AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL082482

An Intercomparison of Verification Scores for Evaluating the Sea Ice Edge Position in Seasonal Forecasts

Cyril Palerme[1], Malte Müller[1], and Arne Melsom[1]

[1]Norwegian Meteorological Institute, Oslo, Norway

Key Points:
- Three verification methods for evaluating the ice edge position are compared in the context of seasonal forecasting
- Various characteristics of the verification methods have been identified, which should help future forecast evaluations
- Predictive skill of a new seasonal prediction system is assessed using various verification methods

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6

Correspondence to:
C. Palerme,
cyril.palerme@met.no

Citation:
Palerme, C., Müller, M., & Melsom, A. (2019). An intercomparison of verification scores for evaluating the sea ice edge position in seasonal forecasts. *Geophysical Research Letters*, *46*, 4757–4763. https://doi.org/10.1029/2019GL082482

Received 15 FEB 2019
Accepted 18 APR 2019
Accepted article online 23 APR 2019
Published online 13 MAY 2019

Abstract With the improving skill of sea ice forecasts, verification methods are becoming increasingly important to inform model developers and users. In this study, the Modified Hausdorff Distance (MHD), the Spatial Probability Score (SPS), and a variation of the SPS are compared in order to assess their performances for evaluating the ice edge position in the new European Centre for Medium-Range Weather Forecasts seasonal forecasts (SEAS5). On average, the SEAS5 forecasts outperform a climatological reference during about 3 weeks using the MHD, and during about 5 weeks using the SPS. Furthermore, our results show that the MHD is more sensitive than the SPS to the presence of isolated sea ice patches. Moreover, the variation of the SPS introduced here is not seasonally dependent (contrary to the original SPS) and can be interpreted as a distance error of the ice edge position, which is a potentially relevant information for end users.

1. Introduction

The number of ships operating in the Arctic is rapidly increasing due to new economic opportunities made possible by sea ice decline (Arctic Marine Shipping Assessment 2009 Report, 2009; Miller & Ruiz, 2014; Melia et al., 2017). In order to plan commercial activities and reduce the risks associated with Arctic navigation, there is a growing demand for reliable sea ice forecasts at different time scales. While several institutions provide operational short-term forecasts (Sea-Ice Information Services in the World, 2017), seasonal sea ice forecasting is still at an early stage. Nevertheless, seasonal forecasts are increasingly being produced, and there are many efforts for developing and evaluating these forecasts (Stroeve et al., 2014; Smith et al., 2015; Zampieri et al., 2018).

Forecasts of various sea ice characteristics such as concentration and thickness are essential for planning Arctic navigation (Stephenson & Pincus, 2018). However, forecast evaluation requires accurate satellite observations during long time periods. While reliable ice concentration observations covering the period from 1978 to present are available, there are less ice thickness observations (Shutler et al., 2016). In addition, large uncertainties are associated with ice thickness observations (Shutler et al., 2016). The position of the ice edge is an important parameter in polar navigation, which provides information about spatial variability of ice concentration. Several studies have already addressed the question of its evaluation (Dukhovskoy et al., 2015; Goessling et al., 2016; Goessling & Jung, 2018; Melsom et al., 2019), but there is currently no consensus on how to best evaluate the ice edge position.

This study focuses on comparing various verification scores recently suggested as relevant for evaluating the ice edge position, such as the Modified Hausdorff Distance (MHD; Dukhovskoy et al., 2015), the Spatial Probability Score (SPS; Goessling & Jung, 2018), and a variation of the SPS (ratio of the SPS to the ice edge length). These verification scores are used for evaluating the ice edge position in the retrospective forecasts from the new European Centre for Medium-Range Weather Forecasts (ECMWF) seasonal prediction system SEAS5 (Johnson et al., 2019). It is worth noting that the verification scores analyzed here have not been examined by Melsom et al. (2019), who compared a set of 15 sea ice edge displacement metrics for short-term deterministic forecasts (up to 10 days). Furthermore, contrary to Melsom et al. (2019), this study focuses on the evaluation of seasonal probabilistic forecasts.

