

EGU22-8130

<https://doi.org/10.5194/egusphere-egu22-8130>

EGU General Assembly 2022

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



A comparison of explainable AI solutions to a climate change prediction task

Philine Lou Bommer¹, Marlene Kretschmer², Dilyara Bareeva¹, Kadircan Aksoy¹, and Marina Höhne¹

¹Machine Learning, Technical University Berlin, Berlin, Germany (philine.l.bommer@tu-berlin.de)

²Meteorology, University of Reading, Reading, UK (m.j.a.kretschmer@reading.ac.uk)

In climate change research we are dealing with a chaotic system, usually leading to huge computational efforts in order to make faithful predictions. Deep neural networks (DNNs) offer promising new approaches due to their computational efficiency and universal solution properties. However, despite the increase in successful application cases with DNNs, the black-box nature of such purely data-driven approaches limits their trustworthiness and therefore the useability of deep learning in the context of climate science.

The field of explainable artificial intelligence (XAI) has been established to enable a deeper understanding of the complex, highly-nonlinear methods and their predictions. By shedding light onto the reasons behind the predictions made by DNNs, XAI methods can serve as a support for researchers to reveal the underlying physical mechanisms and properties inherent in the studied data. Some XAI methods have already been successfully applied to climate science, however, no detailed comparison of their performances is available. As the number of XAI methods on the one hand, and DNN applications on the other hand are growing, a comprehensive evaluation is necessary in order to understand the different XAI methods in the climate context.

In this work we provide an overview of different available XAI methods and their potential applications for climate science. Based on a previously published climate change prediction task, we compare several explanation approaches, including model-aware (e.g. Saliency, IntGrad, LRP) and model-agnostic methods (e.g. SHAP). We analyse their ability to verify the physical soundness of the DNN predictions as well as their ability to uncover new insights into the underlying climate phenomena. Another important aspect we address in our work is the possibility to assess the underlying uncertainties of DNN predictions using XAI methods. This is especially crucial in climate science applications where uncertainty due to natural variability is usually large. To this end, we investigate the potential of two recently introduced XAI methods —UAI+ and NoiseGrad, which have been designed to include uncertainty information of the predictions into the explanations. We demonstrate that those XAI methods enable more stable explanations with respect to model noise and can further deal with uncertainties of network information. We argue that these methods are therefore particularly suitable for climate science application cases.