General comments

This manuscript is concerned with ice-ocean drag and ocean dissipation that results from moving sea ice keels, a topic worthy of study. I find that the main novel conclusions are quantification of the parameter space that exists in the Arctic. The behavior of lee waves in various parameter regimes appears to be known from previous work (but correct me if I am wrong on this point). This conclusion regarding what parameter space is applicable for the Arctic could be better supported in a number of ways, including providing details of how the NEMO model is validated against the parameters of interest (keel geometry, mixed layer depth, and ice-ocean shear in particular). This conclusion could also be expanded to more robustly consider how that parameter space is related to ice conditions, ocean conditions, or both ice and ocean conditions (see comments below).

It is not particularly clear what new conclusions result from other portions of the analysis. 1) The paper claims to conclude that existing parameterizations are not sufficient and should be re-examined. I do not find this to be a robust conclusion, it is more of a suggestion or possibility to consider. 2) This paper additionally uses 2D numerical simulations, and it is not clear what new conclusions are derived from these simulations. They are useful illustrative examples, and can be retained in the paper, but should be presented as illustrative examples instead of new conclusions.

Overall, I think the novel and most useful aspects of this paper are the quantification of the five parameters for the Arctic environment. I think this paper could be shortened significantly to focus on this narrower set of novel conclusions.

We thank the reviewer for their comments. We added the specific points and clarifications as requested by the reviewer, which are denoted as *italicized* text in our response below.

Furthermore, we have re-done our clustering analysis with four nondimensional variables and numbered the clusters in the order of decreasing area that they occupy. Therefore, the cluster numbers are different from the original draft. We ask the reviewer to reference the newly revised figures at the end of our response.

Specific comments

It is difficult to follow this paper because of its use of variables, e.g, chi, J, eta, instead of their more dynamical names / descriptions, along with the use of cluster numbers, instead of more physical descriptions of these regions. For example, cluster 3 could be referred to as Cluster 3 (central arctic) or Cluster 3 (perennial sea ice zone) as the authors see fit. Lines 125-130: parameters like chi and J that have names should have their names stated here as well as in the abstract. In general, use of these names, e.g., internal wave nonlinearity, in addition to their symbols throughout the text would make this paper easier to read and understand.

While we appreciate the reviewer’s comment regarding the variable names, stating the variable names every single time would be unnecessarily confusing (in terms of sentence structure). These nondimensional variables are fairly common in modelling studies (for both sea ice and bottom topography studies), so we would like to keep them in text as they are. However, acknowledging that this might be confusing to readers who are unfamiliar with such studies we will:

- add a table with the nondimensional variable symbol and its physical representation, and
- add the variable names (following the abstract) to the Methods section where the variables are defined in terms of the dimensional parameters.

We appreciate the suggestion of naming the clusters, but we now present all of the clusters based on the annual, summer, and winter averaged data, yielding 18 clusters. Therefore, naming them would not be tenable and will become rather more confusing. The clusters are now all clearly plotted separately in the Revised Figure 4.

Lines 71-80 are not necessary. These are not examples of GMM relevant to this paper.

We would like to keep these as examples of previous studies that applied GMM clustering methodology to oceanographic datasets. However, we shortened this paragraph acknowledging that some of the details are superfluous. As suggested by Reviewer 1, we instead added some more details about the GMM algorithm in general.

> *"GMM is an unsupervised clustering method that attempts to represent the data as a linear combination of K M -dimensional Gaussian distributions (Reynolds et al., 2009). K is the number of clusters that needs to be specified, and M is the number of input variables. Unlike other clustering algorithms, such as k-means, that are deterministic, GMM is a probabilistic approach to clustering. For each data point, it assigns the probability of belonging to one of the K distributions; hence, one can use these probabilities to assess the model performance. GMM has previously been successfully used in oceanographic problems involving complex, spatially variable processes. For example, Jones and Ito (2019) employed GMM to classify global regions based on ocean carbon budget terms, and Ye and Zhou (2025) applied GMM to global ocean temperature profiles."*

Line 96-97: More details of the model validation and whether it is realistic or not need to be given in this paper. This study still relies on a model representation of keel characteristics (height and steepness). How do we know this is valid? Line 317-323: does the model accurately represent observed ocean stratification for the Arctic? Many arctic models fail to do so. Based on the information given in this paper, the claim that this model is a representative sample is not backed up. Figure 2: Can any of these parameters be compared to observational estimates? Is the geographic distribution of these parameters accurate?

We acknowledge that the input data in this study is taken from another model, which, as the reviewer has pointed out, may not be an accurate representation of the relevant Arctic dimensional parameters. However, the main objective of this study is to identify an approximate range of nondimensional parameters for the Arctic to use in future high-resolution numerical simulations. These high-resolution numerical simulations would then be used to determine how accurate the existing parameterizations (representation of unresolved climate model processes) are. One of the main reasons for this is that current high-resolution numerical studies choose the ranges of parameters somewhat haphazardly, and it is unclear what portion of the range of parameters (or dynamical regimes) has been studied.

To this end, we would argue that it is sufficient to examine the representation of ice keel characteristics from a commonly used regional/climate model that has a coupled sea-ice model. Ultimately, any new parameterizations would need to be implemented into such a climate model, so we need to see what ranges of sea ice characteristics the climate model sees (rather than what are true sea ice characteristics). So, we would want to know what high-resolution numerical simulations will need to be run to ensure that the parameterizations in climate models accurately capture the physics. These need to be based on the values of ice keel characteristics that the climate model sees, not necessarily based on the actual observed values. Therefore, this study is done in service of future studies to narrow down the parameter ranges of the nondimensional parameters to vary in future high-resolution simulations rather than pinpointing the precise distribution of sea-ice characteristics. Of course, it would be significantly better if this kind of analysis could be performed on observational data. However, there are no datasets with a full spatial coverage for all of the required variables that go into calculating the nondimensional parameters of interest.

The main point is how accurate is the actual internal wave drag (by considering small-scale physics in high-resolution numerical simulations) in comparison to the predicted drag (i.e., Eqn. 21 in the paper) *based on the parameter values that the climate model sees*. If we can improve this parameterization (the relationship between the nondimensional parameters and the drag), then we can make climate models more accurate and hopefully reflect the observations better. We will make this point clearer in the introduction and the abstract.
With all of this mind, it is important to add some details about the previous studies that performed model validation for this particular model, as suggested by the reviewer.
We have add this to the limitation portion in the Discussion section:

> *"During the summer, the canonical mixed layer can be absent in in parts of the Arctic, such that even the near-surface ocean layers are stratified (Randelhoff et al., 2017). For the idealized numerical simulations conducted to assess the parameterization of CIW , the absence of the mixed layer does not pose a problem, as it would be just the limiting case of setting the nondimensional parameter η to zero (as the mixed layer depth z0 = 0). However, as noted in Flocco et al. (2024), the pycnocline or mixed layer depths shallower than 10 m cannot be accurately quantified in the NEMO model, hence, they are not present in our clustering results. This means that in addition to the ranges of values of η presented in this study, smaller values of $\eta \rightarrow 0$ would have to be considered when conducting numerical simulations to evaluate the parameterization IW drag induced the ice keels to capture the full range of parameter variability. Based on our preliminary numerical simulations, we would expect that smaller η would enhance the IW generation and the IW drag as the ice keel would be in a more direct contact with the stratified layer, potentially without the buffer of the mixed layer. Therefore, omitting these smaller η values would lead to an underestimate of the effects of the ice keels on the ocean IWs. This could be one of the reasons that we find relatively small CIW values during the summer months (Fig. 11(b,e)).*
>
> *Other differences in the values of the underlying dimensional parameters between the model output the real ocean can also affect the implications of our findings. For example, Flocco et al. (2024) showed that their model in general overestimates the sea ice drift*

speed over most of the Arctic and across seasons in comparison to the observational data from the National Snow and Ice Data Center Polar Pathfinder dataset. This is consistent with climate models typically overestimating sea ice drift (Wang et al., 2023). From Flocco et al. (2024), the overestimation of the sea ice drift by the model is largest during the winter in the marginal sea ice areas. These regions are part of our clusters W2 and W3 (Fig. 4(i, l)) that have relatively large CIW predicted by the parameterization (Fig. 11(c,f)). For cluster W3, this is in part because of the larger Froude number Fr (Fig. 6(f)) that is proportional to the relative sea ice velocity u0. Therefore, an overestimate of the ice drift by the model could, in turn, overestimate the IW drag CIW.