2. Data and Methods

In this study, the set of retrospective forecasts from the ECMWF seasonal prediction system SEAS5 (Johnson et al., 2019) is evaluated. The SEAS5 retrospective forecasts, covering the period 1981–2016, start on the first

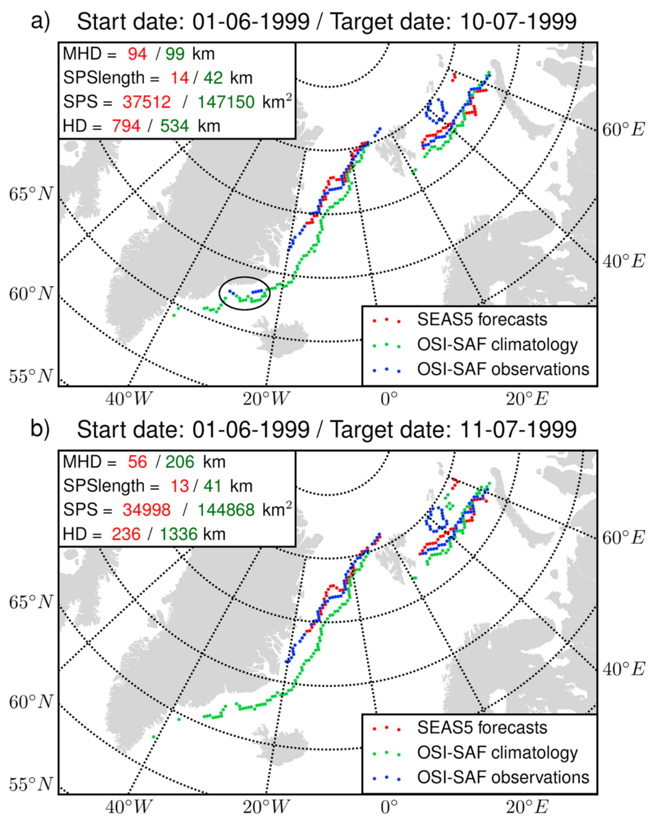

Figure 1. Sea ice edge position in the SEAS5 forecasts (red), OSI-SAF observations (blue), and OSI-SAF climatology (green) during two consecutive days (a and b) from the same SEAS5 forecast. The SPS, $\mathrm{SPS}_{\mathrm{length}}$, MHD, and HD are reported in the upper left corner of the maps in red and green for the SEAS5 forecasts and the OSI-SAF climatology respectively. The black circle in panel (a) shows the position of the isolated sea ice patches in the OSI-SAF observations. OSI-SAF = Ocean and Sea Ice Satellite Application Facility; MHD = Modified Hausdorff Distance; HD = Hausdorff Distance; SPS = Spatial Probability Score.

of every month and run for 215 days with daily outputs. Twenty-five ensemble members are produced, and the sea ice concentration is provided at a spatial resolution of 0.25°. The SEAS5 data set has been used in this study due to the long period covered by the retrospective forecasts, its large number of ensemble members, and its relatively high spatial resolution for a global seasonal prediction system.

The SEAS5 sea ice forecasts have been evaluated using satellite passive microwave observations from the version 2 of the global sea ice concentration climate data record 1979–2015 (Lavergne et al., 2019) of the Ocean and Sea Ice Satellite Application Facility (OSI-SAF). This product provides sea ice concentration estimates as well as the associated uncertainties at a spatial resolution of 25 km.

The sea ice probability (SIP), defined as the probability that sea ice concentration exceeds 15%, has been estimated for each grid cell. For the SEAS5 forecasts, the SIP has been assessed using the fraction of ensemble members with a sea ice concentration higher than 15%. For the OSI-SAF observations, the SIP has been calculated using the sea ice concentration and its associated uncertainty (assuming a Gaussian distribution). Furthermore, a climatological benchmark forecast from the OSI-SAF observations is used as a reference in this study (hereafter referred to as OSI-SAF climatology). In the OSI-SAF climatology, the SIP is assessed using the fraction of years among the ten years preceding the forecast target time with a sea ice concentration exceeding 15% during the same day of the year. The sea ice edges have then been computed using the 50% SIP contour for all the data sets.

