Thank you very much for your detailed and thoughtful comments. Below you will find our responses where the reviewer's comments are in *italics* followed by our responses in blue.

Reviewer #1

This is an interesting paper on the adjustment necessary for the applied use of drift forecasts from short-range forecasting systems. However it is not possible to make a meaningful, statistically significant conclusion on its validity based on an extremely limited sample dataset of 4 drifter buoys, operating for a maximum of 2 weeks during Fall 2018. Given the authors are employed by the forecasting centers producing the model outputs it is possible to do a much more comprehensive analysis with the addition of drifter data from open sources such as International Arctic Buoy Programme (IABP). This will allow further testing to ensure that the results are valid both seasonally, and for varying compactness of the MIZ.

The abstract identifies that knowledge of drift transport in the MIZ is critical for applications including offshore operations and emergency response. It then states that the proposed approach can be used "for operational purposes in the MIZ". There is insufficient evidence presented in this paper to warrant this statement, and there is no attempt to explain or justify this in presenting the results or conclusions. It is also odd that in including this statement, that there is no attempt to verify this applicability through the operational monitoring sections of the Norwegian Meteorological Institute or Environment and Climate Change Canada with their Norwegian or Canadian Ice Services. The abstract attempts to link the approach to operational monitoring, however the term "operational" is used throughout in the limited definition of the research community in meaning only the routine production of data, not the quality assurance and support also included in operational monitoring services.

The recommendation is therefore for major revision, including a more thorough analysis with additional data sources.

We would like to begin by clarifying that the study is more of a proof of concept than a comprehensive test of the drift model. The primary motivation of the study is to see how well two environmental prediction systems can predict the motion of buoys in the marginal ice zone. Current best practices for prediction in the MIZ performed poorly and we found that if we used both the ice and ocean data in the MIZ that the predictions improved. There are not a lot of data available in the marginal ice zone (MIZ), which makes studies such as this important in assessing environmental prediction systems in the MIZ. Programs like the IABP are fantastic, but these buoys are deployed in pack ice and will only drift through the MIZ at times and locations depending on dynamics. Thus their presence in the MIZ is quite sparse. Using such data could be an interesting study, but not necessarily equivalent to buoys deployed in the MIZ. Also, determining whether the IABP buoys are in the MIZ is dependent on ice concentration data products and/or ice-ocean prediction systems, with each having their limitations (temporal and spatial resolution for data products and accuracy from ice-ocean prediction systems).

While we clearly do not have enough data to provide definitive values for the leeway parameters, we argue that we do have enough data to show that current best practices do not perform well in the MIZ for this experiment. We further argue that including a leeway term in the ice - the origin of which could be due to physics not included in the ice-ocean prediction system such as surface waves, or errors in the drag coefficients - reduces the errors with the available data. The scope of the paper is to show that a) best practices for predicting drift in the MIZ do not perform well and b) there are arguments for including a leeway coefficient in the ice. A systematic study to improve the values of the leeway coefficients is beyond this scope.

We should also clarify that we use "operational" as short-term prediction of environmental conditions in order to support operational activities. Examples of this are the knowledge of ocean currents and ice velocities required to support search and rescue or oil spill response. While the quality assurance and support are part of the operational services, these aspects are more part of system upgrades and general improvement and indirectly related to the accuracy of individual predictions. To avoid confusion, we have removed the term "operational" from most of the manuscript and replaced it with "short-term prediction in support of operational activities".

As the transport model is dependent on sea ice concentration (SIC), it is heavily reliant on the accuracy of the source of this data and it's spatial resolution.

Yes, the transport model is dependent on SIC, but this is more a product of the SIC dependence of the coupled prediction systems from which the ice and ocean velocities come from. These three parameters (SIC, ice velocity, ocean velocity) are what we get from CAPS and TOPAZ and we assume they form the basis for the transport model. By using both the ice and ocean velocities in the transport model, the model attempts to compensate for SIC errors leading to inaccurate wind stress partitions in the ice-ocean prediction systems. The wind dependent term, i.e. the leeway, will inevitabily be more senstive to SIC, but current best practices used to predict drift in the MIZ assume no leeway in the ice and we feel we show quite clearly that one needs some leeway in the ice. This could be due to errors in SIC on the scales associated with the buoy motion, but also related to missings physics due to the absence of surface waves in the ice-ocean prediction systems as well as inaccuracies in drag coefficients to name a few. To explicitly show the sensitivity to SIC of our transport model, it is straightforward to rewrite Eq. (3) in the manuscript as

$$\mathbf{u}_o = \mathbf{u}_w + \alpha_w \mathbf{U}_{10} + k_i \left[(\mathbf{u}_i - \mathbf{u}_w) + (\alpha_i - \alpha_w) \mathbf{U}_{10} \right],$$

which shows that the SIC concentration (assuming k_i = SIC) is only important for large differences between the ocean and ice components. Indeed if we set $k_i = 0$, equivalent to the ocean-only model in the manuscript, this performs quite well and much better than ice-only (no leeway), and the 80/30 transport model (except for SIC less than 30% where they are equivalent).