We also assess the distribution of ice keel depths h0 with observed values. SI Figure 2 shows the distribution of h0 from the overall Flocco et al. (2024) model output (not just our filtered data within the lee wave radiation regime) with the distribution from a large keel dataset by Metzger et al. (2021). The distributions are in overall good agreement, especially if we consider h0 < 9 m as one bin, considering that data for h0 < 6 m was not presented in Metzger et al. (2021). However, other observational studies (e.g., Cole et al., 2017; Brenner et al., 2021) have found such smaller keel depths (h0 < 6 m). These comparisons suggest that the distribution of the h0 values used in this study is in a reasonable agreement with the observations."

Figure 2, line 230, and elsewhere: Three of these five parameters rely on u0. How is u0 estimated from the model? Is ocean velocity at a specific depth used, or is it averaged over a certain depth range? What drag coefficient formation is used in the model from which u0 derives? Is the drag coefficient constant or variable? Do the details of u0 varying with depth in the turbulent boundary layer and Ekman layer matter to this problem? How has u0 been validated? In general, I am also confused about when u0 is allowed to vary, and when a constant value of u0=0.1 m/s is used (2D simulations I think). Variability in u0 is potentially very important to this problem.

We thank the reviewer for pointing out some of the details that we omitted from the manuscript and will include in the revised one. Mainly:

- u0 is calculated as speed of the ocean current relative to the sea ice directly below the pycnocline, and
- we use the outputs from the reference run from Flocco et al (2024), that is without parameterized IW drag included in the model simulations.

To address the reviewer's other questions:

- It does not matter (for the definition of the nondimensional parameters and the derivation of the parameterization relationship) where, i.e., at which depth, u0 is measured, as long as it is defined and later applied consistently in the parameterizations.
- The specific accuracy of u0 is also not important as we are interested in the range of parameter values, not specific values at specific locations – see the detailed response to the previous comment.
- Yes, u0 does spatio-temporally vary. However, the variability of u0 is captured in the variability the nondimensional parameters. Since we have 6

dimensional parameters and 4 truly independent nondimensional parameters, in order to vary the nondimensional parameters, we need to fix two of the dimensional parameters (we choose u0 and k0). That is the key advantage of collapsing dimensional parameters into fewer nondimensional ones.

-

Figure 2: How often do the parameters satisfy the lee wave radiation condition?

In the annually-averaged data, we excluded about 32% of data points. For the summer and winter months, we excluded 17% and 38%, respectively. We now added this to the text.

Line 153-155: Is there seasonality in any of these parameters? Are these parameters only within the lee wave radiation range during winter? These are important details to convey. I realize that seasonality is addressed later in the paper, but it is important in this context.

Although it was perhaps unclear, our GMM clustering in the paper was done separately for annually-averaged and seasonally-averaged (over summer months and winter months, separately) data. This is now made more clear in text and we added summer and winter panels to all of our GMM analysis figures (Revised Figs. 2, 4, 5, 6, 9, 11). Specifically, Revised Figure 2 addresses this comment by the reviewer. We indeed find seasonal differences in the distribution of the nondimensional parameters. Please note (in addition to the previous response) that we find data points within lee wave radiation regime both during the winter- and the summer-averaged data (and in the annual average): Revised Fig. 2 (a-c). As the numerical simulations are for illustrative purposes only, we kept them to be only for the annually-averaged clusters.

Line 161-162: I'm confused about time averaging in the GMM setup. This sentence seems to state that data are NOT averaged ('at each timestep'), but also that time averages are used ('after fitlering and time-averaging').

We appreciate the confusion that this sentence introduced. GMM clustering is performed on the filtered and time-averaged data, so only on spatially-distributed (in lat-lon space) nondimensionally variables. However, we perform three types of time-averaging: (a) annual average, so over all the data points after filtering, (b) over the summer months, and (c) over the winter months. For each of the time averaging interval, we compute separate time-averaged values for the nondimensional parameters, and then use them to perform the GMM clustering. So, the clusters for each of the three time-averaging periods are "independent" from each other in a sense that they do not know about the values of other time averages (of course, they are not truly independent because parts of the ocean might behave similarly despite the seasons).

How sensitive are the main conclusions of this paper to some of the choices and assumptions made? For example:

Figure 3: Do the conclusions of this paper change if only 5 clusters are used?

With regard to the number of clusters, when applying GMM, having too few clusters would typically lead to more muddled results. This is why when performing GMM clustering, one produces the BIC score curve (Fig. 3) that balances having too many clusters (overfitting) and having enough clusters so that certain characteristics are captured. Based on our BIC score, we would argue that even 6 clusters is too few from the statistical point of view, but we chose this number to balance the fit of the GMM clustering and interpretability (i.e., not having too many clusters). In practice, once the GMM or other clustering is performed, clusters are often combined in the discussion (e.g., Sonnewald et al, 2019) as we have done here. So, the clustering with many clusters is the first step in letting the machine learning identify statistical patterns that are then interpreted by a researcher into a more coherent narrative. Please refer to the text below and Revised Fig. 5 and Supplementary Fig. 1.

> *"In this study, we made a particular choice of six GMM clusters based on the statistical information from the BIC score and by considering the interpretability of our results. For the purpose of the discussion here, we focus on the annually-averaged data, though similar conclusions can be made for winter and summer clusters. Based the BIC curve (Fig. 3), we should have chosen a much larger number of clusters ($K \approx 20$ clusters) in order to improve the GMM's ability to capture the variability in the data. Having more clusters would have allowed us have better-constrained clusters, i.e., reduce the standard deviation of the nondimensional variable ranges within each cluster, especially for Cluster A5, and less overlap between clusters (Fig. 9). However, this would have been too many clusters to interpret in terms of physical regimes. On the opposite end of the spectrum, we can also divide the Arctic broadly into three regimes: (1) central Arctic, i.e., Clusters A0 − 1, (2) marginal ice regions with larger η (smaller keel depth, deeper pycnocline) and without substantial IW generation, i.e., Clusters A3 − 5, and (3) marginal ice regions with intermediate η that support IW generation, i.e., Cluster A2. However, such broad characterization would return a wide range of nondimensional parameter values for the central Arctic, which perhaps an insightful result. In order to examine the tradeoff between the accuracy of representing the statistical distribution of the nondimensional parameters (i.e., large K) and ease of interpretation (i.e., small K), we show the spatial distribution of GMM clusters for $K = 4$, 5, and 7 in SI Fig. 1 for comparison with $K = 6$ in Fig. 5(a). Too few clusters (e.g., $K = 4$) leaves the entire central Arctic region as a single cluster. However, for $K = 5 - 7$, the GMM algorithm returns a relatively consistently clustered centered Arctic, separating the Amerasian and Eurasian basins, and the outer eastern Eurasian seas. Increasing the number of clusters K in this range (5 − 7) seems to predominantly break marginal regions into even smaller clusters. Based on this analysis, we choose $K = 6$, though realizing that this choice is rather subjective."*

Line 197: A pycnocline width of 0.5 m is likely fine. But do the results depend on this parameter? A comment might help.

This width is chosen to have a thin but finite-width (for the stability of numerical simulations) pycnocline. In practice, the pycnocline width could be another dimensional parameter to consider (so there would be another nondimensional parameter). However, this would complicate the number of high-resolution numerical simulations to be conducted in the parameter sweep (having to vary another parameter). Therefore, since the pycnocline width does not appear in the IW drag parameterization, we choose to keep it constant. Changing its value would, again, affect the numerical simulations, and it might be interesting in the future to consider real stratification profiles from the Arctic to get better estimates of this value (though I am not sure if measurements would have sufficiently high vertical resolution). To reflect this discussion, we will add a sentence to the section where pycnocline width is defined.

Line 211: If u0 were a factor of 2 larger or smaller, would this affect the conclusions?

Again, we are initializing the numerical simulations by choosing the nondimensional parameter values, not the dimensional parameter values. Based on the choice of the nondimensional parameter values and the values of the two fixed dimensional parameters (u0 and k0 in our case), the other dimensional parameters are set. If u0 was chosen to be larger or smaller, the other dimensional parameters would be adjusted accordingly based on the choice of the nondimensional parameter values to be within the same dynamical regime.

