Dear Dr. Michel Tsamados,

Thank you for your letter and valuable suggestions, we have read and cited the reference you provided and have added more introduction on the statistical forecast (see lines 72-80 in the revision).

Thanks for the reviewers' comments on our manuscript. Those comments are valuable and very helpful. We have read through the comments carefully and have substantially revised our manuscript. Based on the instructions provided in the decision letter, we have uploaded the revised files, a point-by-point response, and a marked-up manuscript version. The responses to the reviewer's comments are marked in blue and presented following. We thank you for allowing us to submit a revision and we highly appreciate your time and consideration.

Best Regards,

Yunhe Wang

==

Response to Reviewer 1

General Comments:

This manuscripts investigates the predictability of seasonal sea ice in the Pacific-Arctic sector. The authors develop and use a regional linear Markov model with a suite of atmosphere-ocean variables. They find that prediction skill of this model is improved relative to a similar previous pan-Arctic model and that there are key sources of predictability for sea ice in different seasons of the Pacific-Arctic sector. The paper is well-written, the structure makes sense, and the figures are clear. I find the use of this regional statistical model interesting and the results are clear and novel, but I think there are a few flaws in the development of this model that may have an impact on the key results.

Thank you very much for reviewing our manuscript. We appreciate your valuable comments. We are confident that the revision made based on these comments has improved our manuscript. Please see the details in the following replies to the specific comments.

The main concern I have is if the authors are missing a key source of predictability for summer Arctic sea ice: sea ice thickness. Numerous studies going back to the seminal work by Blanchard-Wrigglesworth et al., (2011) have shown that sea ice thickness is a key source of predictability for summer Arctic sea ice. In this Markov model, it is excluded, with no clear reason why. I think the authors should add this variable or motivate why it is excluded in a clear fashion. For instance, rather than stating in words why it is not included, it would be nice to see supporting figures that show what happens when it is included. Or, the main figures of the paper should be revised/updated with this variable included. Similarly, the authors exclude subsurface ocean heat content as a variable. This too has been shown to be a key source of predictability for wintertime Arctic sea ice. In the detailed comments below I have suggested how the authors can improve this by including SIT (PIOMAS) or OHC (ORAS5).

Thanks for the valuable suggestions. Following the suggestions, we have added OHC and SIT in the model and recombined variables in the model experiments. The new variable-combinations was shown in Table 1.

Table 1. Variable combinations in cross-validated experiments. V1 represents the No. 1 variable-combination. √ represents the variable included in the corresponding combination.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
SIC	√	√	√	√	√	√	√	√	√	√	√	√
OHC		√		√	√	√	√	√	√	√	√	√

SST			√	√						√	√	√
SIT					√							√
SAT						√						√
Surface net turbulent heat flux								√			√	√
Surface net radiative flux							√			√		√
850hPa GPH, U, V									√			√

The updated predictive skill measured by the percentage of grid points with significant anomaly correlation coefficient (PGS) and mean RMSE for each lead time in each season are calculated and shown in Figure 3. We find that both OHC and SIT contribute substantially to predictability. Based on the PGS and RMSE, we primarily chose three superior model configurations marked by black boxes in Figure 3 for each season, respectively. To determine which model configuration produces the best prediction in each season, we spatially average the SIC prediction skill from these superior models with 1- to 12-month leads (Figures 4 and 5).

As a model construction principle, we choose the minimum number of variables and modes to achieve the same level of skill, avoiding possible overfitting. Based on the PGS and RMSE, we chose V9M16 as the best model in winter since it shows the highest PGS. Similarly, we chose V11M20 in spring and V5M10 in summer. In Autumn, The model skill from V5 is obviously superior at 1-5 lead months, while V12 dominates prediction skill beyond the 8-month lead. We decided to choose V5M7 because it has a relatively higher mean skill and fewer variables and modes. The updated model shows more skill than the model developed in the original manuscript.

The model experiments also indicates that OHC contributes more model prediction skill than SST in all seasons (Figure 3). The model built on the data matrix of SIC, OHC performs better in winter and spring (Figure 3), which indicates that the OHC is a considerable source of memory to provide sea ice prediction skill in the cold season and plays a key role in the evolution of the sea ice conditions. In addition, adding SIT to the model has a substantial contribution to the prediction skill in the warm season, indicating that sea ice thickness is a key source of sea ice predictability within the Arctic Basin in the warm season especially in summer. However, SIT has a negative contribution to the prediction skill in the cold season (Figure 3). [lines 338-428]

Based on the updated model experiments, we have modified all relevant contents throughout the manuscript.

