This paper examines the European radiative forced responses to Arctic climate by using a combination method of two machine learning techniques, the k-mean clustering and convolution neural network. Specifically, the authors classified six patterns and discussed how these six patterns responds to the European radiative forcing. This paper is interesting, and the topic is crucial for the community; however, the manuscript is not well organized and the results are not well highlighted. Therefore, I do not suggest this paper to be published in Weather and Climate Dynamics before a major revision is made.

Thank you very much for your valuable review and for recognizing the interest and importance of the topic for the community. We appreciate your comments and make efforts to improve organization and presentation of our results to highlight the manuscript's contributions more effectively.

To clarify, our study indeed focuses on examining the Arctic climate response to European radiative forcing. We analyze the anomalies calculated as the difference between the experimental runs (pre-industrial era climate with European negative radiative forcing) and the control (pre-industrial era climate without forcing) across key fields such as sea ice concentration (SIC), 2-meter temperature, PMSET, EP flux, and zonal mean zonal wind.

We used a machine learning framework and developed a new analysis method that helped us understand the mechanism through which the observed anomalies happened. That is, we developed our methodology in a way that helps us answer questions like “what are the main mechanisms behind the observed arctic climate response to the forcing.”

We are committed to revising the manuscript to address your concerns, including reorganizing the content for better clarity and emphasis on the results. We will also provide additional context to further illustrate the novelty and significance of our findings to the reader.

Once again, we sincerely thank you for your time and constructive feedback, which we believe will greatly improve the quality and impact of our manuscript.

Major comments:

The authors applied multiple encoding methods and clustered them into six groups of large-scale circulation patterns. However, the main six patterns in section 3.1 are not well-discussed in their physical meaning. Even though the main occurrence seasons are described, the dynamical interactions of each pattern are not investigated. Is there any existing large-scale circulation pattern that is similar with these group? If not, why and how different the patterns found are compared with the existing patterns. For example, by composing the same timing and location with surface temperature or other variables, can we obtain more meaning from these patterns? This will make the later discussions, such as section 3.2, easier since the physical meaning of each pattern is known.

Thank you for highlighting the importance of discussing the physical meaning and dynamical interactions of the identified large-scale circulation patterns. We appreciate the opportunity to clarify how our analysis inherently considers these aspects and tries to find the clusters' significance in forming the observed anomalies as the result of the forcing. To address your comment, we will first delineate the analytical framework we employed to investigate the paper's main scientific questions. This framework will be further highlighted upon more clearly in the revised manuscript.

We considered the large-scale circulation regime as the main mechanism by which our local forcing influences remote regions such as the Arctic. This is why we performed clustering to group the similar large-scale circulation regimes. As our analysis demonstrates that the forcing may not induce new circulation patterns, we examined how changes in the grouped circulation regimes within the Experimental run contributed to observed anomalies in key Arctic climate variables. Given our focus on Arctic climate variables, we utilized the PMSET pattern associated with each circulation as an additional target similarity measure for our clustering. The PMSET pattern, indicative of how each circulation regime transfers energy into the Arctic, is crucial for determining Arctic climate conditions. Therefore, in our clustering approach, circulation regimes within a similar group share not only similar spatiotemporal patterns in MSLP and $\tau_{300-700}$, but also characteristics related to their associated poleward energy transport. Moreover, by reformulating the anomalies to attribute them to different circulation groups, we identified two main mechanisms through which a circulation cluster can contribute to the anomaly: 1) through changes in the occurrence frequency of that cluster as a result of the forcing (FSDC), and 2) through slight adjustments in the cluster's mean characteristics as a result of the forcing (WCVC). This framework helps us delineate how changes in circulation patterns, whether in occurrence frequency or in the cluster's mean characteristics, contribute to the observed anomalies in the key climate variables.