Figure 8: The winter surface layer depths in the Arctic shown here are approximately 50 m depth on average, and deeper for cluster 5. There can be a range of surface layer depths across the Arctic, with wintertime values of 40 m or even shallower. How much difference would a 10 m change in MLD make to diagnosed mixing, or any of the other conclusions?

That is precisely the objective of a follow-up study (currently underway), where we will vary the relevant parameters (based on the ranges of values that we identified in this paper plus some additional runs from observational measurements that do not overlap with these ranges) and diagnose mixing, dissipation, and wave drag.

I would interpret Figure 4-5 as having Cluster 4 representative of marginal ice zone / seasonal ice zone conditions. Is it the ice keel dimensions and not the ocean conditions that distinguish this cluster from 3 and 5? This is more relevant than listing the names of the seas that clusters occupy (line 267).

Our analysis in terms of the nondimensional numbers that combine both the ocean and ice characteristics does necessarily let us distinguish this. However, looking at Figure 4 in Flocco et al (2024), the area covered by (the former) Cluster 4 is characterized by relatively deeper mixed layer depth, relatively deeper keel depth, and relatively weaker density jump across the pycnocline. So, we would not say that it is more strongly driven by sea-ice dimensions as opposed to the ocean conditions. However, it does appear to be a transitional (perhaps marginal) zone.

This paper would benefit from some context about ocean mixing in the mixed layer, pycnocline region, and interior. Is diagnosed mixing weak, average, or strong? Are lee wave processes a dominant source of mixing? The diagnosed dissipation rate ranges from order 10^-15 to 10^-7, which is a wide range of values. The largest values for the ocean interior are 10^-10, which I believe is small to average compared with other studies (e.g., Fine and Cole, 2022, Decadal observations of internal wave energy, shear, and mixing in the Western Arctic Ocean, J. Geophys. Res., https://doi.org/10.1029/2021JC018056).

We now more included a more thorough comparison of the observed KE dissipation rates in the Arctic with the rates calculated from our numerical simulations:

> *"Many observational studies estimate internal wave dissipation rates in the Arctic to be within the 10−10 − 10^{-9} m^2 s^{-3} range (Scheifele et al., 2018; Kawaguchi et al., 2019; Fine and Cole, 2022; Fer et al., 2022). These values are smaller than observational measurements of up to 10^{-8} m^2 s^{-3} in the upper parts in other global ocean regions (Waterhouse et al., 2014). In our idealized numerical simulations, we find the dissipation rates below the pycnocline to be generally on the order of 10^{-9} m^2 s^{-3}, which is within the range of observed values (Fig.12(b)). However, larger values of ϵ have been found close to the sea ice, in particular when the mixed layer is thin (Fer et al., 2022; Reifenberg et al., 2025). We also find larger values of kinetic energy dissipation within the pycnocline region and above the pycnocline in the case of numerical simulations with smaller η (shallower pycnocline) (Fig. 12(d)). These results suggest that depending on the sea ice and flow characteristics, there could be a spatio-temporal variability in the dissipation and subsequently diapycnal mixing rates in the Arctic. Of course, there are many real-ocean processes that are missing in our idealized numerical simulations, e.g., propagation of internal waves originated elsewhere and near-inertial waves that can also be generated by the wind stress at the ocean surface. Therefore, we caution against overinterpreting the numerical values of ϵ in our numerical simulations."*

Also in terms of context, I find it hard to visualize which regions have elevated mixing due to lee wave generation from the tables and discussion of the figures. Consider making a map showing pycnocline or interior mixing across the Arctic based on the clusters (obviously involving some assumptions). Line 446-455: The regurgitation in terms of relationships between different parameters is one view, but what is missing is the practical aspects. Overall, is lee wave drag or mixing significant for the Arctic? Where and when is it significant?

In order to make such maps, we would need to conduct a parameter sweep over the ranges of values of the nondimensional parameters with high-resolution numerical simulations, as proposed for a follow-up study. Otherwise, there are no ways to make good predictions for the mixing values generalized from this study, as no appropriate parameterizations for the mixing values exist. We, in fact, use this study to point to the need for such a follow-up study based on the parameter ranges and the numerical simulation set-up that we discuss here.
For the predicted lee wave drag, Flocco et al (2024) showed the values based on McPhee and Kanth (1989) parameterization. We invite the reviewer to see their figures 4 (b,c), 5 (b,c), and 6 (b,c) and we do not reproduce these figures as they are already published. However, the central premise here is that the existing parameterizations may not be correct because they are based on assumptions to derive them analytically (or do not exist as in the case for mixing), so they need to be investigated further, as we plan to do in the follow-up study.
We also have done a more thorough comparison of the predicted Ciw drag parameterization values with form and skin drag coefficients from observations. Please see our response to the reviewer's other comment below.

Line 422-429: These lines relate to one of the main conclusions of the paper, that the existing parameterizations are not quite correct. These lines do not give sufficient detail to support that conclusion. I do not think that the factor of 5 difference in internal wave drag at cluster 2 is particularly robust given the overlap in the 25-75th percentiles, and the general spread in the distributions of several orders of magnitude. The simulations that these are compared to also do not account for any variability in parameters, and are constructed with mean values. There might be a suggestion that the parameterization could be further investigated, but it is presented elsewhere as more of a conclusion.

We believe that this comment stems from some misunderstanding (and perhaps lack of clarity on our part) about our argument. Our point here is that comparing across clusters, there is a discrepancy between the relative values of C_IW predicted from parameterization (Eqn. 21) and the relative values of KE dissipation that we measure in the numerical simulations. That is, from the crosses in Fig. 11, the existing parameterization predicts 5-10 times greater internal wave drag for Cluster 2 mean parameters (the ones that we use in the numerical simulations) in comparison to the internal wave drag for Clusters 3-5 mean parameters. However, based on the results from the numerical simulations, Cluster 2 has 1-2 orders of magnitude smaller KE and KE dissipation rate compared to Clusters 3-5. If the IW drag is larger, one would expect larger KE and KE dissipation rates, but we find that predicted larger IW drag values (from existing parameterization) are associated with smaller KE/KE dissipation rates (from our numerical simulations). Hence, we suggest that the existing parameterization needs to be revisited. We will amend our writing to reflect the explanation above and soften the language regarding revising the parameterizations. To that end, there are other reasons (e.g., parameterizations only consider 2D flows, whereas 3D turbulence is most likely important) that point to the need to re-examine this parameterization.

We also now include a more illustrative figure (please see our Revised Figure 12) to demonstrate the explanation above to be clearer.

We also now added a statement as to why the current IW parameterization might not be correct, in particular in the regions with large value of J.

> *"It is important to note that the IW drag parameterization by McPhee and Kantha (1989) was developed for a two-dimensional model assuming small keel height h0, i.e., small topographic criticality parameter J for a fixed stratification N0 and relative keel speed u0. Recent study that assessed the parameterizations of form drag for flow over seamounts (Johnston et al., 2025) found that the disagreement between numerical simulations and the two-dimensional parameterization also derived for small mount heights (Bell Jr., 1975) increased as J increased. Specifically, they found that the parameterization underestimated the form drag more in comparison to the numerical simulations for larger values of J: at J = 1, the parameterized drag was only about one third of the value estimated from the simulations. This result suggests that the current parameterization for CIW by McPhee and Kantha (1989) might also be underestimating the drag at larger values of J. In this study, we find many points in across the Arctic with $J \gtrsim 1$: 67% of all grid points in the annual average, 91% during the summer months, and 7% during the winter months. Our findings indicated that this supercritical regime $J \gtrsim 1$*

could be an important parameter regime, especially during the summer, that is not well-represented by the current parameterization and needs to be re-evaluated through future numerical studies."

Line 422-429: Again, context is needed. Can these values of the internal wave drag coefficient be compared with previous studies? Typical ice-ocean drag coefficients are order 10^-3, which is not mentioned in this paper and should be. Compared to that value, in all but the most extreme examples, the drag from lee waves is negligible.

We now have redone our analysis to compare the parameterized (predicted) Ciw values with values of form and skin drag coefficients from observations. Please see our Revised Figure 11 and the statement below.