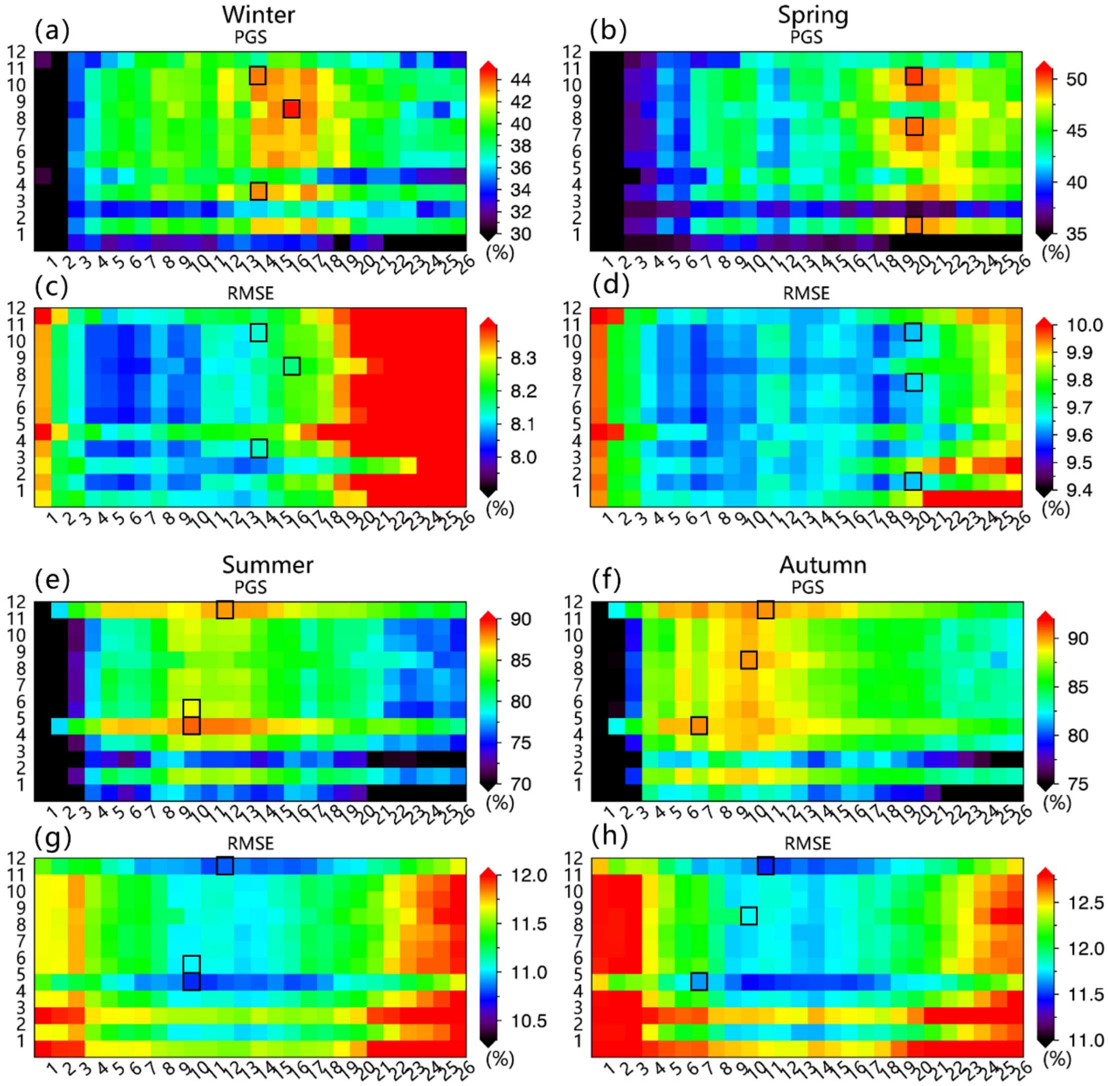

Figure 3. Mean PGS and mean RMSE between the observations and predictions in four seasons. (a) Mean PGS is obtained by averaging all lead months for winter predictions. The x-axis represents the number of MEOF modes, and the y-axis represents the combination of the variables corresponding to Table 1. (b, e, and f) are the same as (a) except for spring, summer, and autumn respectively. (c, d, g, and h) are the same as (a, b, e, and f) except for RMSE.

