1 Forecast verification metrics

A robust verification scheme is essential to gain insight into how the developed forecasting product performs. Both from the point of view of a developer which aim to increase the skill of the prediction but also from the user which may utilize the verification score to assess the quality of a given forecast [1]. In the context of Sea Ice forecasting, a spatial field of continuos or discrete sea ice concentration is predicted, the latter being the case for the current work. Given the uneven distribution intra sea ice concentration classes as well as sea ice compared to ice free open water, simply comparing pixels for correctness would be biased by the large portion of open water and result in difficult to interpret values devoid of physical reasoning. Furthermore, as the rate of maritime activity such as commercial shipping increases in the Arctic due to the sea ice decline [5], having user relevant metrics can aid and alleviate the risks surrounding Arctic navigation. As such, several studies have proposed calculating the position of the ice edge as a user relevant metric which also provides information of the distribution of the Sea Ice Concentration [2, 4, 3]. However, there is no agreement with how to best calculate the position of the Ice Edge, with the currently available metrics posing different advantages/disadvantages [8, 7]. For the purpose of this thesis, The ice edge position and length will be calculated according to [7, Melsom 2019 et.al], whereas the IIEE originally proposed by Goessling. H. [4] will also be utilized.

Syk udokumentert påstand, må modereres

1.1 Defining the Ice Edge

The ice edge for a given Sea Ice Concentration product is derived on a per pixel basis, and defined as the grid cells which meet the condition

$$c[i,j] \geq c_q \wedge \min(c[i-1,j], c[i+1,j], c[i,j-1], c[i,j+1]) < c_e \tag{1}$$

i.e. a pixel is marked as a ice edge pixel if the current pixel itself is larger than some given concentration threshold c_e and the minimum of the pixel's 4-neighbors is less than the same threshold. Moreover, the marked grid cells each contribute to the total length of the ice edge, with each pixel's length contribution determined based on the number neighbors also marked as an ice edge pixel. Consequently, a neighborless pixel is assumed to yield a contribution the length of the diagonal to the ice-edge ($l = \sqrt{2}s$) where s is the side length of the pixel. A pixel with one neighbor a contributes a mixed horizontal - diagonasl length $l = \frac{s+\sqrt{2}s}{2}$. Finally a pixel with two or more neighbors contributes with a pixel side-length $l = s$.

1.2 Integrated Ice Edge Error

The Integrated Ice Edge Length (IIEE) is an error metric which compares the forecast to some ground truth target [4]. The metric is defined as

$$\mathrm{IIEE} = \mathrm{O} + \mathrm{U} \tag{2}$$

where

$$\mathrm{O} = \int_A \max(c_f - c_t, 0)dA \tag{3}$$

and

$$\mathrm{U} = \int_A \max(c_t - c_f, 0)dA \tag{4}$$

with c_t and c_f being the target and forecast concentration respectively, attaining a value of 1 if the concentration for a given pixel i above a set threshold, and 0 elsewhere. From the definition of the metric, it can be seen that the IIEE is a sum of the forecast overshoot and undershoot compared to the ground truth target. For the current work, the IIEE is an easily interpreted metric as it quantifies the total forecast error and reports on the error spatially.

Furthermore, the IIEE can be combined with the length of the Ice Edge which was derived in the previous section 1.1. Thus, the metric is seasonally normalized, assuming that the IIEE and Ice Edge Length is seasonally correlated.

2 Impact of increased resolution on the IIEE

From both the definition of Equation (1) and (2), it can be seen that there is a dependance on the number of pixels which constitutes the ice edge. However, what effect would a change of resolution, i.e. change in number of pixels, have on the IIEE / ice edge length ratio? To answer this question, the "two day" ice edge targets where compared against persistance on all valid two day forecasts samples for the period 2019 - 2021. Furthermore, the original target resolution of 1km will be assessed, as well as a regrided product downscaled onto a 10km resolution.

The Pearson correlation coefficient was computed directly from the DataFrames containing the metrics for both 1km and 10km resolution, using the statistical Python package Pandas [9, 6]. For clarity, the correlation between mean_length and IIEE for 1km and 10km was calculated. From the resulting computations, the correlation coefficient gets reported as $r_{mean\_length} = 0.9458$ and $r_{IIEE} = 0.9995$. Furthermore, the IIEE divided by the mean length correlation is reported as $r_{normalized\_IIEE} = 0.9747$. Finally, Figure (1) display the mean monthly IIEE computed along the inspected three year period.

By inspecting Figure (1) in conjunction with the reported correlation coefficients, it can be seen that increasing the working resolution of the dataset from 1km to 10km has a negligible impact on the reported metrics. Though the 1km ice edge is about 13 times longer than the 10km, coarser resolution ice edge, it does not impact the overall stability of the normalized IIEE metric. However, the normalized IIEE is reduced for the 1km grid, as a consequence of the increased ice edge length due to resolution. Note that the from the definition of the IIEE given in Equation (2), the unnormalized metric is dependent on pixel spatial size. Thus, the number of pixels constituting the sum is inverse proportional with the size of each pixel, hence keeping the stability of the values. However, a discrepancy may arise due to the coarser resolution pixels covering a larger area, thus losing out on the fine details. However, for the current work, the ration of 1km IIEE and 10km IIEE is 0.9986, indicating that they are close to equal.

