## **Reply to referee 1**

We would sincerely like to thank the reviewer for the comments. We have made the suggested changes. We have also put the table information directly into the latex file so it no longer shows up with a line underneath.

## **Reply to referee 2**

We would sincerely like to thank the reviewer for the very thorough job and excellent comments. We provide responses below for the items that required a response

### **General comments**

We have reviewed and revised Section 7.2.3, 8, 9 and 10 to improve readability as suggested by the reviewer.

B) We did not test daily data. On the one hand, the reviewer has made the good point that by only using the noon data we may be missing events happening at different times of the day. However, averaging the states would blur these events (e.g., an opening in the ice cover). On the other hand, by training the model with a relatively large set of noon data, we may have a representative selection of instantaneous events in the training data.

We thank the reviewer for bringing up this point. It relates to the question of time-correlated events, which would be interesting to investigate in a separate study.

C) Thank you for pointing this out. We have double checked the figure numbers

D) We agree extending the break-up date period analysis would have simplified the analysis and improved the final results. The methodology consisted of replacing missing values by the freeze-up date in order to handle multi-year ice situation when no break-up dates are available. We acknowledge the late breakup dates bias from ERA5 is not captured by the break-up analysis (for this area requested by reviewer) and is artificially degrading the results compared to CIS ice charts.

E) The comparison against ice charts is not its own section because the break-up and freeze-up are assessed, using the same method as for the ERA5 data. Hence, it is a subsection of the break-up and freeze-up accuracy assessment. We have also changed the placement of the S2S comparison, putting it inside the results section (section 7).

# Abstract

# Introduction

All corrections have been made.

#### Data

All corrections have been made.

### Forecast model architecture

All corrections have been made.

### **Description of Experiments**

All corrections have been made. We also included a brief explanation of model weights. Thanks a lot for the comments about Figure 3. Very helpful!

### **Monthly Averaged Results**

All corrections have been made.

### Spatial maps of sea ice presence

We removed the part about the Basic model having higher ice presence probability in the northern portion of the domain as this is a minor point.

#### Freeze-up and Break-up Accuracy

All corrections have been made. A new subsection has been made in the section 2 (Data) to describe the ice charts. The accuracy here is different than that shown in the binary accuracy maps because it is calculated by checking if the ice concentration passes a threshold and stays above or below the threshold for a given length of time, as opposed to checking each day. This is now described in section 6.3 (Freeze-up and Break-up Accuracy). We also clarified the wording leading to confusion about a forecast coming from an Ice Atlas.

### S2S system

All corrections have been made. A new subsection has been made in Section 2 (Data) for the S2S data. The description of false negative rate and false positive rate has been clarified to explain how it is computed. We erroneously stated that it was calculated over the entire domain. However, these rates were calculated over the given set of forecast data.

#### Discussion

All corrections have been made

## Conclusions

All corrections have been made. We modified the wording for paragraph lines 444-452, and clarified the wording for the use of a reanalysis. Regarding the question about using high-resolution data, this is something we are currently doing. The Seq2seq model can be trained in that manner, but it is more computationally demanding. Hence there will be trade-offs between resolution, data record length (in years) and domain.

## References

The reference to Vitart et al. was taken from the Nature website and cited using \citep in latex. Other changes to the references have been made. We thank the reviewer again for their attention to detail.