> *"We can estimate the relative importance of the IW drag by comparing the CIW values from parameterizations with skin Cs and ice-ocean drag coefficients Cio estimated from observations (Fig.11(a-c)). We take the skin drag coefficient value of Cs = 7×10^{-4} from the measurements under unridged summer ice by Reifenberg et al. (2025). We show the range of values for Cio estimated from various observational studies: $1.3 - 12.3 \times 10^{-3}$ (Beaufort Sea, annual cycle by Brenner et al., 2021), 4−6×10−3 (Amudsen and Nansen Basins, summer by Fer et al., 2022), $1-10\times10^{-3}$ (Canada Basin, annual cycle by Cole et al., 2017), and $1 - 10 \times 10^{-3}$ (average value of 3.4×10^{-3}, Nansen Basin in July by Randelhoff et al., 2014). However, notably, Kawaguchi et al. (2024) measured that Cio can be as large as 0.13 at times. We also make comparisons to the canonical ice-ocean drag coefficient value of CD = 5.5×10^{-3}, which is approximately in the middle the Cio range (Fig.11(d-f)). We find that the CIW values are typically smaller than the measured Cs and Cio values. In the regions of perennial sea ice (e.g., cluster A0, A1, S0, S2, W0), CIW is much smaller in comparison to the form and skin drag coefficients; we find it to be negligible in comparison to CD (Fig. 11(d-f)). However, in some marginal ice zones (e.g., A2, S4, W2), CIW can be 10 − 20% of CD . The values of CIW are in general larger in the winter. In particular, the CIW values for cluster W3 (along Greenland and in the Chukchi Sea) can be as large as or exceeding CD . So, even though the IW drag may be relatively not as important in the pack ice regions in the center Arctic, it could be important in the marginal zones, especially in the winter."*

Line 440: Are these sea ice regimes, ocean regimes, or sea-ice and ocean regimes? I believe it is the latter. The results section seems to focus more on their interpretation as ocean regimes more than ice regimes. Additional interpretation in terms of perennial ice cover (cluster 3) vs. seasonal ice cover (cluster 4-5) is also needed. Can arguments be made about ice regimes? What does this imply about a future arctic environment with increasing seasonal sea ice?

Yes, it is, of course, the ice-ocean coupled regimes. We use the term "flow characteristics" rather loosely here, as we are interested in the flow processes (i.e., internal waves, turbulence), but the dynamics are influenced by both the characteristics of the sea ice and the underlying upper ocean. This is also reflected in our choice of dimensional parameters (both pertaining to the ocean and sea ice) to characterize this problem. The term "sea-ice regime" is also used rather loosely here as it is the sea ice movement that is driving the generation of IWs in the ocean interior. To avoid any confusion, we will re-write this portion to refer to the whole coupled system in terms of the “dynamical regimes”.

We also thank the reviewer for the suggestion about the perennial vs. seasonal ice cover regimes. We now added the following to our discussion:

> *“Our estimates of KE dissipation rates from the idealized numerical simulations and the estimates of CIW from the current parameterization both indicate a differences between the regions of perennial sea ice (Cluster A0) and seasonal sea ice (Clusters A1 and A2). In particular, we find that these seasonal sea ice regions are predicted to have larger CIW (Fig. 11(a)), but possibly smaller KE dissipation rates (Fig. 12). This can be important as the proportion of perennial sea ice in the Arctic has been decreasing in the past three decades (Serreze and Meier, 2019).”*

Lines 446-455: Please be explicit about what new results are revealed by these 2D simulations. It is my understanding that none of the behavior in different parameter-space regions is novel, it is simply that the quantification of those parameter-spaces for the Arctic has not been previously investigated. The 2D simulations should be presented more explicitly as an illustrative example instead of as new conclusions. I will also make a second point that I do not think conclusions about mixing or kinetic energy values from these 2D simulations are particularly robust. Not enough of the variability in parameter-space has been included to draw robust conclusions. I view these 2D simulation primarily as useful illustrative examples.

We agree with the reviewer that the 2D simulations are used as illustrative examples. While that was our intention, we now emphasize it more clearly in the paper.

However, one major point is that this numerical set-up for sea ice modelling in Oceananigans is novel because Oceananigans significantly increases the speed of numerical simulations through its GPU architecture. Therefore, using our set-up, it is possible to conduct a parameter sweep over a wide range of parameter values in a more time-efficient manner in comparison to the previous CPU-based numerical solvers.

Technical comments

Abstract: Too many variables for an abstract. Use of the more common Ri might be okay, but I suggest removing the symbols for others.
The abstract is now restructured such that we talk about the variables in terms of their physical meaning rather than symbols.

Lines 2-4: Run on sentence with two uses of ‘because’
Thank you –fixed.

Line 57: Be more specific than ‘keel size’. Is this horizontal extent? Height?
We mean horizontal extent – it is now specified.

Line 105: Horizontal spacing BETWEEN KEELS, L.

Now clarified.

Line 227: Equation 13 has w without a prime, u with a prime. Which is correct?
The prime represents that the background mean flow has been subtracted. The mean flow is only imposed in x-direction, hence u' represents the deviation from the mean flow. For w, there is no mean flow, so w (without the prime), represents the total deviation. This is a common notation in describing perturbations about mean state.

Line 287: 'small KEEP depth'
Changed.

Figure 2d: log10=0 corresponds to eta=1, which is an important value for this parameter. The colorbar should be changed to emphasize this range of values, and the variations between eta=10 and eta=100 are less important.
This has now been adjusted – please see the Revised Figure 2.

Figure 5: cluster 0 and cluster 5 are hard to distinguish, especially when printed out.
We now adjusted cluster colors to be more distinctive – please see the Revised Figure 5.

Tables 1-3: These are hard to understand as just numbers on a table. I suggest plotting the values as a graph where shading or symbols can also be used to show standard deviations. Table 2 may need to be a logarithmic scale, but I think it is still useful to consider presenting it this way. Figure 6 is much better than Table 1! I think table 1 could be removed, and mean values could be added to Figure 6.
We now have the values from Table 2 plotted in the new Figure 12. We also believe that having Table 1 in addition to Figure 6 is important to give the exact values (not just ones that are inferred from plots such as Fig. 6), especially for the reproducibility of numerical simulations in the future. Therefore, we chose to retain Table 1.

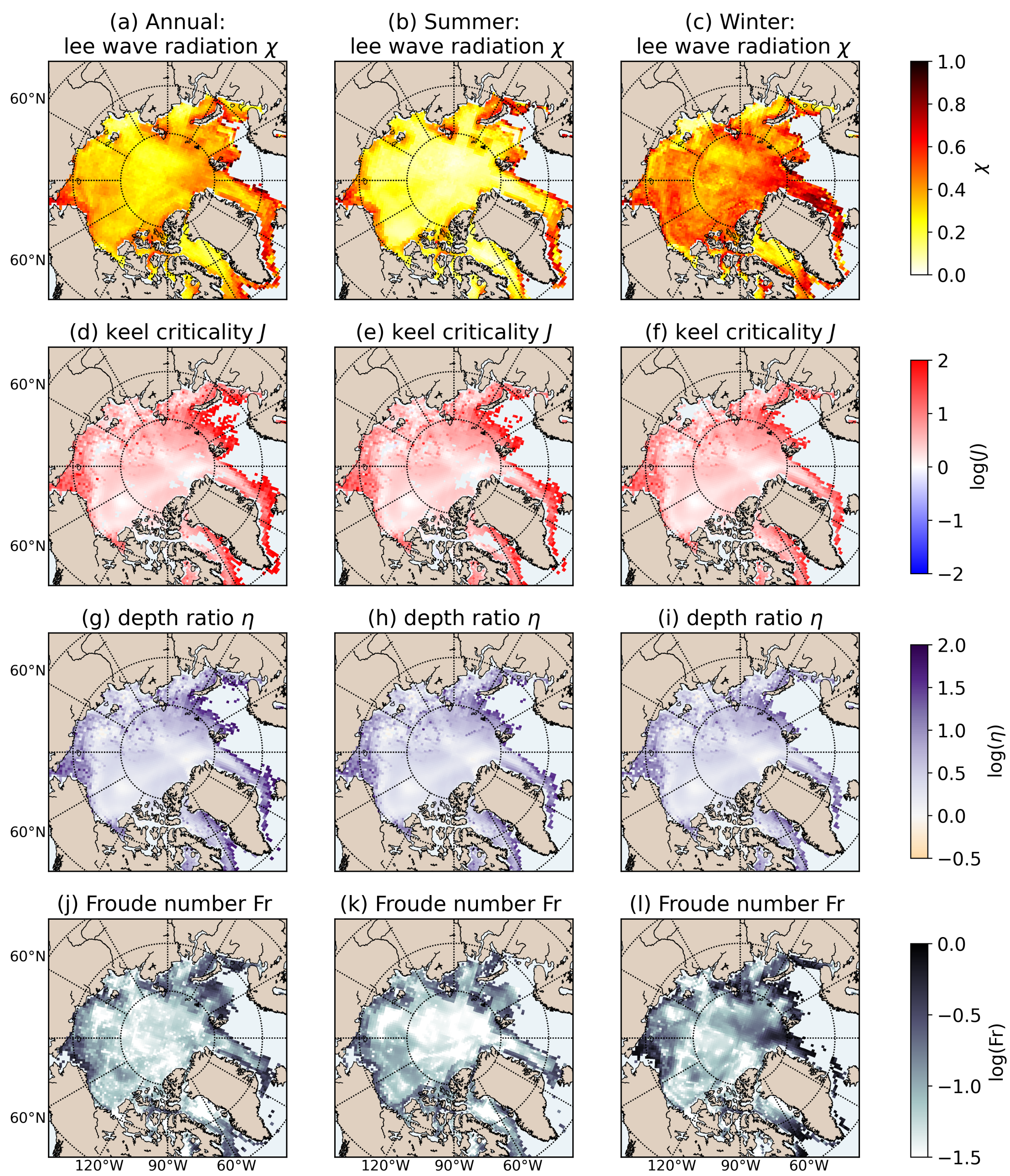
Tables 2 and 3 should make it clear in their description that these results derive from the idealized 2D model.
Tables 2 and 3 already state that these are derived from numerical simulations.