We showed the cluster centers in the data space for the MSLP field in Figures 6 and 8. The class centers give us information on the main cluster spatial patterns. It was nice that we saw some cluster centers that are similar to the well-known circulation patterns. This means that our clustering with a focus on similarities in data points' spatiotemporal patterns and PMSET leads to the circulation regime classification that somehow resembles the well-known large-scale circulation patterns. However, we were careful in associate our clusters with specific well-known circulation patterns, because they have been defined based on different concepts. We will add this argument to section 3.1.

One of the main goals of the paper is to identify the contributions of various classes to the anomalies observed as a consequence of the applied forcing. We aim to understand the dynamical interactions of each pattern that lead to the observed anomaly. In addressing them, we developed our class contribution formulation, focusing specifically on the dynamic interactions of each cluster that lead to the observed anomalies. This approach, as detailed in our manuscript, emphasizes the interplay between class dynamics and the resultant climatic effects, rather than solely on generic class interactions. We will add the emphasis of this point in the revised manuscript.

In conclusion, we developed a method for clustering circulation regimes, utilizing the associated poleward energy transport as a similarity measure. The clusters derived from this method differ from those generated by conventional approaches, leading us to be careful in associating our clusters with specific, well-recognized circulation patterns. Furthermore, through our class contribution method, we explored various class interactions that contributed to the observed anomaly, thus analyzing the cluster dynamics in the context of the observed anomaly. These insights will be emphasized more in the revised manuscript. We are grateful for your constructive comment, which will undeniably strengthen our manuscript and ensure the clarity and impact of our contributions to the community.

> Another suggestion is that the authors discuss Fig. 13 and 14 in details in section 3.1, and point out the most meaningful patterns that we need to focus on. Then in the section 3.2 and so on, the authors do not need to discuss all the patterns, which makes the comparison simpler.

Thank you for your constructive suggestions. In response, we will reorganize the manuscript to enhance its clarity. The revised results section will begin by comparing the mean states of the Control and Experiment runs (highlighting anomalies) across key climate variables. This will be followed by the large-scale circulation clustering and the class contribution analysis. This modification will show the analytical framework of the paper more clearly.

One of the findings of the paper is the varied contributions of different clusters to the observed anomalies across different seasons and locations. Therefore, in most case, one cannot identify one cluster as the sole source of the observed anomaly. This has been highlighted with some examples in lines 630-645. We agree that we didn't emphasize enough on this finding, and we will highlight this argument more in the revised version. Moreover, we will try to put all the most meaningful results in a more compact format like a table to make these part of the paper more concise.

These modifications should make the paper more accessible and emphasize the novelty and significance of our contributions.

> The authors have too many figures and no clear storyline is provided. Also, the tittle is about Arctic climate responses to European Radiative forcing, which is too broad and no specific results. The authors should summarize the main responses with a more specific title and describe the corresponding physical process, rather than discuss all the results a little bit. In section 3.1, the authors should decide which patterns to focus and try to understand the physical meaning of the patterns. For instance, the patterns have stronger occurrence in summer are not important for poleward wave propagation. Or in Lines 685-694, the authors focus mainly on one pattern and discuss the possible dynamical process of it, which should be more emphasized or considered as the main conclusion of this manuscript.

Thank you for your valuable feedback, which has highlighted the need for a clearer storyline and a more focused analysis within our manuscript. We acknowledge the complexity introduced by the extensive results, which may have obscured the central storyline of our study. In response, we will reorganize the paper to make the analytical framework of the paper more clear and strength the story line.

Regarding the title, as we analyze the impact of the negative radiative forcing on the Arctic climate by focusing on various climate variables, finding a specific title is challenging, and we had to use a broad title. However, we will probably consider refining the title to better reflect the paper's content and focus.