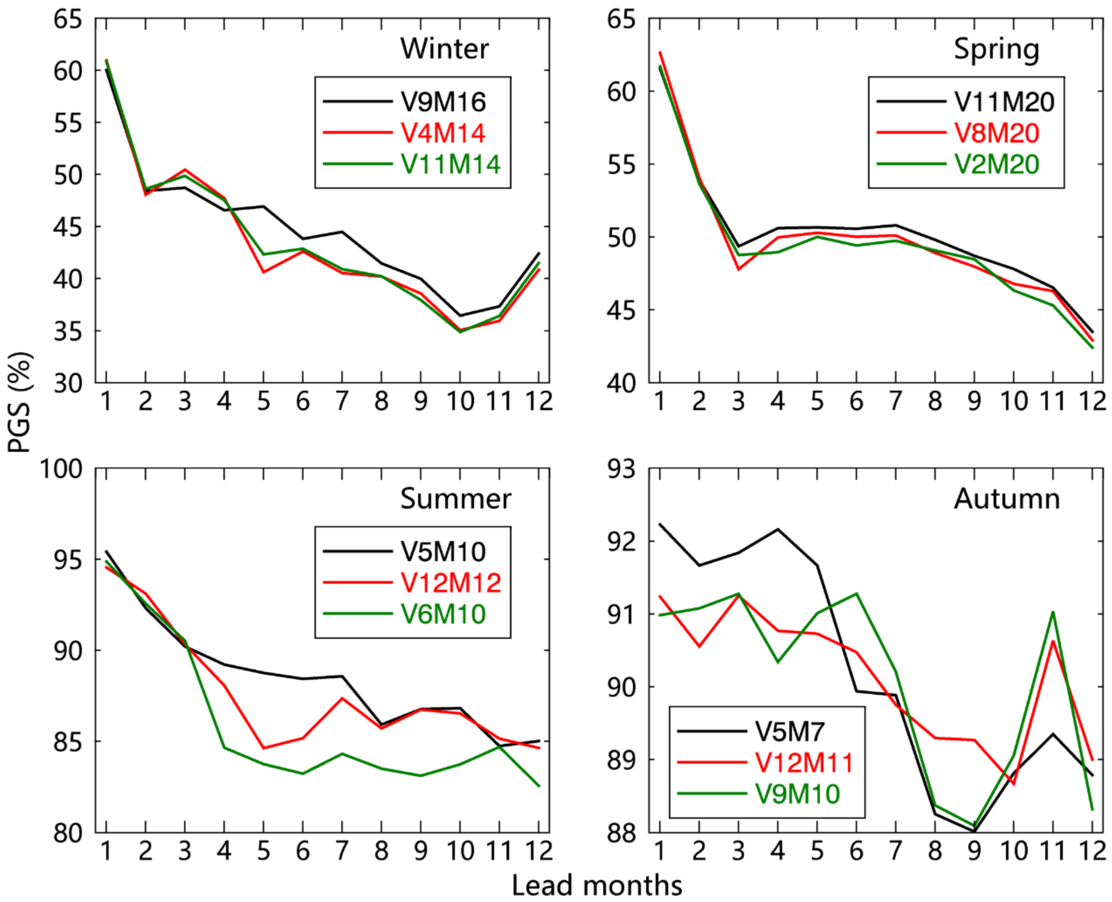

Figure 4. PGS for the preliminary selection of superior models in each season.