Figure 9: Text is too small. The color-scheme should be improved for readability, especially for summer. Panel (o) is also difficult to determine which clusters are which. Different seasons do not need to have different color schemes.

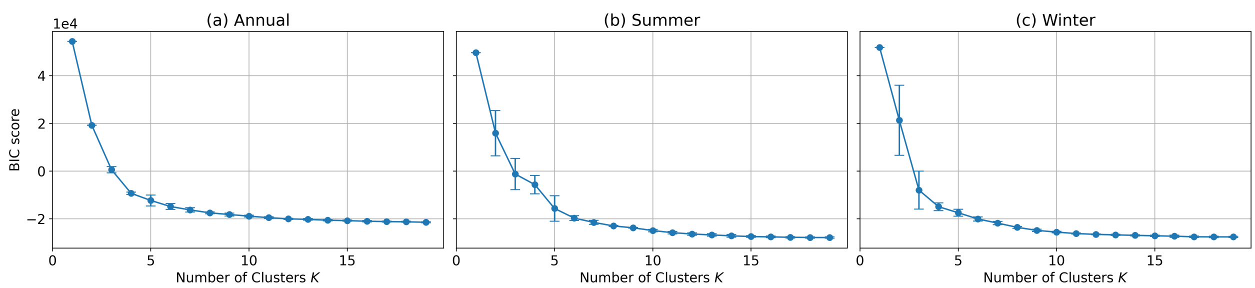
- We have now adjusted the text size – please see our Revised Fig. 9.
- The overlapping of the clusters for certain variable combinations is inevitable. The main point of Figure 9 is to showcase the pairwise range of values for the nondimensional parameters and to demonstrate the parameter regimes that are important and not important to the Arctic. Instead, Figure 6 (boxplots) clearly shows the ranges of values for each cluster separately for comparison across the clusters.
- We choose to retain the different color schemes for different seasons (annual, summer, winter) because those clusters are derived from conducting the GMM clustering on the annual-, summer-, and winter-average data separately. Therefore, representing them with the same color might be confusing as it would

imply that cluster 1 is the same region for annual, summer, and winter. However, that is in fact not the case – please refer to our Revised Figure 4 to see that these are spatially different clusters.

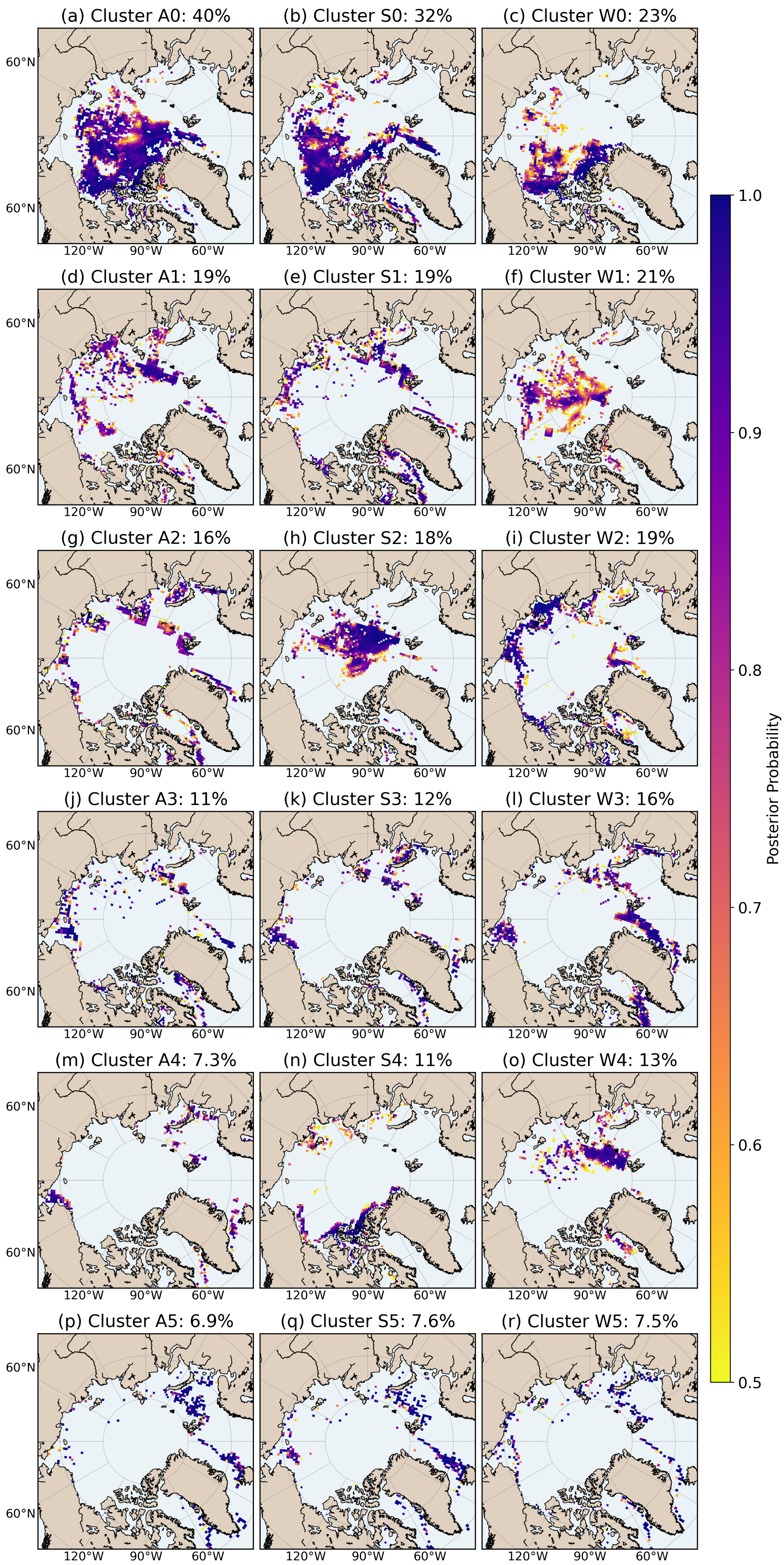
Figure 9 and 10: please adjust one figure so that the corresponding panels share the same axes. Example: Figure 9d uses a linear axis in eta and Figure 10d uses a logarithmic axes. Also consider adding dots for the mean values in each cluster and season to Figure 10.
We have now combined Figures 9 and 10 such that Figure 10 is now a fourth column. Please see our Revised Figure 9. All of the axes on the corresponding panels for each row are now the same. We tried adding dots of the mean values to the fourth column (Ciw), but the plots became too busy. Therefore, we kept the revised configuration in hopes that one can compare across each row to interpret the relevant parameter spaces.