As mentioned earlier, the main focus of the study is to evaluate the effect of different circulation clusters on the observed anomalies. This perspective also guides our writing approach, where we identify the most significant patterns in the anomalies and attempt to attribute them to specific clusters. In some case, as you also noted (lines 685-694), the observed anomalies (the anomalies over Barents-Kara sea in autumn) can be attributed to a single cluster. Then, the discussion about attributing the anomaly to the main contributors is reduced to one cluster. However, in most cases, the situation is more complex, with different clusters playing a role. This is why in discussions on the attribution of significant observed anomalies, we typically consider more than one cluster. However, as you suggested, adopting another perspective for discussing the results could involve focusing on a single cluster at a time and detailing its contributions to the anomalies. However, while this approach facilitates a deeper understanding of each cluster's interactions, it complicates the task of comparing the relative contributions of the clusters to the observed anomalies. This is because the emphasis shifts towards the clusters themselves, rather than attributing the observed anomalies across multiple clusters.

In our paper's discussion, we demonstrate the capability of our methodology to identify potential mechanisms behind each significant anomaly. Some examples include:

- Lines 588-592: Our method successfully attributed the SIC loss in Northeast Canada during autumn to mechanisms hidden in seasonal mean anomalies.
- Lines 593-604: This section illustrates our method's ability to attribute an anomaly associated with a single cluster to changes in both the occurrence frequency of the cluster and the main characteristics of the class. Ultimately, we can explain how these changes, driven by the forcing, impact the observed anomaly.
- Lines 605-613: We attributed the 'warm Arctic and cold midlatitude' pattern to changes in the main characteristics of C1.
- Lines 614-634: C3 is identified as the primary driver of sea ice loss in the Barents-Kara seas. Although a reduction in C3's occurrence frequency in autumn slightly decreases upward wave propagation, this effect is minor compared to the dominant increase in wave propagation driven by changes in the class's mean characteristics.
- Lines 635-640: This section provides an example of how different classes interact to form the observed anomaly.
- Lines 641-650: We discuss our method's limitations in identifying mechanisms and offer possible explanations for these limitations.

In the revised manuscript, we will also incorporate a concise summary table that captures the main findings derived from our methodology, emphasizing our study's contributions to understanding the complex mechanisms driving the observed anomalies in the Arctic. This table, alongside a refined narrative, will ensure that our methodology's capabilities and limitations are presented in a coherent and accessible manner.

We hope that the revised manuscript will present a more focused, comprehensible exploration of our research findings.

> In section 2.4, the authors trained the AAE with only Control run data and lead to the conclusion in Lines 219-220, “This implies that the negative forcing over Europe **does not introduce new discernible spatiotemporal patterns** in the Northern Hemisphere

extratropical large-scale circulation." However, this conclusion seems to be overstated. The difference of reconstruction loss distribution (Fig.3) between the Control and Experiment runs are limited, indicating the same encoder can be used for both Control and Experiment runs. That is, if there are new discernible spatiotemporal patterns in the Experiment run data, they can also be captured by the encoder of the Control run. I suggest the authors also do another way round (train with Experiment) to confirm that the pattern found in both data can be inter-changeably used. Also, I would suggest to weaken the sentence in Lines 219-220.

Thank you very much for this suggestion and for engaging so deeply with our methodology. We appreciate the opportunity to further discuss the logic behind our approach and consider your valuable feedback.

We understand your concern regarding our conclusion and the potential benefits of training an Autoencoder with Experiment run data as well. Our initial choice to train the AAE exclusively with Control run data was driven by the aim to investigate whether the Experiment run introduces any new spatiotemporal patterns that could not be captured by a model trained on the control run. Moreover, in light of your feedback, we will soften the statement in Lines 219-220 to more accurately reflect the limitations and scope of our findings. More specifically, we will revise it from "This implies that the negative forcing over Europe does not introduce new discernible spatiotemporal patterns in the Northern Hemisphere extratropical large-scale circulation." to "This implies that the negative forcing over Europe **may not** introduce new discernable spatiotemporal patterns in the Northern Hemisphere extratropical large-scale circulation."