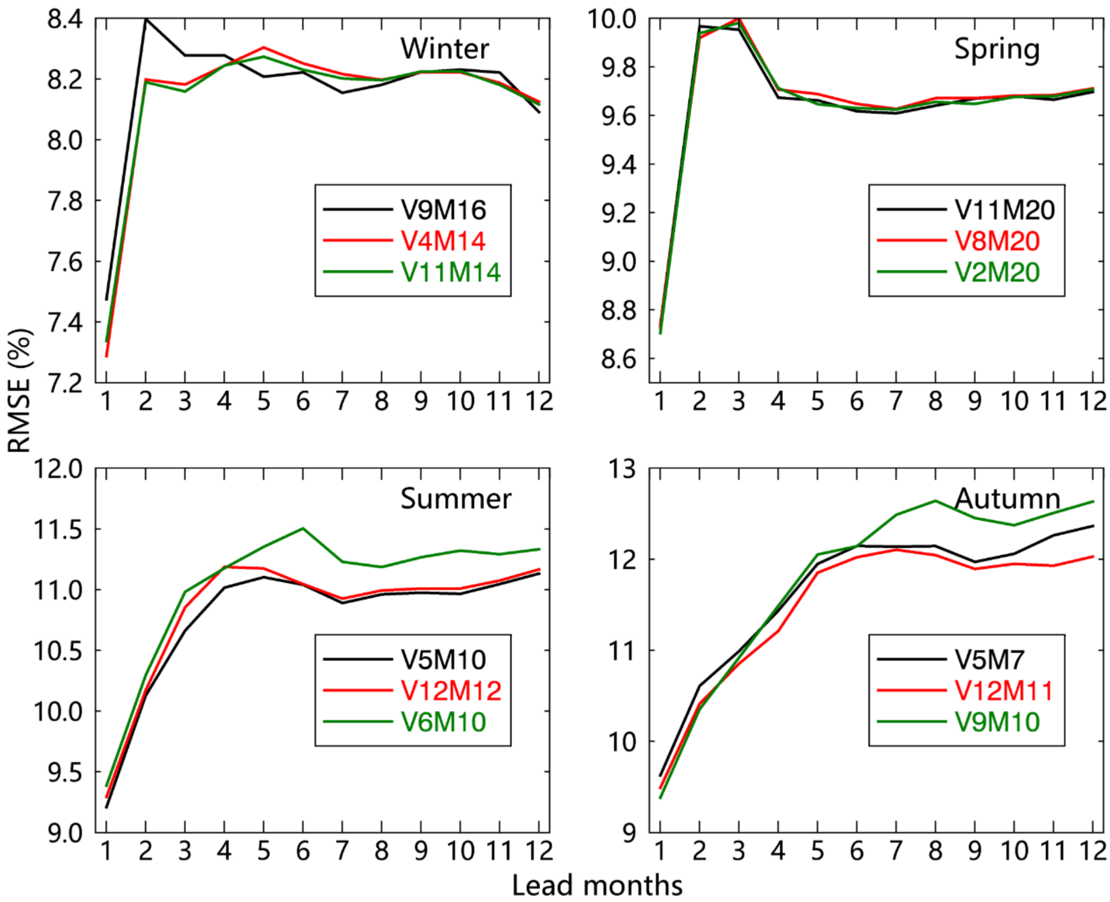

Figure 5. Same as Figure 4 but for RMSE.

Specific Comments:

Data and Methodology:

The Model:

L147-148: It seems this spatial domain (40°N to 84°N, 120°E to 240°E) captures much (if not all) of the Pacific-Arctic sector. But, it feels a bit arbitrary where exactly the longitudinal domains end. For instance, if the same analysis is performed from 90°E to 270°E, are the results similar? Given that the authors base some of this domain choice on regions that exhibit different sea ice variability (i.e., Figure 1), if the authors instead calculated regional sea ice area domains and computed the standard deviation of each region, then grouped the regions by similar magnitudes of variability, will the results be similar? To me this seems like a more precise way of quantifying the Pacific-Sector as regions like the Canadian Archipelago or Laptev Sea could be included or excluded. Basically, I'd like for the authors to conduct this sensitivity analysis and either provide a short explanation of why the results are insensitive to the exact domain or add some supporting information to justify this choice.

Thanks. According to your suggestions, we have conducted the sensitivity analysis of the model domain on prediction skills in the Bering Sea and the Sea of Okhotsk with the same model configuration and different size of model domain. The model domain is defined by 90°E to 270°E, 120°E to 240°E and 135°E to 225°E respectively. The results show that the prediction skill patterns based on three model domains show high similarity in the Bering Sea and the skill based on the model domain (120°E to 240°E) is highest at all month leads. The prediction skill in the Sea of Okhotsk based on the model domain (135°E to 225°E) is highlighted at 5- to 8-month leads, while not well at 11- to 12-month leads. Although the models with different size of model domain have different prediction skills in the Bering Sea and the Sea of Okhotsk, the differences are not significant because all those model domains contain highly similar signals of climate variability. Therefore, the regional model is not highly sensitive to the size of model domain within the Pacific-Arctic sector. [lines 584-605].