References

[1] B. Casati et al. "Forecast verification: current status and future directions". In: *Meteorological Applications* 15.1 (2008), pp. 3–18. DOI: 10.1002/met.52.

[2] Dmitry S. Dukhovskoy et al. "Skill metrics for evaluation and comparison of sea ice models". In: *Journal of Geophysical Research: Oceans* 120.9 (Sept. 2015), pp. 5910–5931. DOI: 10.1002/2015jc010989.

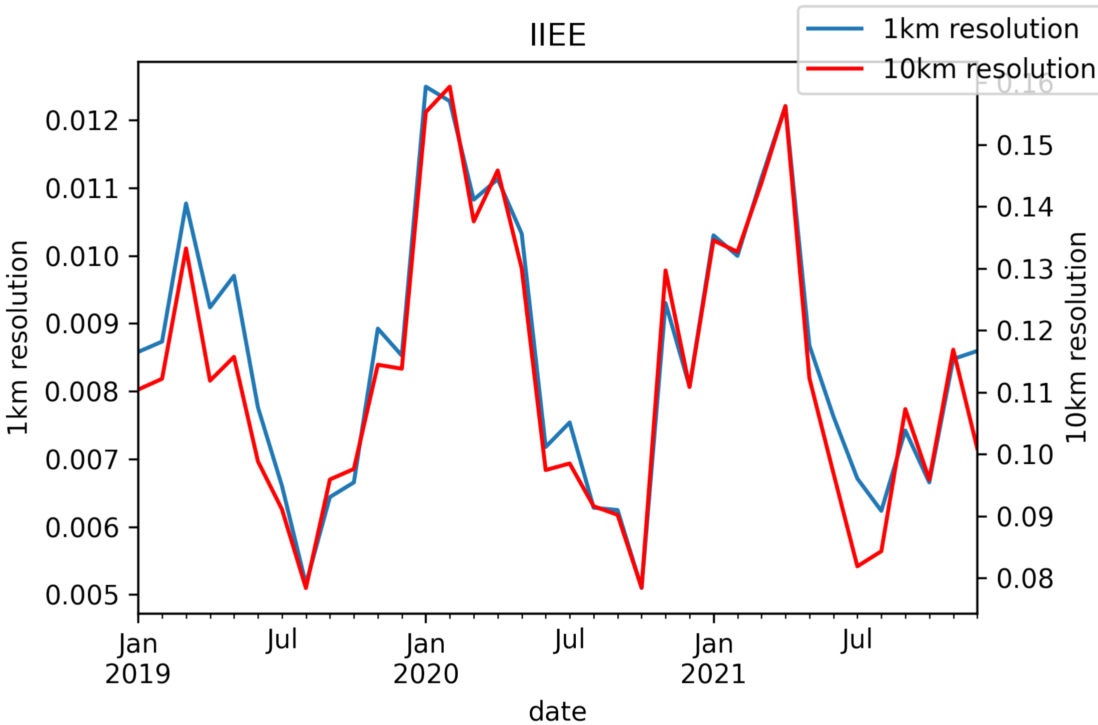

Figure 1: Monthly mean of Normalized IIEE spanning 2019 - 2021

[3] H. F. Goessling and T. Jung. "A probabilistic verification score for contours: Methodology and application to Arctic ice-edge forecasts". In: *Quarterly Journal of the Royal Meteorological Society* 144.712 (Apr. 2018), pp. 735–743. DOI: 10.1002/qj.3242.

[4] H. F. Goessling et al. "Predictability of the Arctic sea ice edge". In: *Geophysical Research Letters* 43.4 (Feb. 2016), pp. 1642–1650. DOI: 10.1002/2015gl067232.

[5] Joshua Ho. "The implications of Arctic sea ice decline on shipping". In: *Marine Policy* 34.3 (May 2010), pp. 713–715. DOI: 10.1016/j.marpol.2009.10.009.

[6] Wes McKinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[7] Arne Melsom, Cyril Palerme, and Malte Müller. "Validation metrics for ice edge position forecasts". In: *Ocean Science* 15.3 (May 2019), pp. 615–630. DOI: 10.5194/os-15-615-2019.

[8] Cyril Palerme, Malte Müller, and Arne Melsom. "An Intercomparison of Verification Scores for Evaluating the Sea Ice Edge Position in Seasonal Forecasts". In: *Geophysical Research Letters* 46.9 (May 2019), pp. 4757–4763. DOI: 10.1029/2019gl082482.

[9] The pandas development team. *pandas-dev/pandas: Pandas*. Version latest. Feb. 2020. DOI: 10.5281/zenodo.3509134. URL: https://doi.org/10.5281/zenodo.3509134.