Your suggestion to also train a separate Auto encoder with Experiment run data to assess the interchangeability of patterns captured in both Control and Experiment runs is indeed an interesting approach. While it does not directly investigate whether there is any new spatiotemporal pattern in the Experiment run, it could potentially provide additional insights into the similarity or divergence of spatiotemporal patterns between the two scenarios.

We implemented training an auto-encoder with solely Experiment run data points as a supplementary analysis. The reconstruction loss of this autoencoder, calculated by inputting data points from both the Experiment and Control runs, is illustrated in Figure 1. Similar to AAE, the new auto-encoder managed to reconstruct data points from both the Control and Experiment runs with comparable levels of accuracy. This outcome suggests a notable degree of pattern interchangeability between the Control and Experiment runs, as captured by the new model. We will add the resulting reconstructed loss and associated discussion in the appendix of the paper. This will not only address your suggestion but also strengthen our conclusions by providing a more comprehensive understanding of the spatiotemporal pattern dynamics between Control and Experiment runs.

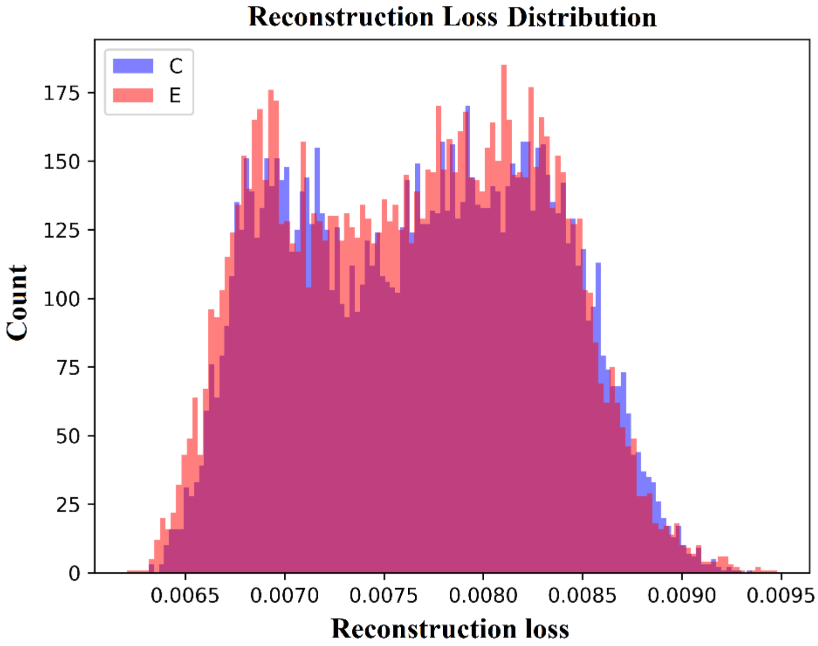

Figure 1: Reconstruction loss distribution for the Control (C) and Experiment (E) runs generated by an auto-encoder, with a similar architecture as AAE, trained only on the Experiment run data points with the same training strategy as AAE.

There is little discussion on the mean states between Control and Experiment runs. The authors should first discuss how different it is when the forcing is imposed through traditional methods, such as composite or even EOF/PCA. Since the authors directly start using the six patterns for explaining the forced signature, the results are hard to follow. The authors should give a basic state of how the forced signal look like. And why we need to use such complicated methods to study the forced signal?

Thank you for your insightful comments. We agree that addressing them will significantly enhance the reader's comprehension of our findings.

We will revise the manuscript to begin the results section with a detailed discussion of the mean states and anomalies of various climate variables (PMSET, 2m temperature, SIC, zonal mean zonal wind, and EP flux) as observed in the Control and Experiment runs. The mean state of the different climate variables were discussed at the beginning of the related sections in the pre-print version. Starting our results section by discussion on the climate variables mean states and anomalies, will set a clear foundation for subsequent analyses on large-scale circulation and cluster contributions. We believe this approach will greatly improve the clarity of our analytical framework and storyline.

