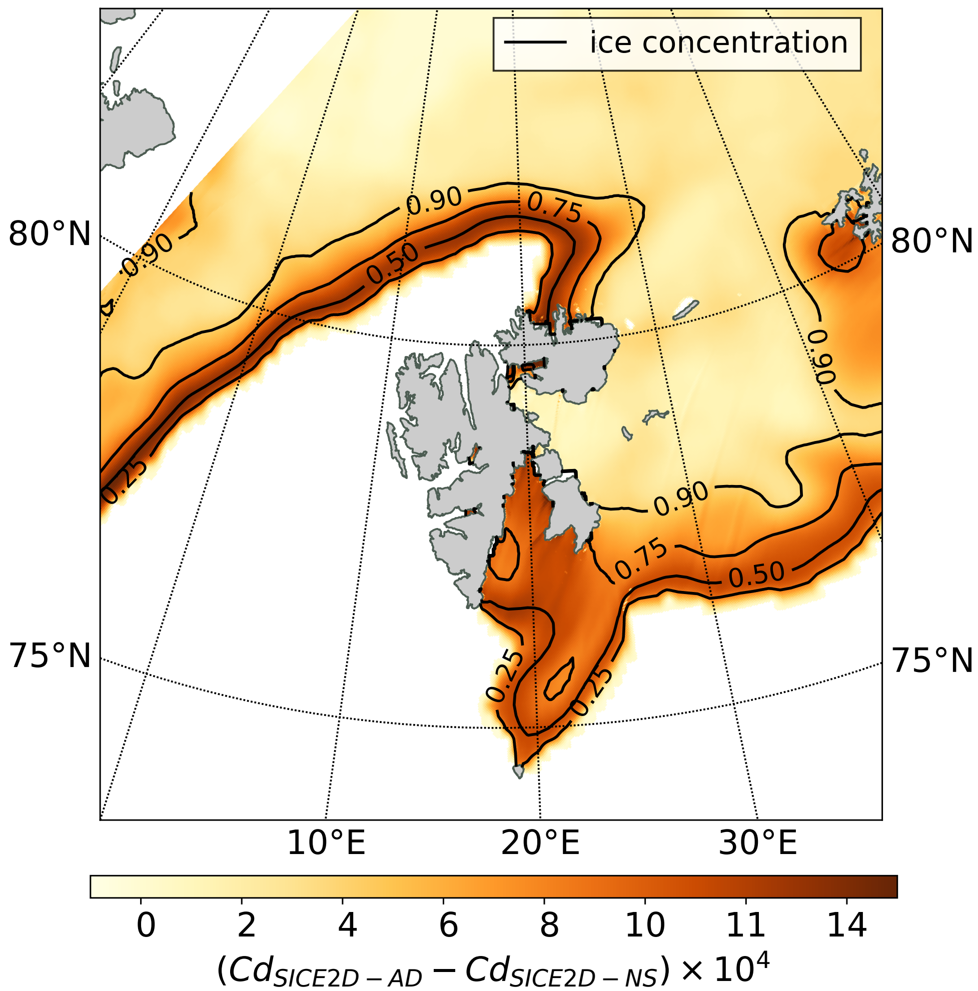

Figure 8. Impact of the form drag term on the average drag coefficient. The shading shows the difference between the average drag coefficients over ice covered areas from the SICE2D-NS and SICE2D-AD experiments for 10 March 2013 0000 UTC. Contours show the ice fraction. Open sea and land points are masked.

