2020IVMWO

SEA-ICE session, 15:00-17:00 UTC, 13th Nov

Discussion Notes

Q = Question
A = Answer
C = Comment

###
Talk 1: Peterson, Andrew, Environment and Climate Change Canada:
Using Integrated Ice Edge Error (IIEE) and Spatial Probability Score (SPS) to assess spread-error relationships in an ensemble sea ice forecast

Q: Are the metrics computed on a grid-point basis?
A: Yes, Brier scores are computed per grid cell, then integrated.

Q: Have you considered using neighbourhood methods instead of grid-point methods?
A: No.
C: The neighbourhood aspect (relaxing the requirements of exact spatial co-location of forecast and observation) is implicitly taken care of by the probabilistic approach.
C: Yes, that is likely true -- if the ensemble spread is representative of the errors in co-location. Likely, still areas where this might not be entirely true.

C: Suggestion to apply IIEE to the 50% SIP contour (a.k.a. the “median ice edge”) rather than to the 15% contour of the ensemble-mean SIC. Otherwise the ice extent is positively biased toward the ensemble members with large ice extent.

Q: How was your IIEE-based ensemble spread defined?
A: IIEE of each ensemble member relative to observed edge, then averaging IIEE values. Then, SPS minus that average IIEE gives the spread.
C: Some discussion around this definition of spread may follow offline.

###
Talk 2: Zampieri, Lorenzo, Alfred Wegner Institute, Germany:
Verification of subseasonal sea-ice prediction at both poles

Q: The initial errors are quite large, although the verifying analysis should be very close to the analysis used for the initialization. What lead time corresponds to the very first data point?
A: First day. Note that all fields were interpolated to the coarse S2S dataset grid (1.5°); high-resolution evaluation might show smaller initial error (or not?). No clear answer for the large initial errors yet.

Q: Each centre uses its own product, some gridded observational product or an analysis (that incorporated an observational product). Should one make a systematic comparison of ice edge discrepancies among them, similar to what Drew showed in his presentation?
A: Good idea.
(See general discussion, where this was followed up on.)

C: Regarding initial edge error, the interaction with SST assimilation might also play a role.

C: Maybe also the orography plays a role: Arctic ice is bounded, being land-locked by the continents (in winter it is easy to correctly predict full cover), whereas Antarctic ice has no boundaries.
C: Maybe, but in Arctic in winter, in the Atlantic sector, the edge is still quite long. But yes, in the Antarctic the ice may be more susceptible to wind shifts (e.g. jet stream related). In this context, it would be interesting to know the relative contributions of ocean preconditioning versus atmospheric predictability on the Antarctic sea ice predictability.

Q: Do systems assimilate observation products directly?
A: In some cases intermediate analyses are involved, e.g. OSTIA at ECMWF. But some assimilate observations directly, e.g. UKMO.

Q: Are S2S data available for everyone to use?
A: Yes! There is just a 3-weeks delay for current forecasts. All reforecasts all available. Information on the database can be found here: https://confluence.ecmwf.int/display/S2S .

Q: Can the SPS be decomposed like the IIEE?
A: One can first split into IIEE and non-IIEE contributions (IIEE contribution based on the 50% SIP contour, SPS minus that is the non-IIEE contribution to the SPS). The IIEE contribution can then be decomposed as usual).

Q: Is the systematic error part (AEE fraction) larger in the Antarctic?
A: Not clear.

###
Talk 3: Niraula, Bimochan, Alfred Wegner Institute, Germany:
Reference forecast of sea-ice edge using damped persistence of probability anomaly

Q: In the S2S analysis, lots of the marginal seas are masked out. Have you looked at the performance of your damped anomaly persistence in marginal/regional seas?
A: Yes. Actually some of the results presented used the better-resolved OSISAF mask. Also, we have repeated the S2S analysis with a less conservative mask (excluding one of the models with a particularly conservative mask), revealing that our benchmark performs even slightly better in those near-coastal regions.

###
Talk 4: Cheng, Angela, Environment and Climate Change Canada:
A symmetric spatial verification method for sea ice edges

C: Graphs along the ice edge, showing where there are positive and where negative deviations, are nice!