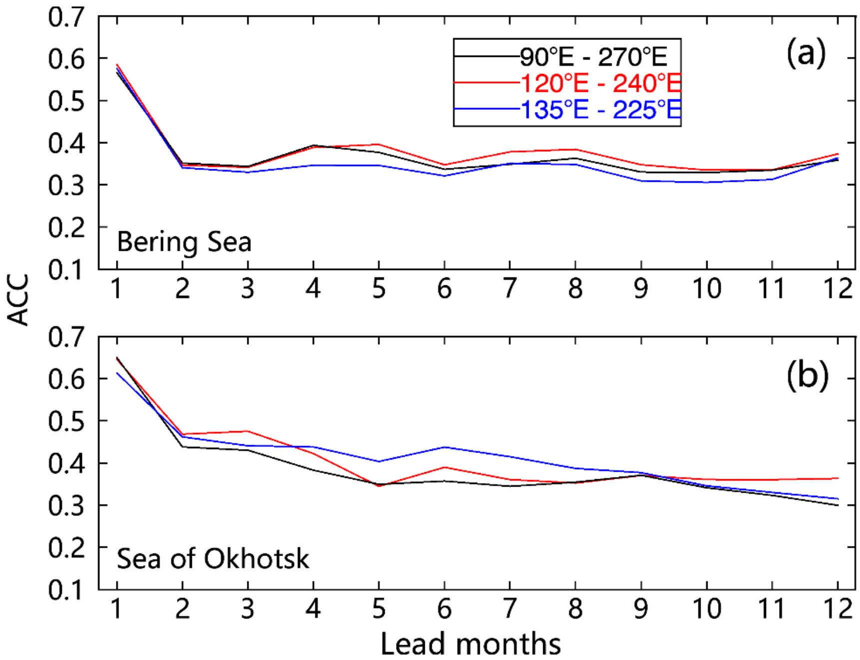

Figure 12. Cross-validated model skills of the regional Markov model with the same model configuration and different size of model domain. (a) The skills are measured by the ACC between predictions and observations with trends from 1980 to 2020 as a function of lead months in the Bering Sea. (b) is the same as (a) except for the Sea of Okhotsk. The model is configured by V9M16 in winter, V11M20 in spring, V5M10 in summer, and V5M7 in autumn. The ACC values are averaged over the area marked in black box in Figure 1.

L148-149: Similarly, can the authors perform the same analysis on sea ice concentration from 0 to 100% instead of 15% to 95%? It is unclear why the authors chose this limit. I can understand the 15% threshold (eliminating area and extent), but why 95%? It would be nice to see the analysis performed without these cutoffs and if they differ, an explanation as to why.

We are sorry for the unclear content. We did perform the analysis on sea ice concentration from 0 to 100%. These cutoffs are only used to eliminate the grid points of mostly open water or mostly complete covered by ice from the ice field since these grid points represent no ice-variability and therefore no predictability. Therefore, we exclude the grid points where 96% of the time SIC is less than 15%, and also exclude the grid points where 96% of the time SIC is larger than 95%. Please check revised sentences in [lines 202-208].

L159-161: I can understand why the authors chose this combination of atmospheric and oceanic variables. But, it appears one key variable is missing: sea ice thickness. Why have the authors omitted this variable? Over the past few years, numerous studies have shown that sea ice thickness is a key source of predictability of Arctic sea ice, particularly for the summer. This, to me, is my biggest concern with this manuscript. I'd like the authors to include sea ice thickness (SIT; from PIOMAS most likely) in their analyses, as I think the results will be different. Perhaps this model may even be more skillful.

Thanks for the valuable suggestions. We have included SIT as model variables and reconstructed the model. Please see above responses.

This also pertains to the ocean domain. What about including sub-surface ocean heat content (OHC)? This has also been shown to be an important source of predictability for Arctic sea ice, particularly on the Pacific side in the wintertime. The authors can include this variability using Ocean Reanalysis System 5 (ORAS5). I think including this variable will improve the weaknesses of the model (as stated in Lines 590-594).

Thanks for the valuable suggestions. We have included OHC as model variables and reconstructed the model. Please see above responses.

Finally, it is shown that SIC contributes to the trends and gives prediction skill. But, why even include this at all? It would be more interesting if the authors could develop this regional Markov model without SIC so that it is predicting SIC/SIA. Perhaps I am missing the motivation to include this. If that is the case, please add a justification in this section.

Thanks, as the trends are parts of the total variability that the signal source for MEOF analysis to capture, we decide to retain the SIC trends in anomalies while building the model. This does not mean that our model can't capture the variability signal from detrended anomalies. To evaluate the contribution of long-term trends to the regional Markov model skill, we conducted post-prediction analysis on the predictions of SIE in the Pacific-Arctic sector. The result shows that the model retains skill for detrended sea ice extent predictions up to 7 month lead times in the Pacific-Arctic sector.

The success of the Markov model is attributed to its ability to capture the co-variability in the coupled air-sea ice-ocean system. Thus, including sea ice in the matrix ensures that we select relevant modes representing the co-variability. This approach is an advantage of our model, which is different from regression models.

L166-186: The introduction of the Markov model could be slightly more clear. In particular, the bolded variables are not present in Eqs. (2)-(4). Does that make them not matrices? As it stands, I think they should be matrices.

