Title: Arctic Climate Response to European Radiative Forcing: A Deep Learning Approach
Authors: Sina Mehrdad, Dörthe Handorf, Ines Höschel, Khalil Karami, Johannes Quaas, Sudhakar Dipu, and Christoph Jacobi

I would like to thank the authors for their efforts in addressing my comments. The structure and storyline of the manuscript have been improved, allowing for a clearer delivery of the results. However, regarding causality, I am hesitant about the statement the authors made in their replies: 'Deriving causality in the context of regional forcing on Arctic climate is complex and multifaceted.' In a coupled climate system, it seems nearly impossible to clearly distinguish causality. Is it necessary to provide causality, or should one acknowledge that deriving it is very difficult, and/or not necessary? Overall, I am satisfied with this version but have a few additional comments for the authors to consider.

For the machine learning approach, I think using a convolutional auto-encoder architecture to study this topic is both exciting and promising. I also appreciate that the authors provided baseline models (e.g., SOM and EOF) for comparison. However, extending this point further, what new scientific insights can these methods bring to us? For example, the authors mentioned that 'We compare this with the traditional data space representation, which is dominated by the seasonal cycle.' These results may seem straightforward to most meteorologists or climate scientists, as some of their research is based on signals after removing the seasonal cycle. Additionally, the stratosphere-troposphere coupling has been well-known for decades. Does using deep learning reveal something new? Related to this, the authors stated that 'a high-pressure system over Northern Eurasia and Scandinavia in autumn, observed in the Experiment run, led to reduced upward wave propagation.' Is this a new finding that previous studies did not document? The authors may want to elaborate more on this aspect and highlight the new scientific findings from this manuscript, as well as those not found in previous studies.

For the forcing and its connection to remote impacts, the authors argue that the large-scale circulation regime is the main underlying mechanism. My follow-up question is: how can one justify that the forced signal stands out from internal variability, given that large-scale circulation is intrinsically affected by internally-driven components? Many previous studies have stressed out the large internal variability could mask out the forced signal. The authors also mentioned that the forcing does not always produce uniform results. Resonating with previous comments, the authors could emphasize the role of machine learning approaches in this context and highlight what new insights these methods bring to the table and deal with the non-stationary results from the same forcing.