The SEAS5 forecasts starting between 1999 and 2014 have been examined in our analysis. This period allows to compare the SEAS5 forecasts starting at the end of 2014 with OSI-SAF observations in 2015, and to use daily observations for producing the OSI-SAF climatology. In this study, the ice edge position is evaluated in the area covered by the East Greenland Sea, the Barents Sea, and the central Arctic Sea between the longitudes 45°W and 70°E (Figure S1 in the supporting information).

The verification scores evaluated are the MHD (Dubuisson & Jain, 1994), the SPS (Goessling & Jung, 2018), and a variation of the SPS. The SPS is a probabilistic verification score designed for evaluating the ice edge position and is defined as the spatial integral of the Half Brier Score

$$\mathrm{SPS} = \int_x \int_y \left(\mathrm{SIP}_{\mathrm{observations}}(x,y) - \mathrm{SIP}_{\mathrm{forecasts}}(x,y)\right)^2 \mathrm{d}y\mathrm{d}x \tag{1}$$

The SPS is the extension of the Integrated Ice Edge Error (IIEE; Goessling et al., 2016) for probabilistic forecasts and is equivalent to the IIEE if the terms "$\mathrm{SIP}_{\mathrm{observations}}$" and "$\mathrm{SIP}_{\mathrm{forecasts}}$" are replaced by binary values in equation (1) (e.g., 1 if the sea ice concentration is higher than 15% and 0 otherwise). The IIEE represents the area where the observations and the forecasts disagree on these binary values.

The Hausdorff Distance (HD) and its variations have been widely used in pattern recognition (Dubuisson & Jain, 1994; Huttenlocher et al., 1993; Sim et al., 1999). In this study, we have used the Modified Hausdorff Distance (MHD) introduced by Dubuisson and Jain (1994), which has recently been suggested as an efficient method for evaluating the sea ice edge position (Dukhovskoy et al., 2015). The MHD between two contours A and B is defined as

$$\mathrm{MHD}_{A,B} = \max\left\{ \frac{1}{|A|}\sum_{a\in A} d(a,B), \frac{1}{|B|}\sum_{b\in B} d(b,A) \right\} \tag{2}$$

where $d(a,B)$ is the minimum distance (here the Euclidean distance) between the point a ($a \in A$) and the B contour and $d(b,A)$ is the minimum distance between the point b ($b \in B$) and the A contour. $|A|$ and $|B|$

10.1029/2019GL082482

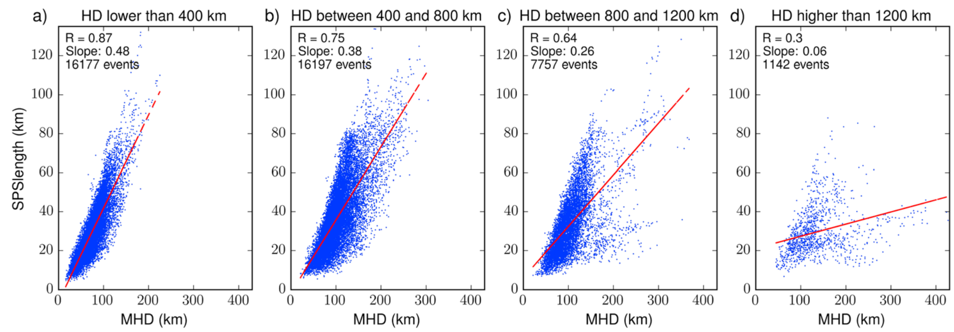

Figure 2. Scatter plots comparing the $\text{SPS}_{\text{length}}$ and the Modified Hausdorff Distance depending on the Hausdorff Distance (HD). In this figure, only the comparison between the SEAS5 forecasts and the OSI-SAF observations has been reported. The linear regressions are represented by the red lines, and R is the Pearson correlation coefficient. SPS = Spatial Probability Score; OSI-SAF = Ocean and Sea Ice Satellite Application Facility.

are the number of points on the contours A and B, respectively. In order to test the influence of isolated sea ice patches on the verification scores, the Hausdorff Distance (HD) has been used in this study to detect the presence of outliers. The HD between two contours A and B is defined as

$$\text{HD}_{A,B} = \max\left\{\sup_{a\in A} d(a,B), \sup_{b\in B} d(b,A)\right\} \tag{3}$$

The HD represents the largest of all the distances from a point in one of the ice edges compared to the other ice edge. Due to the supremum term (equation (3)), the HD is much more sensitive to outliers than the MHD (Dubuisson & Jain, 1994; Dukhovskoy et al., 2015).