We also implemented EOF based clustering for comparison purposes. The EOF analysis was applied to the whole Control and Experiment runs data points treating them as one time series. The first five leading EOFs were only used for spanning the reduced representation. The corresponding PC deriving from these EOFs serves as the features within the reduced representation. The five leading EOFs account for 64.7% of the total variance within our dataset. The centers of these classes, along with their monthly occurrence frequencies, are depicted in Figure 2 and Figure 3, respectively. The class centers and occurrence frequencies resulting from this clustering approach are similar to those obtained by performing k-means clustering directly on the data space. We will elaborate on this comparison in the Appendix section of the revised manuscript.

Additionally, we will include t-SNE visualizations on different data representations in the Appendix to further argue the advantages of our chosen methodology over traditional approaches. This visualization aims to provide a compelling argument for the proficiency of our method by highlighting the distinct data embeddings and the insights they offer.

We are confident that the addition of these analyses and the reorganization of our results section will not only strengthen our argument but also significantly improve the overall quality of our manuscript.

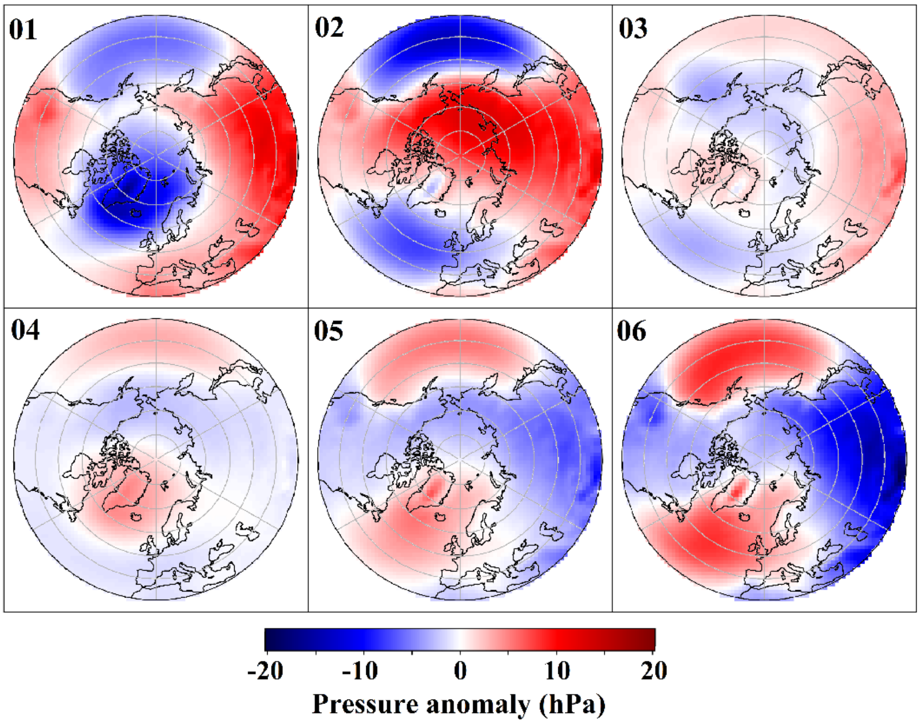

Figure 2: Mean MSLP fields for clusters derived from k-means clustering (6 clusters) on PC time series derived from the 5 leading EOFs. The fields for each cluster are plotted as anomalies relative to the mean MSLP field for all data points, encompassing both the Control and Experiment runs.

