Response to Referee 1

In this revised version, the authors have made substantial improvements to the manuscript, and I believe that most of the major issues identified in my previous review have been satisfactorily addressed. Overall, the paper has become a much more complete and scientifically meaningful contribution. I appreciate the authors' considerable efforts in revising the manuscript.

In particular, the inclusion of the coastal polynya dataset from Lin et al. (2024), and the presentation of results both including and excluding polynyas (LF without polynyas; LFWP), significantly clarifies the interpretation of the analysis. I also find it appropriate that the authors removed 2 m air temperature, which appeared to act mainly as an apparent or indirect predictor in the previous version, and instead introduced physically meaningful wind-related predictors. Furthermore, the addition of the new coastal analysis using a 10 × 10 km^2 grid substantially increases the value of the study compared with the original analysis that relied only on the coarse 2° × 5° grid.

We thank the referee for their careful review of the revised manuscript and for their valuable comments. Our responses to each comment are provided point by point below.

However, there remains one important issue that, in my opinion, still requires further revision or at least more careful discussion.

In the newly added coastal analysis, the manuscript now concludes that ocean current speed (OS) is the top driver of lead variability. In the previous version, OS was not identified as the dominant predictor, and therefore concerns regarding the reliability of the ORAS5 ocean current dataset were relatively less critical. In the revised manuscript, however, this issue becomes much more important. As I noted in my previous review, I still encourage caution in interpreting ORAS5-derived coastal current variability as a robust physical driver at the regional scale, particularly in Antarctic shelf and slope regions where observational constraints remain very limited.

ORAS5 is certainly a useful ocean reanalysis product and likely provides a reasonable representation of large-scale circulation. However, it has not been sufficiently validated to confidently represent monthly-to-interannual variability of coastal currents in Antarctic coastal regions. With a horizontal resolution of approximately 0.25°, ORAS5 cannot adequately resolve shelf-break dynamics. This limitation is particularly important for the present study, where the phenomena under investigation are likely influenced by shelf-break dynamics and the Antarctic Slope Current (ASC) in coastal regions.

The manuscript now concludes that ocean current variability is the dominant driver of coastal lead variability. However, the ocean current field itself remains uncertain, and there is a possibility that the ORAS5 current variability partly acts as a proxy for atmospheric forcing.

A dataset such as B-SOSE may be more appropriate for this type of analysis because it has higher resolution, is specifically optimized for the Southern Ocean, and is adjoint constrained. At least for shelf/slope dynamics, B-SOSE likely provides a more realistic representation than ORAS5. Nevertheless, even B-SOSE still faces difficulties in Antarctic coastal shelf regions, and the improvement may not necessarily be dramatic.

One possible approach would be to test the robustness and sensitivity of the coastal results by either removing ocean current (ORAS5) as a predictor or repeating the analysis using B-SOSE.

However, if the authors decide to retain the present analysis framework, then the interpretation of the coastal results should more explicitly and carefully discuss the limitations and uncertainties described above. This remains the main issue that concerns me in the revised manuscript.

We agree with the reviewer that ORAS5 is subject to larger uncertainties than other predictors in our analysis. We note that we used permutation importance rather than impurity-based importance specifically to reduce sensitivity to such issues. Resolving the question of how the dataset's own uncertainty propagates into the importance ranking itself would require a full ensemble-based uncertainty propagation through the random forest. We have added a sentence to the Discussion (Sect. 4) acknowledging this as a limitation and identifying it as a direction for future work.

"A limitation of our coastal analysis can be the use of ORAS5 ocean surface current speed as a predictor in the Antarctic shelf and slope regions. ORAS5 is a global, 0.25° ocean reanalysis designed for basin-scale applications, and its resolution and observational constraints may not be sufficient to fully resolve shelf-break dynamics and the Antarctic Slope Current (ASC). While permutation importance indicates that ORAS5 contributes substantially to predictive skill, the dataset's own uncertainties are not propagated through the importance estimation, and future work using ensemble realizations or alternative independent products would help quantify how sensitive this ranking is to input-data uncertainty."

As a very minor comment: The labels in the newly added Figs. 10 and 11 could be enlarged slightly to improve readability. In Fig. 11, the colors used in the figure and those shown in the legend appear slightly different (the legend colors are darker). It would be better to make them consistent.

We have enlarged the labels to improve readability and corrected the color inconsistency between the figure and the legend.