Contrary to the SPS, the MHD can be used for comparing data sets which are not on the same grid. However, even for computing the MHD, the SEAS5 SIP fields have been interpolated on the OSI-SAF grid in order to compare the different verification scores in a consistent way. Furthermore, only the grid cells with no land in both data sets have been used in this analysis to ensure an unbiased comparison between the SEAS5 forecasts and the OSI-SAF observations (the resulting land-sea mask is shown in Figure S1). Therefore, coastal sea ice has been excluded in this analysis, which tends to moderate the influence of isolated ice patches on the verification scores.

Due to the different units of the evaluated verification scores (area for the SPS and distance for the MHD), we introduce the $\text{SPS}_{\text{length}}$ as the ratio of the SPS area to the ice edge length. It is worth noting that the SPS is strongly influenced by the ice edge length (see Figure S2; Goessling & Jung, 2018; Zampieri et al., 2018) and is therefore less relevant than the $\text{SPS}_{\text{length}}$ for analyzing the seasonal variation of the forecast errors. For calculating the $\text{SPS}_{\text{length}}$, the length of the ice edge has been defined as the mean value of the ice edge lengths from the two data sets (observations and forecasts). In each data set, the ice edge length has been assessed using the method introduced by Melsom et al. (2019) and described in the supporting information.

While the SPS is computed from the SIP fields, the MHD and the ice edge length used for calculating the $\text{SPS}_{\text{length}}$ are estimated from the ice edge coordinates (determined using the 50% SIP contour). Therefore, while the SPS takes into account the full range of probabilities, the MHD and the ice edge length only use the 50% probability threshold.

3. Results

3.1. Sensitivity of the Verification Scores to Isolated Ice Patches

Figure 1 shows the ice edge position from the three data sets (SEAS5 forecasts, OSI-SAF observations, and OSI-SAF climatology) and the corresponding verification scores during two consecutive days from the same SEAS5 forecast. This event has been chosen due to the large variation of the MHD compared to the other verification scores during the 2 days, and because it is easy to interpret visually. Two parts of the ice edge in the OSI-SAF observations were located between Iceland and the southeast coast of Greenland (about 68°N, 25°W) the first day, but not the second day. These ice patches in the observations contribute to decrease the MHD for the climatology due to the presence of many ice edge points in the climatology near these ice