Yes, they are matrices, we have changed it, thank you so much. In addition, to make the introduction of the Markov model more clear, we have changed some content. For example, we changed ‘The initial multivariate space is formed to capture the predictable variability in the atmosphere-ice-ocean system by MEOF analysis. Since our focus is on short-term climate variability, the climatological seasonal cycle for the period from 1979 to 2020 was subtracted to obtain monthly anomalies for all variables.’ to ‘We create anomaly time series for all variables from 1979 to 2020 by subtracting climatologies of the same period from monthly mean data and normalizing all anomaly series.’ [lines 232-240]

Model construction and assessments:

Construct an optimal model for each season:

L280-283: I think this cross-validation experiment will be greatly impacted by the addition of SIT and OHC (and/or removal of SIC). This section may need to be reworked with these variables included.

Thanks for the valuable suggestions. We have included SIT and OHC as model variables and reconstruct the model. See above responses.

Conclusions:

L521: I think this paper would benefit from a clear “Discussion” section and then the “Conclusion” section can be streamlined and the results would be clearer.

According to your suggestion, based on the structure and logic of the manuscript, we have moved some content of discussing the impact of linear trends on the model skill that includes sections 3.4 and 3.5 to the section of 'Discussion'. We also added the analysis of sensitivity of model domain on the prediction skill in the section. [lines 497-605]

Technical Comments:

Introduction:

L17: Add hyphen to Pacfic-Arctic? Also, consider keeping it uniform across the paper: Pacific-Arctic instead of switching to Arctic-Pacific

We have changed it throughout the manuscript.

L41: Consider changing “shrinking” to “decreasing” and adding extent, so that it states “The decreasing Arctic sea ice extent contributes…”

We have changed it. [line 42]

L48: Consider changing to “The rapid retreat of summer Arctic sea ice extent has also created…”

We have changed it. [line 49]

L60: Change “which” to “and”.

We have changed it. [line 65]

L88: This sentence does not make sense. Can the authors reframe this? I am not sure what “is initialized in the spring”. Is this in reference to the spring predictability barrier?

We are sorry for the unclear content. We have reframed this sentence to ‘many studies have shown evidence for an Arctic sea ice spring predictability barrier that causes forecasts initialized prior to May to be less skillful and imposes a relatively sharp limit on regional summer sea ice prediction skill (Bushuk et al., 2017a; Day et al., 2014b; Yuan et al., 2016).’ [lines 101-104]

L522: Add hyphen (check hyphens throughout the manuscript)

We have changed it throughout the manuscript.

==

Response to Reviewer 2

General Comments:

The authors present a regional sea ice forecasting method using a linear Markov model. The methodology is variation of an established pan-Arctic version of this model. The novelty is a focus on the Bering Sea and the Sea of Okhotsk and using varying predictors by season. The method shows good prediction skill compared to the pan-Arctic version and to an anomaly persistence model.

Overall, this is a good study and is worthy of publication, but some additions would make this paper more compelling. I would like to see more discussion on the choice of predictors. There have been many studies on sea ice forecasting, only a couple of which are referenced here. A more in-depth literature review could be a good place to start when discussing the choice of predictors. The study area is also a bit confusing. It is stated that this method forecasts in the Bering Sea and the Sea of Okhotsk, but Figures 1, 2, 6, & 7 show forecasting in most of (if not all) of the Chukchi Sea as well. Maybe redefine your study area or clarify your delineation of sea ice regions more clearly.

Thanks for your interest in our research and positive comments. These comments helped us improve our manuscript, and provide important guidance for our future research. Following your suggestions, we have added more discussion on the choice of predictors.

Sorry for the confusion, our model is constructed using the data at grid points in the entire Pacific-Arctic sector including the ice field (Figures 1-8 and 10). Previous studies suggested that low prediction skill occurs in the Pacific sector of the Arctic, particularly in the Bering Sea and the Sea of Okhotsk, compared with other Arctic regions (Bushuk et al., 2017; Yuan et al., 2016). Here the regional Markov model is compared with previous models in the two regions with low prediction skills to assess whether the prediction skill is improved. The results are presented in Figure 9 and Figure 11, respectively. We have made the delineation of sea ice regions more clear in our revision. Please check the details in the following replies of the specific comments.

Specific Comments:

Literature Review: You've listed a couple previous studies of sea ice forecasting methods, but I think you are missing out on a lot of work that has been done, even quite recent publications. I think lines 57-87 could benefit greatly from a more in-depth discussion of previous forecasting methods. Section 2.1 would also benefit from this when discussing your choice of predictors. See references below:

Many thanks for your recommendation. We have read those references and have added more introduction on the dynamic and statistical model, and discussed more on previous typical forecasting techniques [lines 60-80]. In addition, building on the extensive literature studying the sea ice predictability and its related variables, we have added more discussion on predictor choice [lines 120-140, 146-159].