Response to Referee 2

I would have been interested in having more details on some points raised in the first round of review, like on the connections between sea ice concentration and lead frequency, or on the surprising contribution of temperature but I can live with the fact that the authors preferred to remove these points from their analyses if they were not able to reach strong conclusions. Nevertheless, I still consider that a few modifications, that would not take much time, are required before publication.

We appreciate the referee's continued careful evaluation of the manuscript and would like to thank for their valuable comments. Below, we respond to each comment individually.

1/ I do not see the reason of not including the difference map (Observed - Predicted LF), in Fig. 1 in addition to the existing panels. The authors argue that they 'have not added this as a third panel to Figure 1, because the spatial bias pattern is already addressed quantitatively through the dedicated coastal analysis in Sect. 3.5'. However, there are many differences outside the coastal regions and it is instructive for the reader to see them clearly, as shown on the difference map but not on the current panels.

We agree with the referee that the coastal-box analysis in Sect. 3.5 only characterizes bias at four exemplary locations and does not give the reader a full pan-Antarctic view of where the model over- or under-predicts LF. Rather than adding a third panel to Fig. 1 itself, we have chosen to keep Fig. 1 as a two-panel comparison (observed vs. predicted) for clarity, and instead provide the full difference map as a new supplementary figure, Fig. A2, in the Appendix, which we cross-reference directly from Sect. 3.1. This makes the requested spatial bias pattern fully available to the reader while keeping the main-text figure uncluttered. The pre-existing Appendix Fig. A2 (mean fields of predictor variables) has accordingly been renumbered to Fig. A3, and the one in-text cross-reference to it in Sect. 3.5 has been updated accordingly.

We have also added the following sentence to Sect. 3.1:

"The spatial pattern of model bias is shown in Fig. A2 (Appendix), which shows that the largest under-predictions are concentrated along the Antarctic coastal margins, particularly in the Weddell Sea, Prydz Bay, and Ross Sea coastal zones, while the offshore pack ice shows modest over-prediction."

2/ I do not consider that the proposed modification is explicit enough on the limitations associated with the usage of a coarse resolution (point 4 of the previous review).

We addressed the scale-dependence concern by adding a dedicated coastal analysis at 10 × 10 km^2 resolution (Sects. 3.5 and 3.6), rather than repeating the full RF framework at high resolution across the entire Southern Ocean. This choice was deliberate: leads and coastal polynya occurrence are concentrated along the coastal margins and shelf-break zones. We chose the lower spatial resolution in non-coastal areas because we consider the driver contributions here to be governed by broad-scale processes, and aggregating to that scale is intended to capture the coherent regional signal of these drivers. Increasing the spatial resolution substantially would raise the number of training samples without a comparable increase in independent information, since neighbouring grid cells at high resolution are strongly spatially autocorrelated; this raises the risk of overfitting to local noise rather than better resolving the broad-scale process of interest. Refining the grid in non-coastal areas would therefore not add genuine new information and could introduce artificial precision.

We have revised the following sentence to Discussion (Sect. 4):

" A notable portion of LF variance remains unexplained, reflecting both unresolved processes and limitations that arise from temporal resolution. In consequence, this highlights the need for high-resolution lead-resolving modelling to better understand ocean–ice–atmosphere interactions in bathymetrically-controlled and coastal lead zones."

3/ Why the change of terminology from permutation importance to feature importance? Keeping 'permutation importance' would make the method used more explicit, whereas 'feature importance' could be confused with the impurity-based importance.

We agree that the term permutation importance is more specific and more transparent than the broader term "feature importance". We therefore revised the terminology throughout the manuscript to consistently refer to permutation importance.

4/ Section 3.2: the comments about the order of predictors added is not addressed. Do these conclusions still hold when the sequence of predictors changes?

The stepwise predictor-addition analysis was designed to show the marginal gain of each predictor group within a physically motivated sequence and to identify the minimum predictor set required for skillful LF reconstruction. We agree that the intermediate values in such an incremental approach can depend on the order of addition when predictors are correlated. The purpose of this analysis is not to claim a unique causal ranking from the sequence itself, but rather to illustrate how model skill increases as additional physically relevant predictor groups are introduced. Importantly, the main conclusions remain unchanged: directional wind components and ocean current speed consistently provide the largest contribution to model performance. This analysis is also practically useful for future applications where data availability may limit the choice of predictors.

Minor point/ Caption for figure 11: 'Grey cells indicate no statistical significant at the 0.1 level' word missing

The caption has been corrected to: "*Grey cells indicate no statistically significant relationship at the 0.1 level.*"