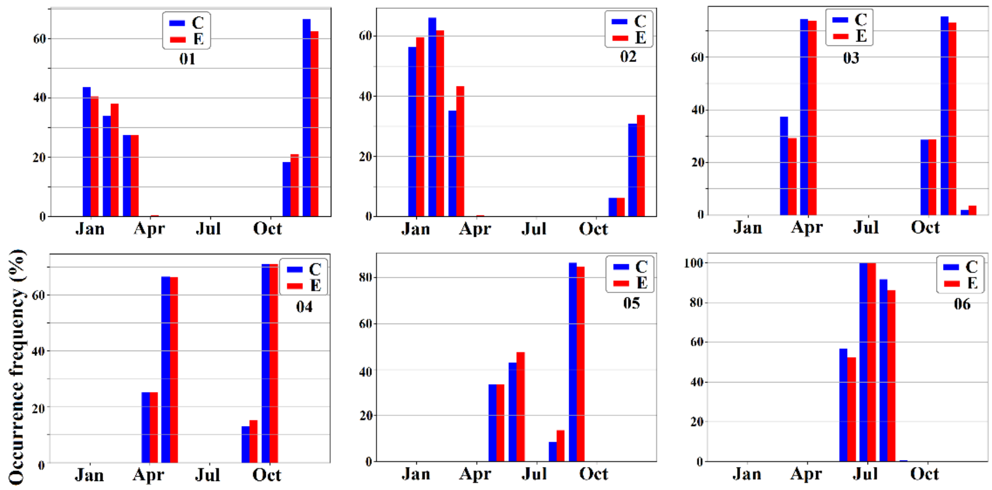

Figure 3: The monthly occurrence frequency of the 6 clusters showed in Figure 2 for the Control (blue) and Experiment (red) runs.

Other comments are following:

> B1-04 is similar with Fig. B3-01; however, their occurrence seasons are so different. Why? Same for the Fig. B5-05. Are they corresponding to the C1 and C2 in Fig.6? Even though the methods are different, the results should match between each other in certain degree.

Thank you for pointing out the similarities and differences in occurrence seasons between the Figures B1, B3 and 6. Your question highlights a key aspect of different clustering methodologies, which is the dependency of clustering outcomes on the chosen representation space. Different representation spaces can influence which data points are considered close to each other according to the specific geometry of that space. Clustering algorithms like k-means then group data points based on their proximity in the chosen representation space. For instance, the clustering performed on PMSET indices (Figure B1) is guided solely by similarities in PMSET indices. That is, data points with similar PMSET indices lie closer together in this representation. This contrasts with the clustering performed on data space (Figure B3), where the Euclidean pixel-wise distance between data points dictates the data points proximity and the resulting clustering.

As you mentioned, there are similarities between Class 04 in Figure B1 and Class 01 in Figure B3, suggesting they could be perceived as negative phases of each other if viewed as EOF patterns. However, they have different monthly occurrence frequency. It is important to notice that they are not EOF patterns, nor do they originate from clustering with the identical focus, as discussed in the previous paragraph. Cluster 04 in Figure B1 is derived from clustering data points based on their similarities in PMSET indices, meaning that data points forming this cluster primarily exhibit similar PMSET characteristics. On the other hand, Cluster 01 in Figure B3 represents one of the mean states of winter circulation. It emerges from clustering based on data space, which primarily groups the seasonal cycle.

To identify similarities in patterns, one should compare clustering results that share a similar focus. For instance, k-means clustering on data space, as shown in Figures B3 and B4, and clustering on the FAE latent space both primarily reveal the seasonal cycle as the dominant pattern within the dataset. Consequently, the class centers in these representations are somewhat comparable. A case in point is Cluster 06 in Figure B3 and Cluster 01 in Figure B5 can be perceive as similar clusters. Similarly, clustering based on PMSET indices (Figures B1 and B2) and clustering on the MCAE latent space exhibit resemblances because both approaches utilize PMSET indices as a similarity measure at different levels. For example, Cluster C5 in Figure 6 of the manuscript and Class 04 in Figure B1 show some similarities.

The divergent outcomes from clustering across different representation spaces underscores the sensitivity of clustering outcomes to the representation space. Conventional clustering methodologies only focus on data point similarities using standard measures of similarity, such as Euclidean pixel-wise distance between data points or PCA-reduced (Principal Component Analysis) space representation. One of the main contribution of this paper is to introduced a method that can produce a representation of the data points that can effectively distance the data points based on both the similarities among the data points and their alignment with a target objective similarity, here the PMSET patterns. We will better emphasize this contribution in the revised version and further use t-SNE visualization to provide deeper insights into how various clustering methods perform across different representations.