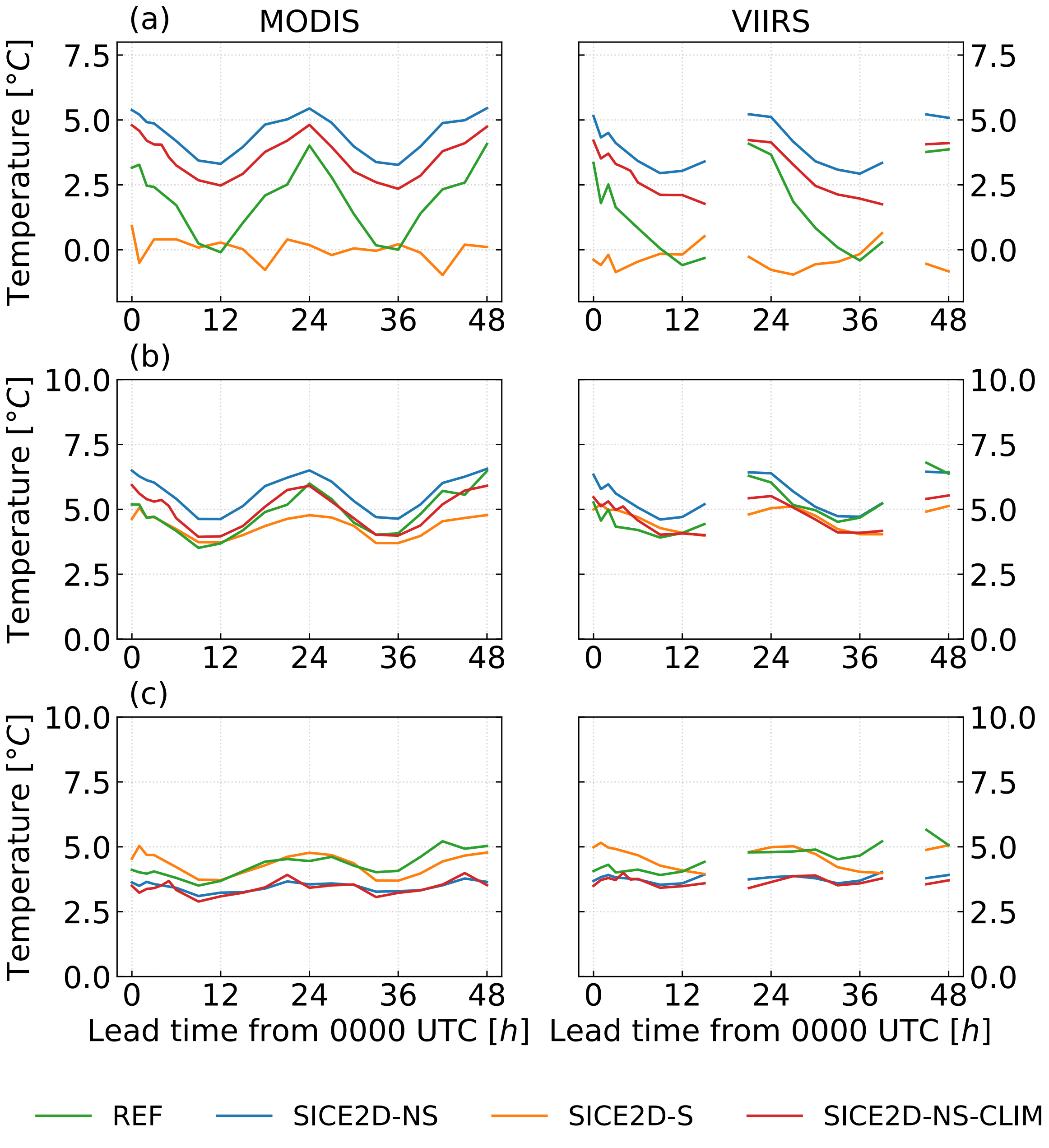

Figure 9. Mean bias, root mean square error and standard deviation of errors in the ice surface temperature forecasts initialized at 0000 UTC calculated using MODIS (left panel) and VIIRS (right panel) ice surface temperature products, as a function of lead time for the period from 1 March 2013 to 30 April 2013 for REF, SICE2D-NS, SICE2D-S and SICE2D-NS-CLIM. a) Mean bias; b) root mean square error (RMSE); c) standard deviation of errors (ESTD). VIIRS swathes for the lead times of 18 and 42 hours systematically cover only a minor part of the AROME Arctic domain. They were excluded as not representative.