Andersson, Tom R., J. Scott Hosking, María Pérez-Ortiz, Brooks Paige, Andrew Elliott, Chris Russell, Stephen Law, et al. "Seasonal Arctic Sea Ice Forecasting with Probabilistic Deep Learning." Nature Communications 12, no. 1 (August 26, 2021): 5124. https://doi.org/10.1038/s41467-021-25257-4.

Chi, Junhwa, and Hyun-choel Kim. "Prediction of Arctic Sea Ice Concentration Using a Fully Data Driven Deep Neural Network." Remote Sensing 9, no. 12 (December 2017): 1305. https://doi.org/10.3390/rs9121305.

Horvath, Sean, Julienne Stroeve, Balaji Rajagopalan, and William Kleiber. "A Bayesian Logistic Regression for Probabilistic Forecasts of the Minimum September Arctic Sea Ice Cover." Earth and Space Science 7, no. 10 (2020): e2020EA001176. https://doi.org/10.1029/2020EA001176.

Horvath, Sean, Julienne Stroeve, and Balaji Rajagopalan. "A Linear Mixed Effects Model for Seasonal Forecasts of Arctic Sea Ice Retreat." Polar Geography 0, no. 0 (October 15, 2021): 1–18. https://doi.org/10.1080/1088937X.2021.1987999.

Wang, Lei, Xiaojun Yuan, Mingfang Ting, and Cuihua Li. "Predicting Summer Arctic Sea Ice Concentration Intraseasonal Variability Using a Vector Autoregressive Model." Journal of Climate 29, no. 4 (December 8, 2015): 1529–43. https://doi.org/10.1175/JCLI-D-15-0313.1.

For instance, in lines 116-117, is there a reason 850 hPa geopotential height and winds were chosen over other pressure levels?

For the choice of geopotential height and winds, we first calculated the correlation between SIC and geopotential height and winds at different levels. The result showed that SIC is highly related to geopotential height and winds at different levels, including 850 to 200 hPa. The correlation between sea level pressure and SIC is relatively small compared to other levels, which may be caused by the noisiness of the surface process. Due to the barotropic nature of the polar troposphere (Chen, 2005; Ting, 1994), only one pressure level in the troposphere is chosen. We finally chose geopotential height and wind vector at 850 hPa to define the low-level atmospheric circulation, whose interaction with sea ice is most vital relative to higher levels. We have added those discussion in lines 148-156.

Methodology:

Lines 94-99: Here you differentiate the Bering Sea and the Chukchi Sea, but as mentioned above it looks like you are indeed forecasting in both areas. More clarification here is needed.

We are sorry for the unclear content. Here we tried to elucidate that sea ice in the Bering Sea is driven by more complex physical processes than that in other regions. Although the sea ice variability and geographic setting in the Bering Sea are different from the Chukchi Sea in different seasons, our regional Markov model has the ability to capture sea ice anomaly signals in each region, which benefits from that the SIC predictions were performed at each grid cell and each season. Our regional model is indeed constructed using the data at grid points in the entire Pacific-Arctic sector including including the Bering Sea, the Sea of Okhotsk, and the Chukchi Sea.

To prevent misleading readers, we have deleted these contents.

How are you detrending the model in Section 3.4? Removing the trend from the data prior to running the model? Detrending the predictions with the observed trend? With the predicted trend? It is difficult to assess how well the detrended model is actually doing without knowing this.

Sorry for the confusion, we conducted post-prediction analysis. Monthly trends were removed from the predictions and observations, respectively. Figure 10b in the original manuscript shows the correlation between detrended predictions and detrended observations. We have added more discussion on the detrending process in the revision. [lines 509-550]

Other:

Lines 324-328: Similar results were found in Horvath et al., 2020, namely that forecasts made in March showed worse skill than those made in January and February. Perhaps this warrants further analysis.

For this model weakness, we have improved the model construction by adding sub-surface ocean heat content (OHC; using Ocean Reanalysis System 5 (ORAS5)) in the upper 300m and sea ice thickness (SIT; from PIOMAS). The result show that the OHC is a considerable source of memory to provide sea ice prediction skill in the cold season and plays a key role in the evolution of the sea ice conditions. The weakness that low prediction skill for forecasts initialized in March was improved by the updated model (Figure 4). See lines 342-428 for details of improving model construction.