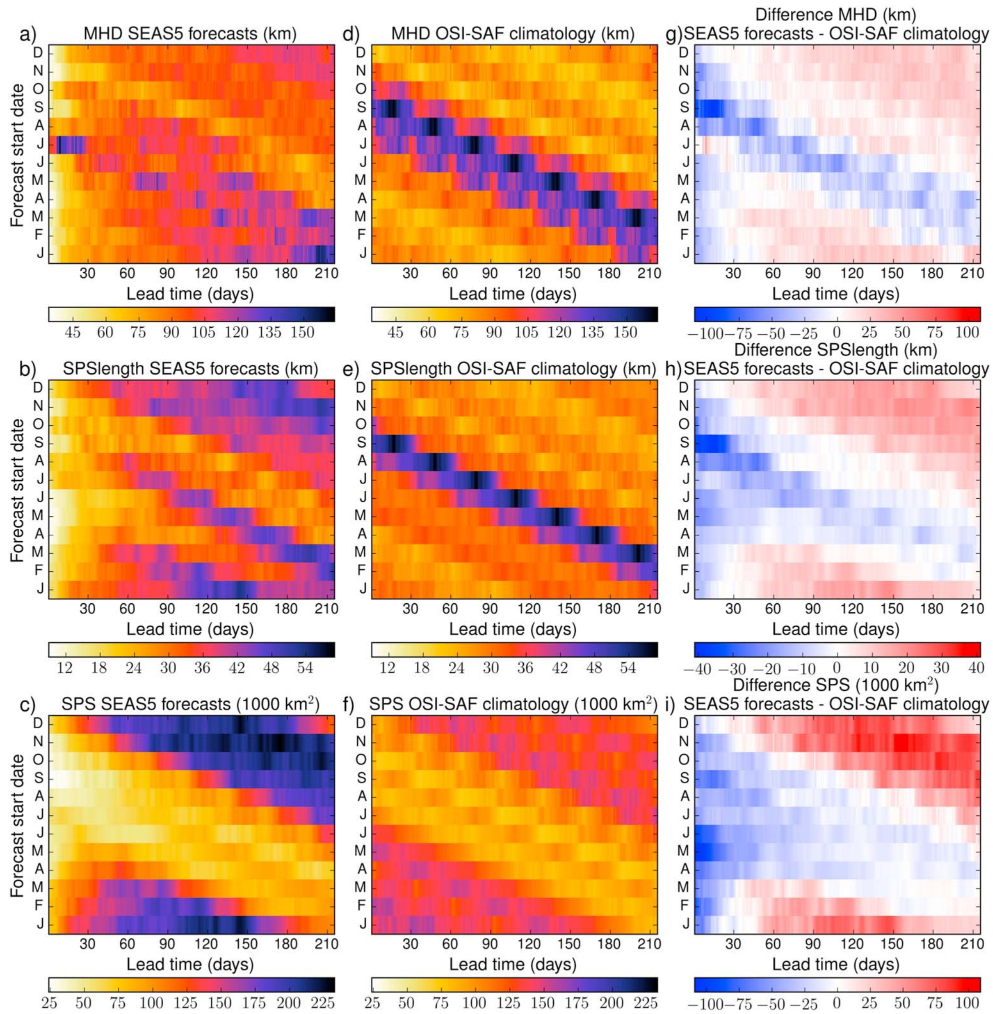

Figure 3. (a–c) Mean verification scores for the SEAS5 forecasts depending on the forecast start month and the lead time. (d–f) Mean verification scores for the OSI-SAF climatology depending on the forecast start month and the lead time. (g–i) Difference between the verification scores for the SEAS5 forecasts and the OSI-SAF climatology. OSI-SAF = Ocean and Sea Ice Satellite Application Facility; MHD = Modified Hausdorff Distance; SPS = Spatial Probability Score.

patches. In contrast, the presence of these ice patches tends to increase the MHD for the SEAS5 forecasts due to the long distance between these ice patches and the nearest ice edge point in the SEAS5 forecasts. However, the values of the SPS and the SPS_{length} were very similar during the 2 days due to the small areas of these sea ice patches. This case study shows that the MHD can be very sensitive to the presence of isolated sea ice patches, even if the area of this remote sea ice is small and therefore does not significantly influence the SPS and the SPS_{length}.

In order to test the influence of isolated sea ice patches on the verification scores, the forecasts have been sorted into four categories depending on the magnitude of the HD from the comparison of the SEAS5 forecasts and the OSI-SAF observations (Figure 2). The HD is used here as an indicator for assessing the likelihood of occurrence of isolated sea ice patches (a large HD corresponds to a high probability of occurrence of isolated sea ice patches). When the HD is lower than 400 km, meaning that there is no ice edge point further than 400 km from the other ice edge, the MHD is well correlated with the SPS_{length} (Pearson coefficient of 0.87). However, when the HD increases, the correlation between the MHD and the SPS_{length} decreases. Moreover, the values of the verification scores and the slopes of the linear regressions in Figure 2 show that the presence of outliers tends to increase the MHD much more than the SPS_{length}. Therefore, the MHD is more sensitive to outliers than the SPS_{length}, which is consistent with the case study from Figure 1.