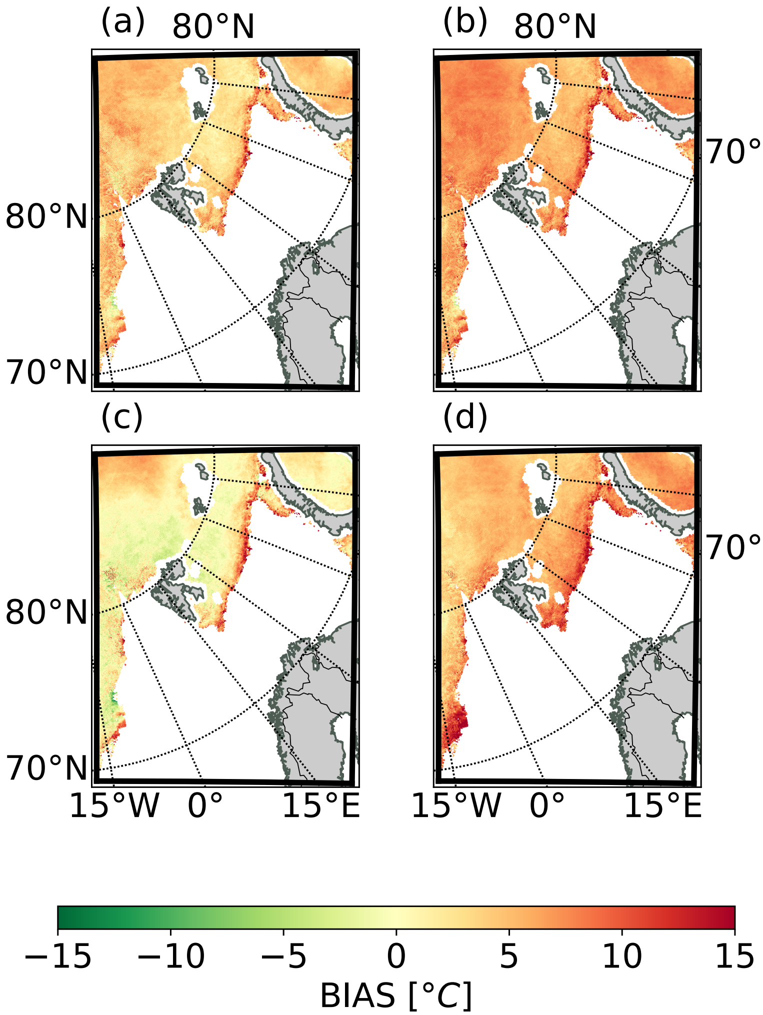

Figure 10. Spatial distribution of the mean error of the ice surface temperature after 24 hours of forecast starting at 0000 UTC for REF, SICE2D-NS, SICE2D-S and SICE2D-NS-CLIM compared to the MODIS product. Mean errors are calculated for the time period from 1 March 2013 to 30 April 2013. a) REF; b) SICE2D-NS; c) SICE2D-S; d) SICE2D-NS-CLIM.