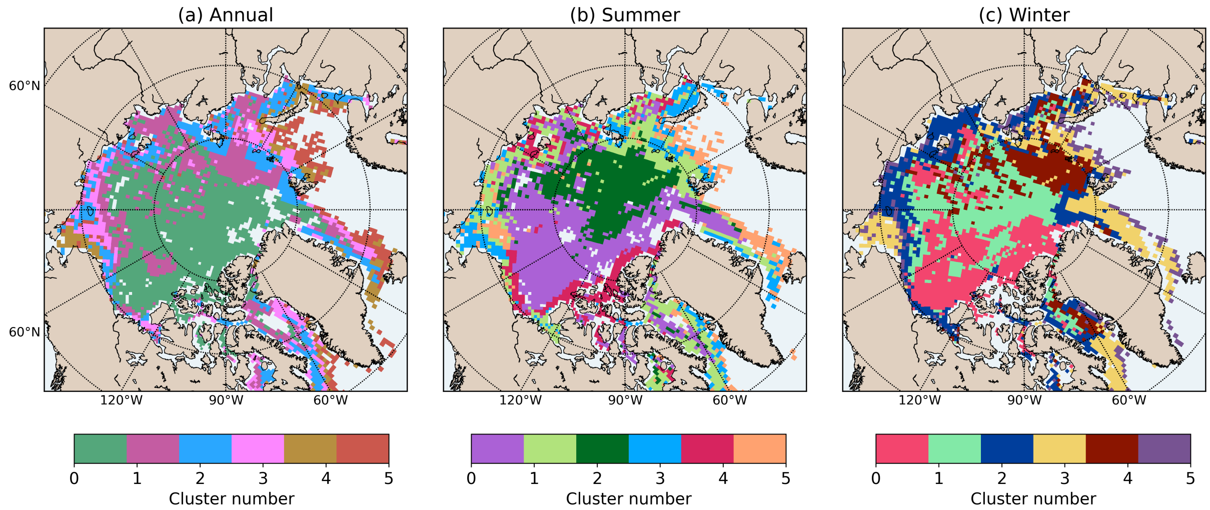
Revised Figure 2: Spatial distribution of time-averaged five nondimensional variables over the Arctic Ocean: (a-c) lee wave radiation parameter χ, (d-f) keel criticality parameter J, (g-i) depth ratio η, and (j-l) Froude number Fr. The values are based on the values averaged over different time intervals: (left) annually, (middle) over the summer months, and (right) over the winter months. Note that colorbars vary across subplots and the magnitudes of (d-f) J, (g-i) η and (j-l) Fr are shown on a logarithmic scale.



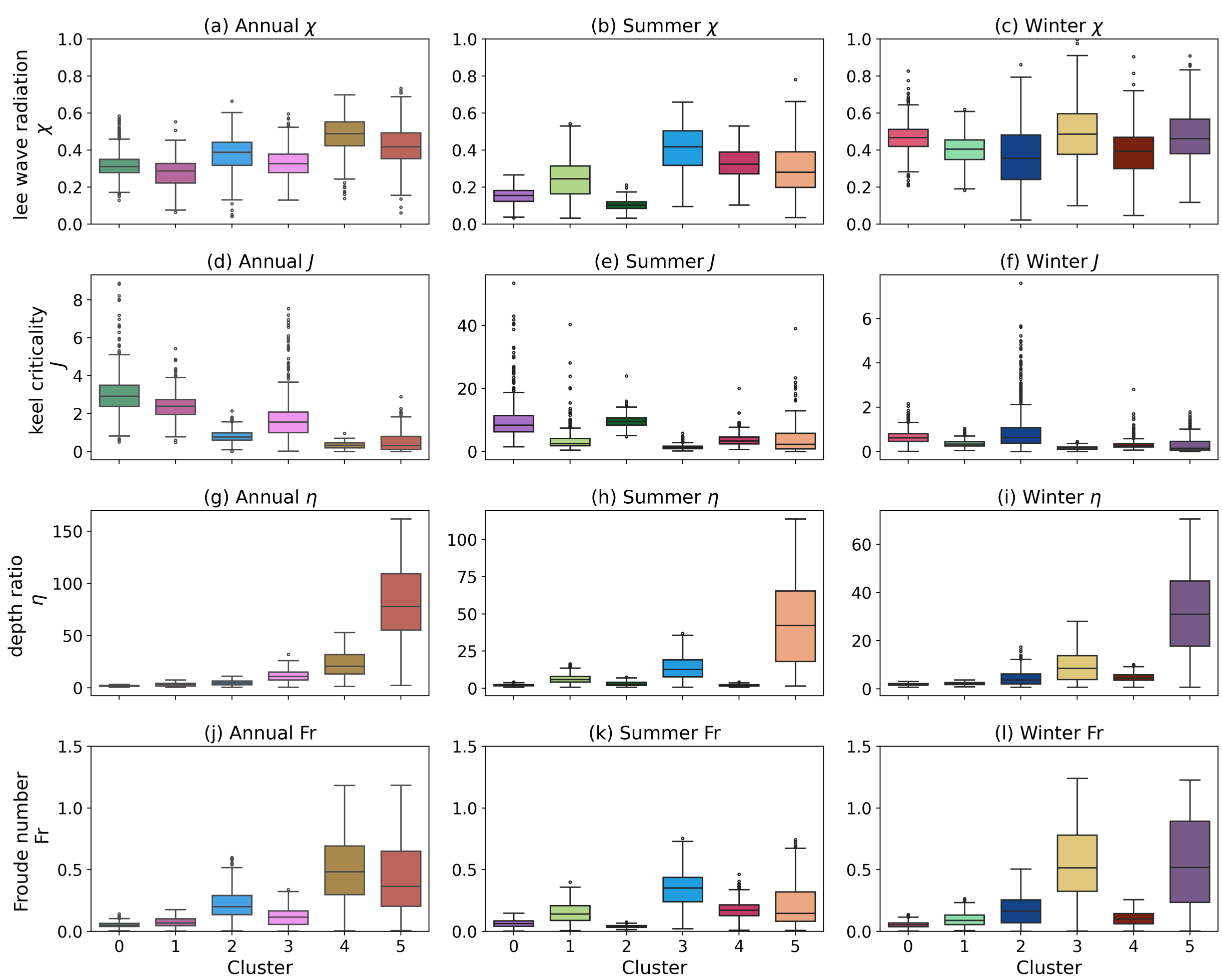
Revised Figure 3: Bayesian Information Criterion (BIC) scores for GMM fitted to the five-dimensional feature space composed of χ, J, η, and Fr. Models were fitted for cluster numbers ranging from 1 to 19. Each model fitting was repeated 20 times with different random initializations to assess variability in BIC values; error bars indicate ± 1 standard deviation. The values are based on the values averaged over different time intervals: (a) annually, (b) over the summer months, and (c) over the winter months.