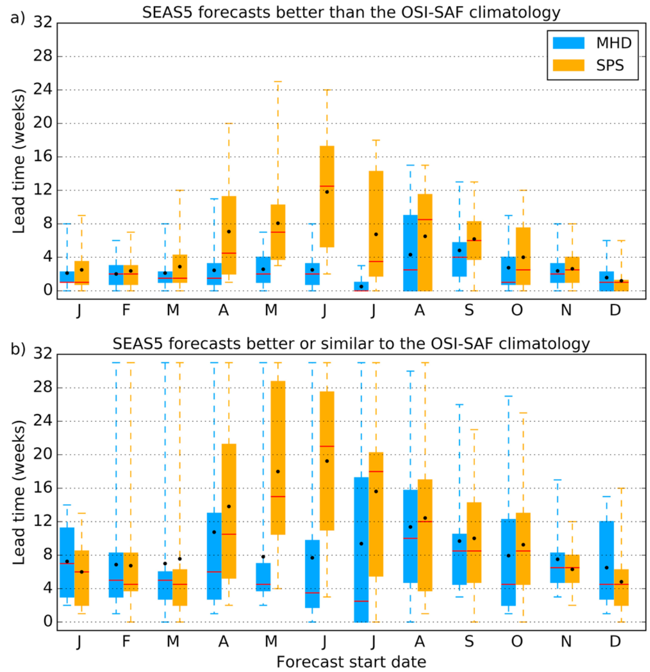

Figure 4. (a) Duration during which the SEAS5 forecasts outperform the OSI-SAF climatology within the 16 years of the analysis. (b) Duration during which the SEAS5 forecasts do not perform worse than the OSI-SAF climatology within the 16 years of the analysis. The median and the mean are shown by the red lines and the black dots respectively, the 25th and the 75th percentiles are represented by the rectangles, and the minimum and maximum values are shown by the whiskers. OSI-SAF = Ocean and Sea Ice Satellite Application Facility; MHD = Modified Hausdorff Distance; SPS = Spatial Probability Score.

3.2. Evaluation of the SEAS5 Forecasts

The variability of the three verification scores (MHD, SPS_{length}, and SPS) depending on the lead time are shown in Figure 3 for each forecast start month. The verification scores have been averaged over the 16 years of the analysis in Figure 3. It is worth noting that the mean value of the verification scores for the OSI-SAF climatology are always the same for a given day of the year, independently of the forecast start month (hence the symmetric patterns of Figures 3d–3f). The MHD and the SPS_{length} show higher forecast errors from the OSI-SAF climatology in August and September than during the rest of the year. This is in contrast to the results from the SPS, which show lower forecast errors from the climatology between July and December than during the rest of the year. These results were expected due to the strong influence of the ice edge length on the SPS (Figure S2).

The varying performances of the climatology strongly influence the difference between the verification scores for the SEAS5 forecasts and the climatology (Figures 3g–3i). In particular, using the MHD, the SEAS5 forecasts starting between February and July have a period with lower performances than the climatology before outperforming the climatology again (Figure 3g). Furthermore, the MHD and the SPS_{length} differ significantly in July for the OSI-SAF climatology, and during the first month of the lead time for the SEAS5 forecasts starting in July. This is likely due to the high fraction of events with a large HD in both data sets occurring in July compared to the rest of the year (Figure S3).

When comparing the performances of various forecasts using the SPS_{length}, the different ice edge lengths in the forecasts can introduce some bias. For example, the SPS of various forecasts could be similar, while the SPS_{length} could differ due to different ice edge lengths in the various data sets. Therefore, we consider that the SPS is more robust than the SPS_{length} for assessing the duration during which the SEAS5 forecasts outperform the climatology, and that is why the SPS_{length} has not been used in Figure 4. In order to determine the duration during which the SEAS5 forecasts outperform the climatology, the daily time series of the verification scores have been averaged over each week of the lead time. This approach allows to provide a duration with a weekly resolution, and to moderate the influence of particular daily forecasts. Moreover, the standard deviation of the verification scores have been estimated for each week of the lead time from the daily time series. For each verification score, the SEAS5 forecasts are considered significantly better than the climatology as long as the sum of the SEAS5 verification score and the two standard deviations (SEAS5 and OSI-SAF climatology) remains lower than the OSI-SAF climatology verification score (Figure S4). The SEAS5 forecasts are considered better or similar to the climatology as long as the SEAS5 verification score remains lower than the sum of the OSI-SAF climatology verification score and the two standard deviations (SEAS5 and OSI-SAF climatology).

The distribution of the durations during which the SEAS5 forecasts outperform the climatology within the 16 years of the analysis has been reported in Figure 4. The results differ significantly depending on the verification score used. Overall, the SEAS5 forecasts outperform the OSI-SAF climatology longer using the SPS than the MHD, particularly for the forecasts starting between April and August. This is consistent with the results from Figures 3g and 3i, which show a more progressive decrease of the skill of SEAS5 forecasts using the SPS than using the MHD. The higher variability of the MHD compared to the SPS could also explain this difference (Figure S5).