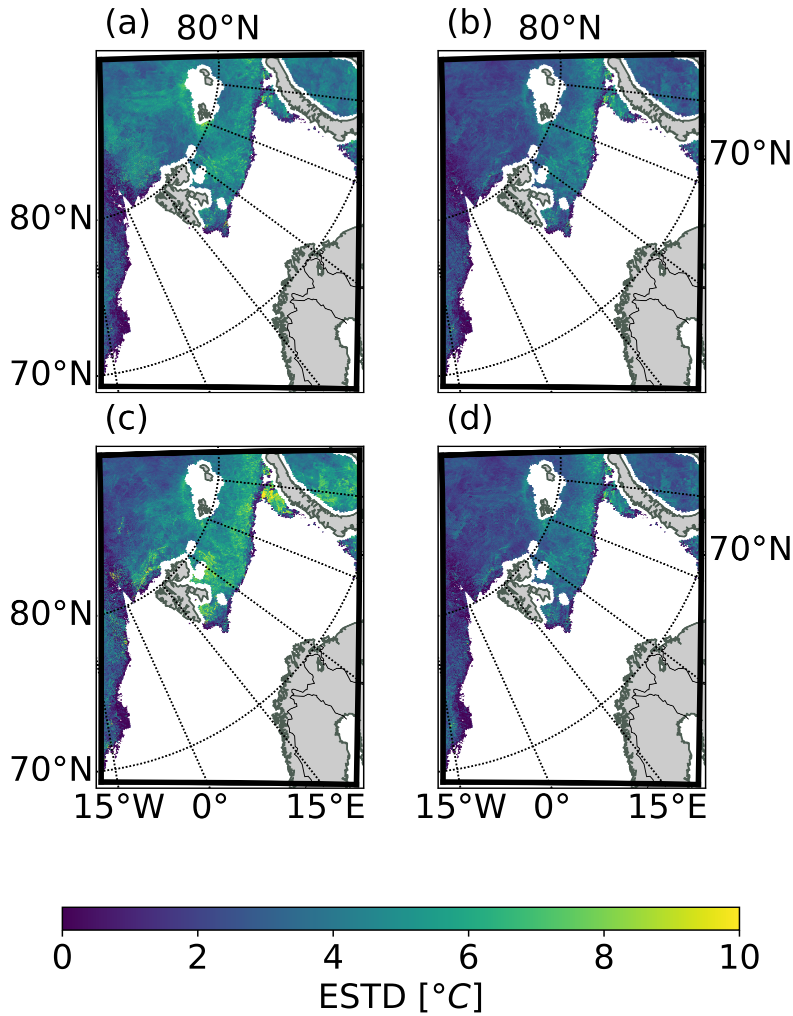

Figure 11. Same as Fig. 10 but showing the spatial distribution of the standard deviation of errors (ESTD).