> Line 231, why 40 indices in total? 4 latitudes with 5 PMSET values and two levels of troposphere? It is easy to miss the two levels of troposphere.

Our decision to choose these specific indices was driven by a desire to comprehensively capture the three-dimensional spatial patterns of the PMSET associated with the data points, which is crucial for understanding the dynamics of the atmosphere energy transport in our study.

To elaborate, the five PMSET indices per profile are intended to represent the main characteristics of the PMSET profiles. Furthermore, the choice of four latitudes was aimed at capturing the latitudinal progression of the PMSET. Additionally, incorporating two vertical levels within the troposphere was intended to provide insights into the vertical dynamics of the PMSET. The two vertical levels allow us to observe how PMSET values change with altitude, offering insight into the vertical distribution of atmospheric energy transport.

We appreciate the opportunity to clarify this in our manuscript. We will revise section 2.5.1 of the manuscript to highlight our reasoning behind selecting these indices.

> Line 349, how does the authors consider the “statistically significant increases” in Fig. 10?

As we mentioned in line 119-122, the independent two sample t-test and Welch’s unpaired t-test were used throughout the study to determine the statistical significance levels when samples have equal or unequal variances, respectively, and samples with p-value less than 0.05 were considered statistically significant. Using the same methodology, the doted points in figure 10 are consider statistically significant

> Line 359-361, “between the daily mean PMSET” of upper and lower troposphere?

Exactly. We appreciate you highlighting this point. In the revised version of our manuscript, we will refine this sentence to more precisely express that our analysis involves comparing the daily mean PMSET values between the upper and lower levels of the troposphere.

As a final remark, we would like to highlight the main contributions of our paper, which we plan to better emphasize on the final revision of the manuscript:

1 - We introduced a deep-learning-based method that generates an effective representation of data points. This representation accounts for both the similarities among the data points and their alignment with a target objective similarity, here the PMSET patterns, to effectively distance the data points. This method was applied to data points consisting of 8 consecutive days of MSLP and $\tau_{300-700}$, capturing dynamic patterns in the circulation data points.

2- We developed a formulation for class contribution, focusing on the dynamic interactions of different circulation patterns that lead to observed anomalies. By reformulating the anomalies to attribute them to different circulation groups, we identified two main mechanisms through which a circulation cluster can contribute to the anomaly: (1) through changes in the occurrence frequency of that cluster as a result of the forcing (FSDC), and (2) through slight adjustments in the cluster's mean characteristics as a result of the forcing (WCVC). This framework delineates how changes in circulation patterns, whether in occurrence frequency or in the cluster's mean characteristics, contribute to the observed anomalies in key climate variables.

3- We utilized this advanced methodological approach to understand the role of different circulation patterns in forming the anomalies. These patterns were grouped based on similarities in their spatiotemporal behavior and their associated PMSET patterns. This grouping and the class contribution formulation helped us understand how these interactions lead to observed anomalies resulting from negative radiative forcing over Europe. Among the main findings are:

- In most cases, the change in a cluster's main behavior is more significant than changes in the occurrence frequency of that cluster when considering the cluster's contribution to the anomaly. For example, the decreased occurrence frequency of a cluster associated with a high-pressure system over Northern Eurasia and Scandinavia in autumn, observed in the Experiment run, led to reduced upward wave propagation. However, a slight adjustment in the mean behavior of this cluster in the Experiment run resulted in an increase in upward wave propagation. This increased wave propagation dominates the cluster's contribution to the observed anomaly in upward wave propagation.
- Our method's capability to attribute observed anomalies, such as the SIC loss in the Barents-Kara Seas in autumn, the warm Arctic and cold mid-latitude in winter, and others.