Q: Modified/Partial Hausdorff are not distance metrics (do not fulfil triangle inequality) -> Is the latter actually needed?
A: Yes! Otherwise the forecast is not proper.
Q: So should MHD/PHD not be used?
A: Unclear.
C: Regarding the new metric presented in this talk, it's not been investigated yet whether this actually is a metric in the mathematical sense.

Q: Have you compared with / considered dynamic time warping, Baddeley delta-metric, and/or to replace the nearest distance by generalized distance transform in any Haussdorff distance family (should be more robust)?
A: Yes, except the latter one.

Q: Is the "forced sequencing" always the more intuitive/accurate way to measure distance?
A: From discussions with forecast users, this seems to be the case.

Q: Will the code be made available publicly, e.g., on GitHub?
A: Probably yes, after more checking and publication.

C: The method automatically gets rid of isolated ice patches.
C: Computing a distance passing over land in these types of metrics could in principle be avoided by looking for closest ocean paths (geodesic), but that may be complex to implement (maybe something's available for robot path planning.
C: Not needed if coasts are included when looking for closest neighbours.

Q: How are isolated patches removed?
A: By feature matching. When there is no match, features are removed.
Q: How about displaced polynya, not overlapping?
A: Will end up being separate features.

###
General discussion

Q: Is the SPS more a spatial CRPS, or integrated Brier score?
A: Both. CRPS correspondence becomes less stringent when SIP does not evolve monotonically from zero to one along a transect (like a cumulative distribution function used for CRPS).

Main discussion point: Differences between gridded observational products and derived analyses (focus on ice edge)
C: Sea-ice verification depends on the choice of gridded sea-ice product used, and there are quite some uncertainties between them. And these are also different from the different analyses of various systems. In the weather context there's a study showing that verifying against one's own analysis is always introducing a spurious advantage, see Parks et al.
C: For S2S verification, both OSISAF & Univ. Bremen AMSR2 (higher res) were used, with small impact.

C: At ECCC, huge difference between SIC analysis and OSISAF, see Drew's talk. Also, that was similar for different thresholds. OSISAF appears to have some melt ponds as open water. Note that the IMS analysis is actually just a binary product. It seems that the concentration definition is not great for various products. Differences between the IMS and CMC analyses are quite large, and the effective/optimal concentration threshold to binarize seems to change with time.
C: ORA-IP comparison (sea-ice in ocean reanalyses): Large differences/errors even in ice-assimilating models.
C: Would be great to have a systematic comparison.
Q: In S2S ice-edge studies, why initial errors in most cases very similar, at 0.25? Might be related to same obs products assimilated?
A: Open question.
C: There's a SIC observation products comparison paper: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6782670 . Also, looking at analyses is sobering: a lot of variations!

- ➔ Topic to be carried further in YOPP sea-ice task team. Could make a nice community paper.

C: Generally, other variables should be considered more, too, e.g. thickness, pressure, age, snow depth. These are all important forecast quantities. Bimo did some work on thickness verification when he was at ECCC.
C: For thickness, not as many observations available as for concentration, and those that do exist do not go far into the past, e.g., Cryosat about 10 years. However, SMOS (0-50cm) is also interesting, and probably more relevant for navigation. Using thickness contours should also be done, easy follow-up to ice-edge studies, same metrics could be used.

C: There has been a lot of work on ice verification metrics (in particular ice edge) recently. One should always use a broad "portfolio". If they show very different things we need to be question the metrics. Also: Use different validation products. Seek for consensus across metrics and across verifying observational data sets.

Q: Can SPS be adopted for verification of atmospheric variables?
A: Might work nicely for things like jet stream location, maybe not so well for patchy things like convective precipipitation. However, given that we're talking about ensemble forecasts, which involve "ensemble smoothing" intrinsically, that might actually also be reasonable.

Q: Could the MetVerif package be supplemented with the sea-ice verification metrics discussed here?
A: Good idea.
C: Ocean Predition Center / NCAR trying to incorporate many new verification metrics, including for ocean, into their package (MetPlus). Happy to have more collaborations/contributions -> sessions next week, more discussion there.