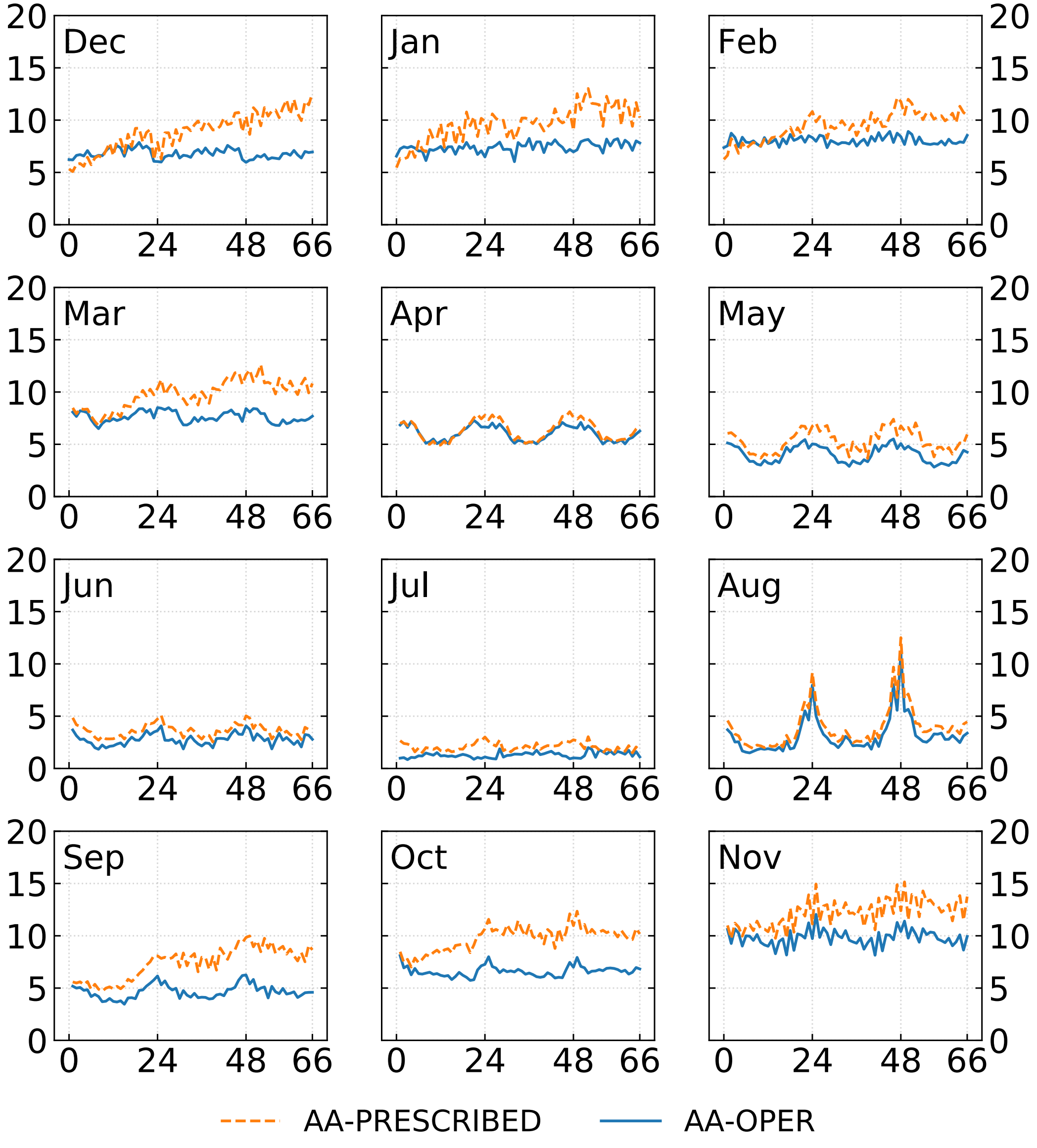

Figure 12. Monthly root mean square error (RMSE) of the ice surface temperature as a function of lead time for forecasts initialized at 0000 UTC for AA-PRESCRIBED and AA-OPER (snow-free SICE configuration). Monthly RMSE are calculated using MODIS ice surface temperature product and cover the time period from 1 December 2015 to 1 December 2016. X-axis – forecast lead time from 0000 UTC (h); Y-axis – RMSE of the ice surface temperature (°C).