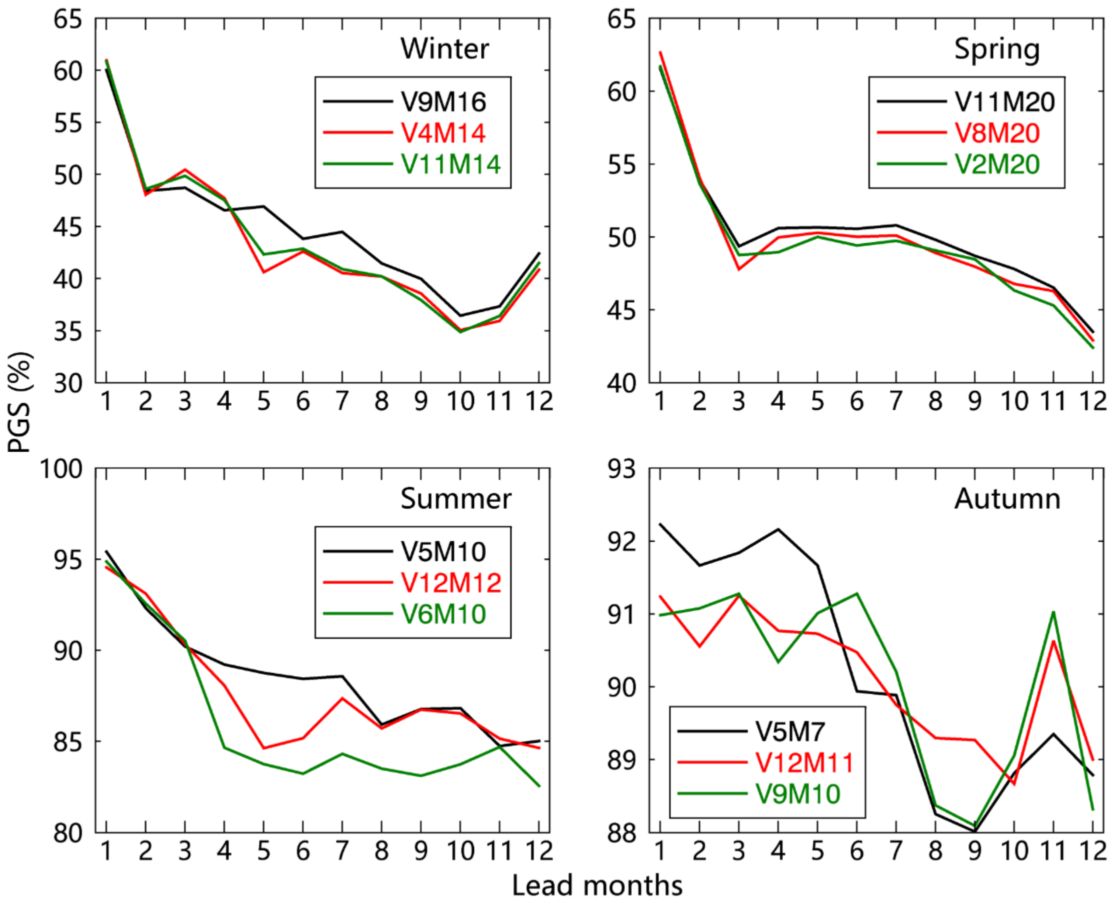

Figure 4. PGS for the preliminary selection of superior models in each season.

Technical Comments:

LN 303-305: The fact that more modes are needed during the cold season is repeated. I suggest combining these sentences to clarify.

Thanks for the suggestions. We have changed 'more modes are needed in the cold season to capture the predictable signal of SIC. This indicates that sea ice in the cold season has requires more modes to capture its variability, likely due to the weaker trends in these months.' to 'more modes are needed in the cold season to capture the predictable signal of SIC, which is likely due to the weaker trends in these months. ' [lines, 360-362]

LN 505: remove 'the'. "In other words,…"

We have changed it. [line 575]

References

Ting, M. F.: Maintenance of northern summer stationary waves in a GCM. J. Atmos. Sci., 51, 3286–3308, doi:10.1175/1520-0469(1994)051,3286:MONSSW.2.0.CO;2, 1994.

Chen, T. C.: The structure and maintenance of stationary waves in the winter Northern Hemisphere. J. Atmos. Sci., 62, 3637–3660, doi:10.1175/JAS3566.1, 2005.

Yuan, X., Chen, D., Li, C., Wang, L., and Wang, W.: Arctic sea ice seasonal prediction by a linear Markov model, J. Clim., 29, 8151-8173, https://doi.org/10.1175/JCLI-D-15-0858.1, 2016.

Bushuk, M., Msadek, R., Winton, M., Vecchi, G. A., Gudgel, R., Rosati, A., and Yang, X.: Skillful regional prediction of Arctic sea ice on seasonal timescales, Geophys. Res. Lett., 44, 4953-4964, https://doi.org/10.1002/2017GL073155, 2017.