Using the SPS, the SEAS5 forecasts starting between April and September have the best performances, with a mean duration during which they outperform the climatology ranging from 6 to 12 weeks. On the other hand, the SEAS5 forecasts starting between October and March outperform the climatology during less than 4 weeks on average. The SEAS5 forecasts starting in June are the best compared to the climatology, with almost 12 weeks on average during which they outperform the climatology, and they are not worse than the climatology during more than 19 weeks on average. Nevertheless, the interannual variability of these results is large, meaning that there is a high variability in the performances of individual SEAS5 forecasts.

4. Discussion and Conclusions

In this study, the position of the sea ice edge in the ECMWF SEAS5 retrospective forecasts has been evaluated using various verification scores. The SPS is correlated to the length of the ice edge (Goessling & Jung, 2018; Zampieri et al., 2018) and is therefore not suitable for analyzing the seasonal variation of the forecast errors without normalization. The SPS_{length} and the MHD are more relevant verification scores for comparing the forecast errors during different seasons. Nevertheless, when predictive skill is assessed by comparing the forecasts to a climatology, we recommend to use the SPS instead of the SPS_{length} since the ice edge lengths are different in the two data sets. Furthermore, it has been shown in this study that the MHD is more sensitive to the presence of isolated sea ice patches than the SPS and the SPS_{length}. The sensitivity of the verification scores to isolated ice patches is an important parameter to take into account when evaluating the forecast errors. Depending on the application, this can be a relevant or inadequate characteristic. For example, a high sensitivity to isolated ice patches can be suitable for evaluating the forecast ability to reproduce coastal sea ice, while this can be inadequate for comparing the general agreement between the forecast and observed ice edge positions.

The duration during which the SEAS5 forecasts outperform the climatology differs greatly depending on the verification scores used. Overall, the SEAS5 forecasts perform better compared to the climatology using the SPS than the MHD. The SEAS5 forecasts starting between April and September have the best performances compared to the climatology using the SPS. They outperform the climatology during more than 6 weeks on average, with the best performances for the forecasts starting in June (about 12 weeks on average). However, there is a large variability in the performances of individual SEAS5 forecasts. Furthermore, due to the varying performances of the climatology, it is also important to evaluate the forecasts using the value of the verification scores in addition to the relative score compared to the climatology.

A useful information for decision making is the mean distance between the forecast and observed ice edges. This information can be provided by the MHD, but the MHD can also be overly sensitive to isolated ice patches depending on the application. The SPS_{length} has the same unit as the MHD, but it is lower by a factor of about 2 (Figure 2a). The lack of parallelism between the ice edges tends to reduce the value of the SPS_{length} compared to the mean distance between the forecast and observed ice edges. In addition, this difference could also be explained by the different ranges of probabilities taken into account for calculating the MHD (only the 50% threshold) and the SPS_{length} (full range of probabilities). Goessling and Jung (2018) have reported that using a deterministic forecast (defined from the 50% threshold for the SIP) instead of ensemble forecasts produces an increase of the SPS of about 36%. We have also evaluated this characteristic (Figure S6), by comparing the SPS_{length} to the ratio between the IIEE and the ice edge length ($IIEE_{length}$).

The $IIEE_{length}$ has been calculated using binary values from the 50% threshold for the SIP (1 if the SIP is higher than 50% and 0 otherwise). We have found that the $IIEE_{length}$ and the SPS_{length} are very well correlated (Pearson coefficient of 0.97) and that the $IIEE_{length}$ is about 44% larger than the SPS_{length}, which is consistent with the findings from Goessling and Jung (2018).

There is a need of estimating the reliability of sea ice forecasts at various spatial scales, and it would be more challenging to evaluate the ice edge position in small areas due to the limited number of ice edge points. Among the verification scores used in this study, we consider that only the SPS could be reliable for this application because it is calculated from the SIP and not from ice edge coordinates. Due to the limitations of the SPS, further developments are needed in order to provide relevant information to end users at local scales.