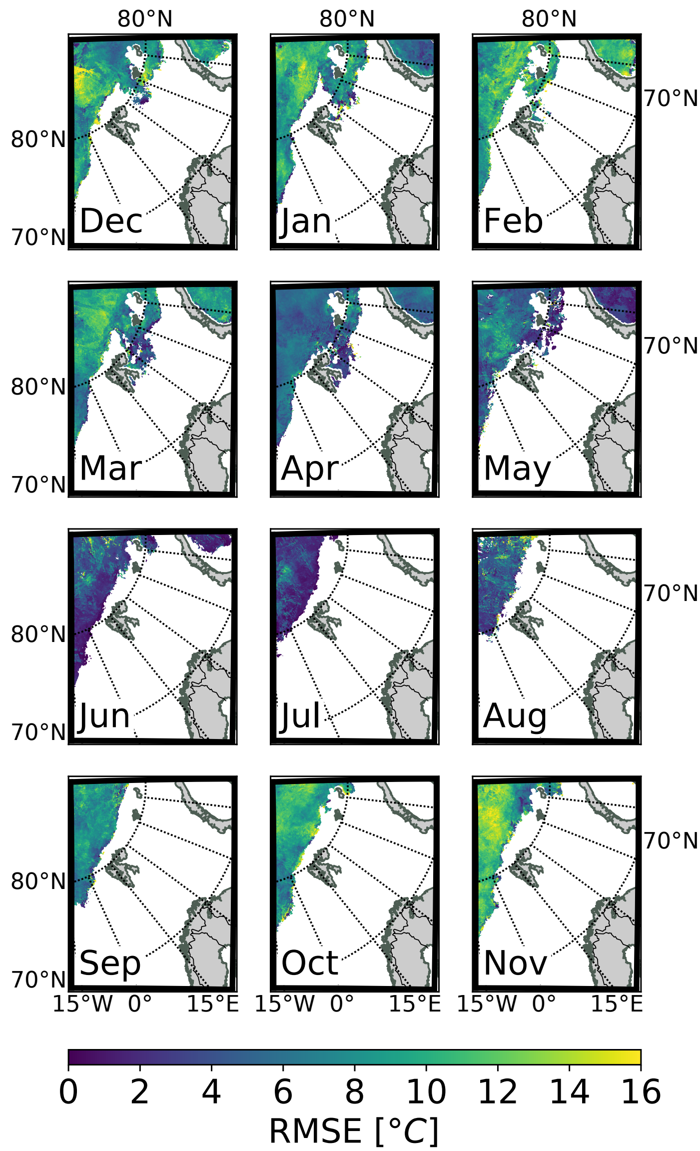

Figure 13. Spacial distribution of the monthly root mean square error (RMSE) of the ice surface temperature after 66 hours of AA-PRESCRIBED forecast initialized at 0000 UTC. Monthly RMSE are calculated using MODIS ice surface temperature product and cover the time period from 1 December 2015 to 1 December 2016.

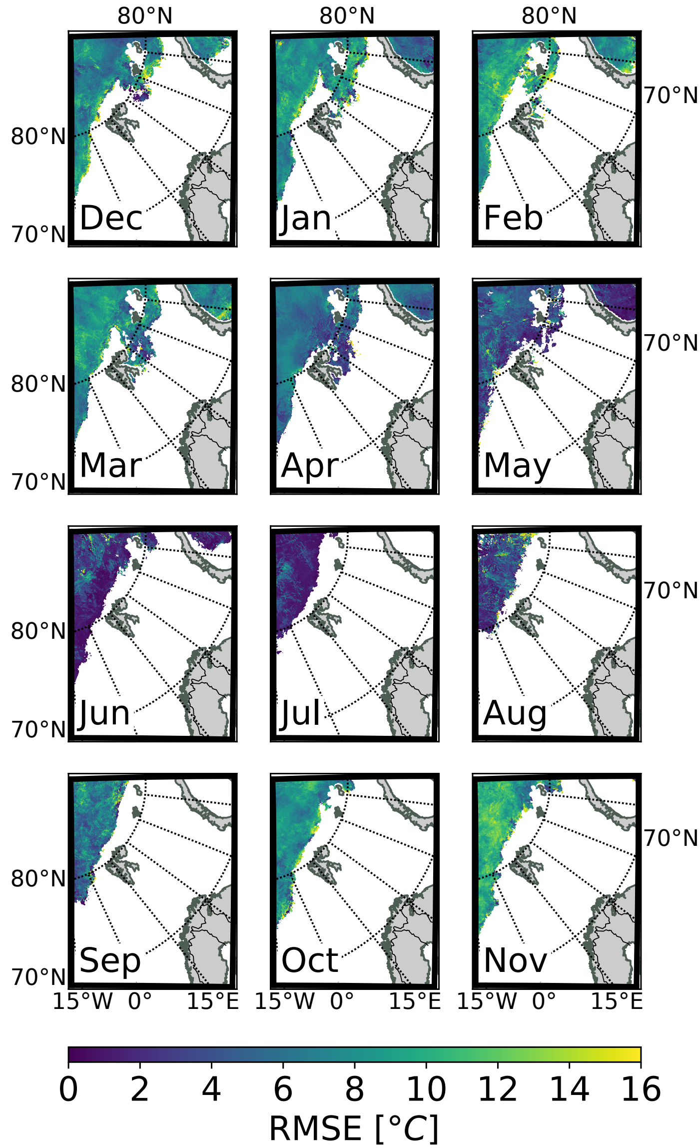

Figure 14. Same as Fig. 13, but for AA-OPER (snow-free SICE configuration).