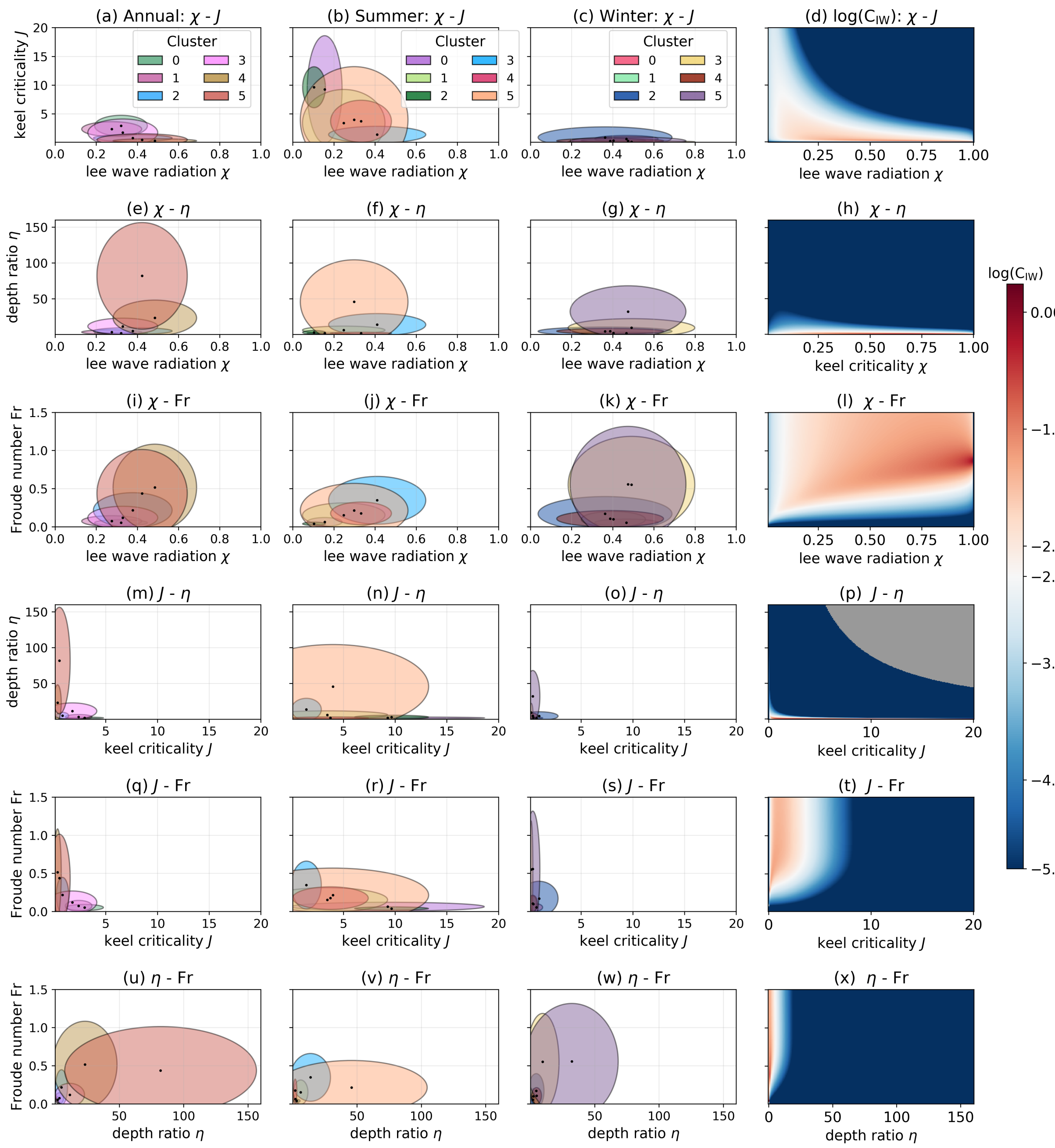
Revised Figure 4: Posterior probability maps for each of the six clusters identified by the GMM, based on time-averaged standardized nondimensional parameters (χ, J, η, and Fr. The clusters are based on the values averaged over different time intervals: (left) annually "A", (middle) over the summer months "S", and (right) over the winter months "W". Each subplot shows the posterior confidence that a given spatial grid cell belongs to the respective cluster. Clusters within each temporal averaging space are all ordered in the descending proportion of data points that belong to each cluster (i.e., most data points belong to cluster 0) and the proportion is shown in each subplot title (e.g., 40% of the annually-averaged data points below to the cluster A0).



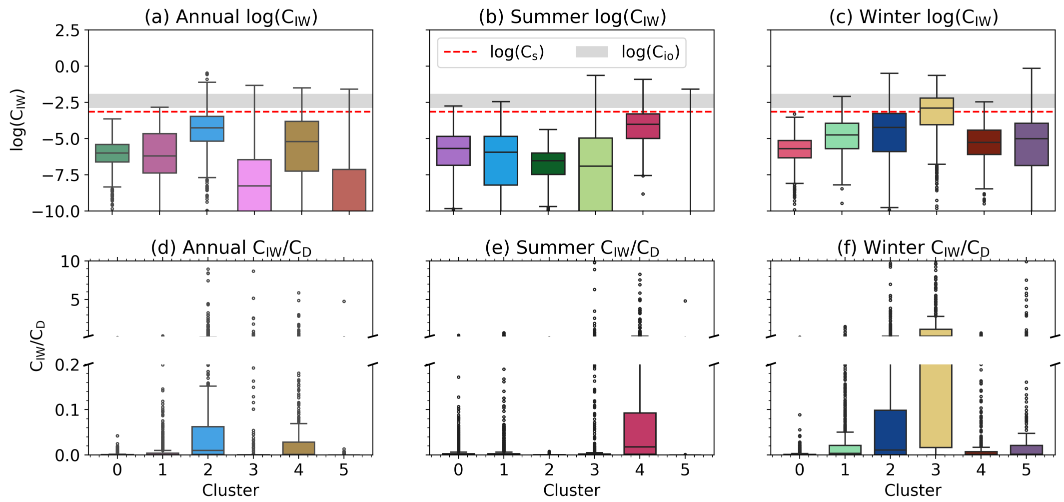
Revised Figure 5: The spatial distribution of six statistically inferred regimes ($K = 6$), each represented by a unique color, over the Arctic Ocean domain, derived from a GMM fitted to standardized time-averaged nondimensional parameter values across all spatial grid points. The values are based on the values averaged over different time intervals: (a) annually, (b) over the summer months, and (c) over the winter months. Clusters within each temporal averaging space are all ordered in the descending proportion of data points that belong to each cluster (i.e., most data points belong to cluster 0).



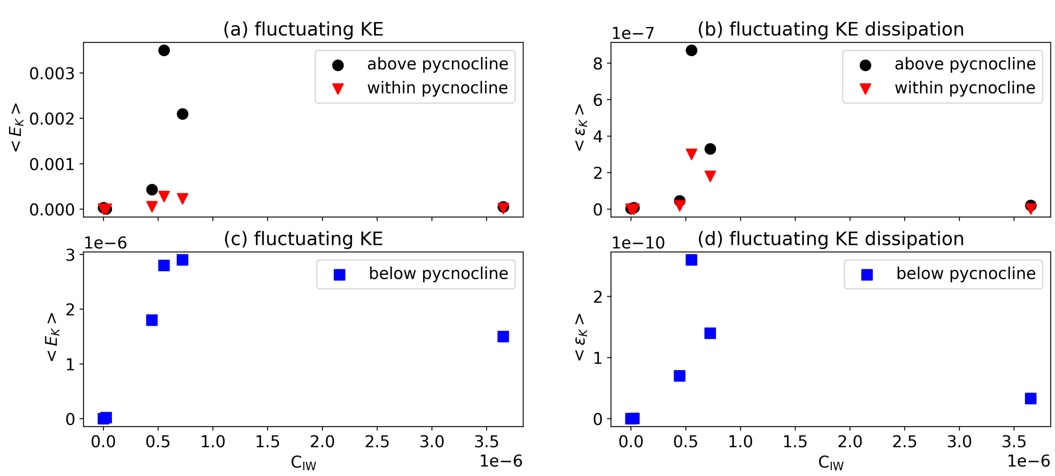
Revised Figure 6: Distribution of nondimensional variables across six GMM clusters based on annual-averaged nondimensional variable values shown in Fig. 4. Box plots summarize the spread of each computed nondimensional parameter (χ, J, J, η, and Fr) across the six clusters identified by the GMM. Each subplot corresponds to a single parameter, with individual boxes showing the interquartile range, median, and outliers for each cluster. The values are based on the values averaged over different time intervals: (left) annually, (middle) over the summer months, and (right) over the winter months.