The 80% threshold for assuming ice is or is not in free drift is based on observations, which cover much smaller areas than the typical 100 square kilometers of passive microwave (PMW) SIC products and 12.5 kilometer resolution of TOPAZ.

The 80% threshold is based on the internal ice stress being negligible to the other forces at these ice concentrations for typical wind values and ice parameters. While it may be originally from observations, it is also parameterized into how the ice component of CAPS and TOPAZ calculate the stress. As the ice model has the same SIC input, the model dynamics should be approximately equivalent to a free drift model when the SIC is less than 80%. In CAPS and TOPAZ, the ice strength formulation from Hibler (1979), $P = P_* e^{-C_*(1-A)}$ and a value for $C_* = 20$ so it can be shown that $P(A = 0.8) = 0.02P(A = 1)$. The reviewer quite correctly points out the spatial scales of the SIC data products that contribute to the analysis are much larger, but these are related to the "constrained" scales of the model and the prediction systems will still make calculations on much smaller scales (grid resolution) which are necessary to support operational activities.

P7 Figure 2 and L153: What is shown here is that both CAPS and TOPAZ fail to properly reproduce the MIZ in their SIC values, as a result proposing a drift correction weighted by SIC runs into an issue. This is due to assimilation of PMW SIC into both models that fails to properly represent the MIZ and ice edge, except if it is extremely compact. It would be interesting to see these 2 models compared against the openly available U.S. Naval Research Laboratory GOFS3.1 forecasts, where assimilation of SIC in the MIZ is augmented with use of the Multisensor Analyzed Sea Ice Extent " Northern Hemisphere (MASIE) product. Although CAPS gets its sea ice state from RIOPS/GIOPS, and it also assimilates ECCC Canadian Ice Service ice chart data which would provide better information on MIZ ice conditions, those ice charts only cover the Canadian Arctic area and not north of Svalbard, so the SIC data coming from CAPS in this study also originated in PMW SIC products. A more thorough analysis should be performed also including MIZ in the Canadian sector of the Arctic, e.g. Beaufort and Labrador Seas.

Prediction systems in the Arctic will have errors in the SIC in the MIZ, which will impact the proportion of the wind stress going to the ice and to the water. This is part of the motivation of this study. How can we make best estimates of drift velocity using our prediction systems? We show that using information from both the ocean and ice model improves estimates of drift velocity in the highly coupled region of the MIZ. The primary aim of the paper is to demonstrate a proof of concept for improved drift estimates of the MIZ. A thorough comparison amongst the many products in the Arctic is beyond this scope.

P5 L123 "various ice floes". Given ice type is important for understanding the drag coefficients and differences in drift behaviour, why is the stage of development of these floes and whether there was any deformation (ridging) present not recorded?

The initial state of the ice floe was observed, but the evolution of the ice floes could not be recorded by the instruments. The ice floes at deployment were relatively flat with minimal deformation observed (CHECK WITH JEAN). This has been added to the text.

P5 L144 "The horizontal resolution is about 12 km in the Arctic." No, it is exactly 12.5 km on the Polar Stereographic grid projection used.

This has been clarified in the text.

P6 Figure 1: "Contours of ice concentration". A contour is a line feature, what is shown is shaded discretized sea ice concentration.

Corrected to "filled contour" as contours are calculated and the regions between the contours are filled with the colours corresponding to the colourbar.

P15 Section 5 Conclusions. The proposed general leeway model needs a more comprehensive evaluation to warrant the conclusions here such as P16 L277 "It is clear from the available data that the inclusion of an ice leeway improves short-term predictions in the MIZ".

We believe we state very clearly that from the available data that an ice-only prediction would be very different after 48 hours than a hybrid approach. As always, no research is definitive and more data and ideas will come forth and things evolve. The main point we aim to make is that the current methods for estimating drift velocity in the MIZ in support of operations such as search and rescue and/or oil spill response can give large errors over short time periods in the marginal ice zone. By using a weighted velocity of the model ice and ocean velocities the errors in this case are reduced. The applicability of these results across the Arctic remains to be seen, but for this particular case study they show to hold true.

The text includes a few typographical and stylistic errors:

P2 L23, P2 L35, P2 L50, P3 L70: Replace "arctic" with "Arctic", as capitalization is used when referring to the geographic region.

P2 L27: Repetitive "typically".

P2 L40, P4 L109: Replace "don't" with "do not".

P2 L45: Replace "it's" with "it has"

Thank you. We have corrected the above typographical and stylistic errors as suggested.