Acknowledgments
M. Müller and A. Melsom received funding from the Nansen Legacy project, which is funded by the Norwegian Research Council under contract 276730. A. Melsom also received funding from the Copernicus Marine Environmental and Monitoring Service under Mercator Océan contract number 2015/S 009-011301. C. Palerme and M. Müller acknowledge research funding from Salienseas project funded by the Norwegian Research Council contract number 276223. This is a contribution to the Year of Polar Prediction (YOPP), a flagship activity of the Polar Prediction Project (PPP), initiated by the World Weather Research Programme (WWRP) of the World Meteorological Organisation (WMO). All data analyzed in this study are freely available. The OSI-SAF sea ice concentration product can be downloaded from the MET Norway FTP server (ftp://osisaf.met.no/reprocessed/ice/conc/v2p0), and the ECMWF SEAS5 retrospective forecasts are available on the Copernicus Climate Change Service (C3S), through its Climate Data Store (https://cds.climate.copernicus.eu). Finally, we thank the two anonymous reviewers for their comments that significantly improved the manuscript.

References

Arctic Marine Shipping Assessment 2009 Report (2009). Arctic council.

Dubuisson, M.-P., & Jain, A. K. (1994). A Modified Hausdorff Distance for object matching. In *IEEE Proceedings of 12th international conference on pattern recognition* (pp. 566–568).

Dukhovskoy, D. S., Ubnoske, J., Blanchard-Wrigglesworth, E., Hiester, H. R., & Proshutinsky, A. (2015). Skill metrics for evaluation and comparison of sea ice models. *Journal of Geophysical Research: Oceans*, *120*, 5910–5931. https://doi.org/10.1002/2015JC010989

Goessling, H. F., & Jung, T. (2018). A probabilistic verification score for contours: Methodology and application to Arctic ice-edge forecasts. *Quarterly Journal of the Royal Meteorological Society*, *144*(712), 735–743.

Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E., & Jung, T. (2016). Predictability of the arctic sea ice edge. *Geophysical Research Letters*, *43*, 1642–1650. https://doi.org/10.1002/2015GL067232

Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993). Comparing images using the Hausdorff Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(9), 850–863.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). Seas5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., et al. (2019). Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, *13*(1), 49–78.

Melia, N., Haines, K., & Hawkins, E. (2017). Implications from opening arctic sea routes. Report, Foresight: Future of the Sea Government Office for Science, University of Reading, Gov.uk, http://centaur.reading.ac.uk/75321/

Melsom, A., Palerme, C., & Müller, M. (2019). Validation metrics for ice edge position forecasts. *Ocean Science Discussions*, *2019*, 1–27.

Miller, A. W., & Ruiz, G. M (2014). Arctic shipping and marine invaders. *Nature Climate Change*, *4*(6), 413–416.

Sea-Ice Information Services in the World (2017). edition 2017 World Meteorological Organization, WMO-No. 574.

Shutler, J. D., Quartly, G. D., Donlon, C. J., Sathyendranath, S., Platt, T., Chapron, B., et al. (2016). Progress in satellite remote sensing for studying physical processes at the ocean surface and its borders with the atmosphere and sea ice. *Progress in Physical Geography*, *40*(2), 215–246.

Sim, D. G., Kwon, O. K., & Park, R. H. (1999). Object matching algorithms using robust Hausdorff Distance measures. *IEEE Transactions on image processing*, *8*(3), 425–429.

Smith, G. T., Jung, T., Gordon, N., Klebe, S., Goessling, H. F., Bauer, P., et al. (2015). The Year of Polar Prediction (YOPP): Challenges and opportunities in ice-ocean forecasting. *Mercator Ocean Quarterly Newsletter*, *51*, 9–12.

Stephenson, S. R., & Pincus, R. (2018). Challenges of sea-ice prediction for arctic marine policy and planning. *Journal of Borderlands Studies*, *33*(2), 255–272.

Stroeve, J., Hamilton, L. C, Bitz, C. M., & Blanchard-Wrigglesworth, E. (2014). Predicting september sea ice: Ensemble skill of the search sea ice outlook 2008–2013. *Geophysical Research Letters*, *41*, 2411–2418. https://doi.org/10.1002/2014GL059388

Zampieri, L., Goessling, H. F., & Jung, T. (2018). Bright prospects for arctic sea ice prediction on subseasonal time scales. *Geophysical Research Letters*, *45*, 9731–9738. https://doi.org/10.1029/2018GL079394