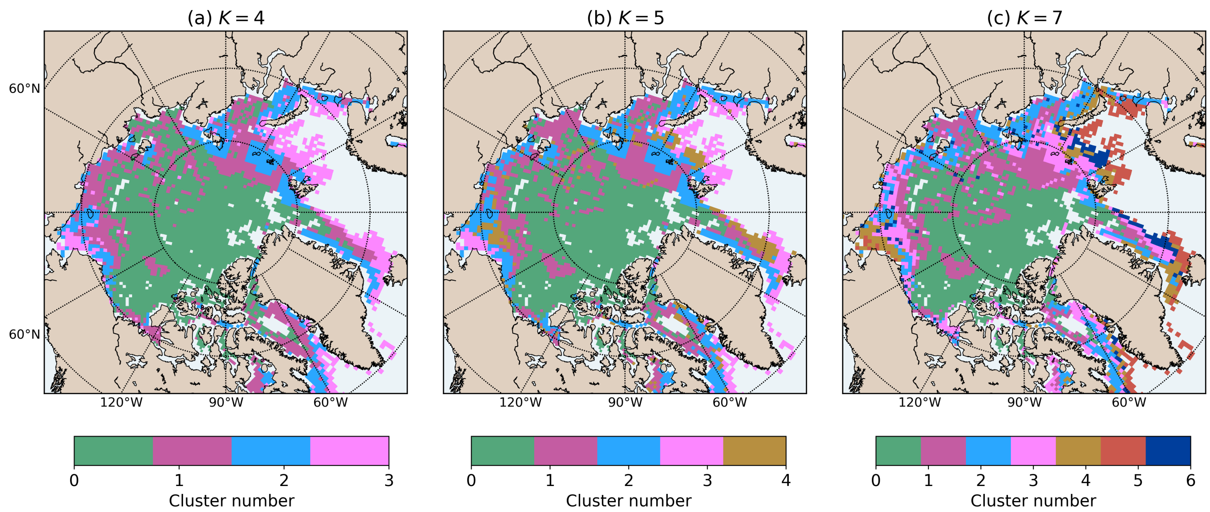
Revised Figure 9: (left three columns) Pairwise ellipse plots showing the cluster-mean values and associated variability across six GMM-identified regimes (columns from left to right: Annual, Summer, Winter). Each colored ellipse is centered at the cluster mean for the variable pair shown and spans two standard deviations along each axis, capturing the internal spread of that cluster. (right column) internal wave drag $C_{\rm IW}$ induced by the ice keel calculated from the parameterization expression over the joint pairwise parameter range. The values of $C_{\rm IW}$ are plotted on a logarithmic scale and the white values (center of the colorbar) corresponds to the the canonical ice-ocean drag coefficient value of $C_{\rm D} = 5.5 \times 10^{-3}$ ($\log(C_{\rm D}) = -2.26$). Subplots on each row correspond to the following pairs: (a-d) χ–J, (b-h) χ–η, (i-l) χ–Fr, (m-p) J–η, (q-t) J–Fr, and (u-x) η–Fr. Note that the grey shaded region in (p) represents undefined values in the parameterization due to the large values of J and η.



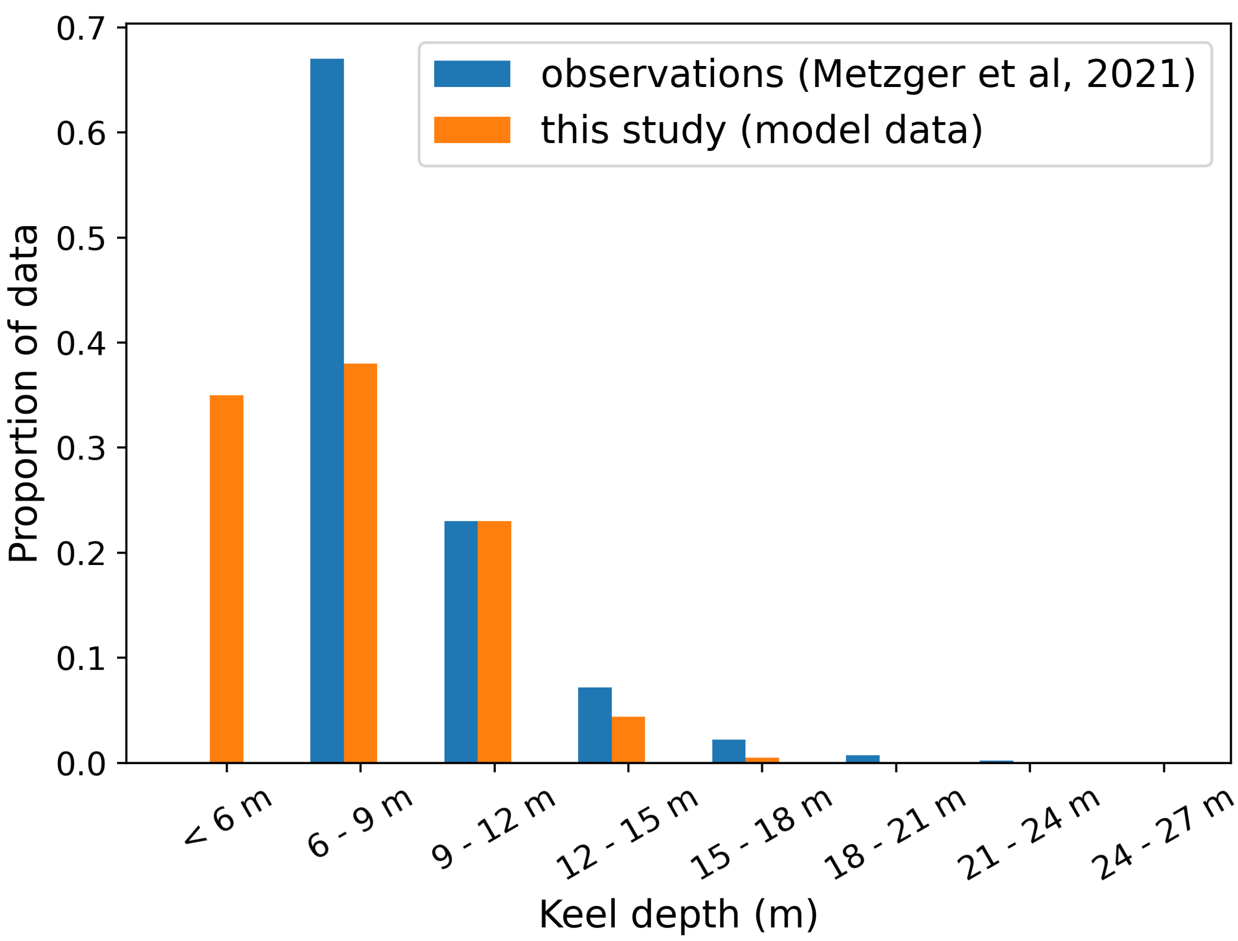
Revised Figure 11: Distribution of internal wave drag C_{IW} induced by the ice keel calculated from the parameterization expression for each of the clusters based on data averaged over: (a, d) annually, (b, e) over the summer months, and (c, f) over the winter months. Panels (a-c) show the values of C_{IW}, with grey shaded region representing the range of ice-ocean drag coefficients C_{io} and skin drag coefficient C_{S} estimated from observations. Panels (d-f) show the ratio between the parameterized values of C_{IW} for each cluster and the canonical ice-ocean drag coefficient value of $C_{\mathrm{D}} = 5.5 \times 10^{-3}$. Note that in (a-c), in order to better see the differences across clusters with larger internal wave drag, the $y-$axis is cropped; so values for some clusters that fall below $\log(C_{\mathrm{IW}}) = -10$ and are too small to be shown. Also, note that in (d-f), the vertical y-axis is broken into two intervals $[0, 0.2]$ and $[0.2, 10]$ in order to show the distributions for clusters with both small and large values (e.g., cluster W3).



New Figure 12: (a, c) fluctuating KE and (b, d) fluctuating KE dissipation from the idealized 2D numerical simulations plotted against the IW drag coefficients C_{IW} calculated based on the current parameterization for the same values of nondimensional parameters as used to initialize the numerical simulations. Panels (a, b) show the energy terms averaged in the regions above the pycnocline (within the mixed layer) and within the pycnocline, and (c, d) averaged below the pycnocline (within the stratified portion of the domain).



Supplementary Figure 1: The spatial distribution of GMM clusters for the annually-averaged data changing input values of the number of clusters K: (a) $K = 4$, (b) $K = 5$, and (c) $K = 7$. The case with $K = 6$ is shown in main text Fig. 6(a). Clusters are all ordered in the descending proportion of data points that belong to each cluster (i.e., most data points belong to cluster 0).



Supplementary Figure 2: Distribution of keel depths based on (blue) the observational study by Metzger et al (2021) and (orange) model output from Flocco et al (2024) that is used as inputs in this study. For the model output, this distribution reflects all data points, not only the ones within the lee wave radiation range ($\chi < 1$).