Exploring the Existence of Atmospheric Blocking's Precursor Patterns with Physics-Informed Explainable AI

Anh Nhu [1] **Lei Wang** [2]

Abstract

Atmospheric blocking is an atmospheric flow pattern that is quasi-stationary, self-sustaining, and long-lasting that effectively blocks the prevailing westerly atmospheric flows. This blocking is directly linked to large-scale extreme events such as heat waves, yet there is no confirmed study on the precursor patterns that signal atmospheric blocking's evolution. In this paper, we investigate the combination of physics, Convolutional Neural Network (CNN), and eXplainable Artificial Intelligence (XAI) to form a scientific hypothesis: precursor patterns of atmospheric blocking do exist. To investigate the predictability and search for signals of the existence of precursor blocking patterns, we integrated the Two-Layer Quasi-Geostrophic (QG) Model, an idealized model of atmospheric evolution, into the training process of CNN and predict atmospheric blocking, reaching the prediction accuracy of 95%, 88%, and 72% at 1, 5, and 12 lead days, respectively. Next, we employ XAI to highlight spatial patterns that guide CNN's prediction. The resulting composite patterns highlighted by XAI algorithms are physically consistent with the composite ground-truth observations at different lead days. This work hypothesizes the existence of atmospheric blocking's precursor patterns, motivating future fundamental research directions focusing specifically on these precursor patterns.

[1]Department of Computer Science, University of Maryland, College Park, MD, United States. [2]Department of Earth, Atmospheric, and Planetary Science, Purdue University, West Lafayette, IN, United States.. Correspondence to: Anh Nhu <anhu@terpmail.umd.edu>, Lei Wang <wanglei@purdue.edu>.

Accepted after peer-review at the 1st workshop on Synergy of Scientific and Machine Learning Modeling, SynS & ML ICML, Honolulu, Hawaii, USA. July, 2023.

1. Introduction

Although there are variations in the definition of heat waves, they are commonly considered to be periods of exceedingly high temperature (from 90 to 95-percentile of the daily climatology) lasting for at least 2 to 6 consecutive days. A persisting heatwave can seriously endanger human lives and damage the environment as well as our society. In fact, heat waves are a leading cause of many weather-related deaths (Robinson, 2001). Tan et al. (2006) confirm the high association between high temperature caused by heat waves and increased daily mortality rate in Shanghai during the heat waves in 1998 and 2003. Unfortunately, we lack a firm understanding of how heat waves evolve. Several studies have investigated and provided evidence for the link between atmospheric blocking and heat waves (Martius et al., 2009; Dole et al., 2011; Bao et al., 2017; Dong et al., 2018; Zschenderlein et al., 2020; Marengo et al., 2021; Kautz et al., 2022). Specifically, atmospheric blocking refers to the situation where high-pressure systems remain quasi-stationary that cover a certain region for an extended period. Due to their persistent high pressure, these systems effectively "block" the normal eastward atmospheric flow, which can initiate the dynamic evolution of extreme weather events, including heat waves and cold spells (Kautz et al., 2022). There are several ways that the formulation of atmospheric blocking can contribute to heat wave developments and intensification. First, during atmospheric blocking, a high-pressure system remains stationary for several days to weeks, trapping air in a particular region. The air trapped causes temperatures to soar and intensify the heat wave conditions. Second, since atmospheric blocking "blocks" normal westerly atmospheric flows, it creates stagnant weather patterns. These patterns prevent the mixing of air masses, limiting the cooling effect from weather systems like storms, leading to prolonged periods of high temperature. Last but not least, normal jet stream behaviors are disrupted by the appearance of atmospheric blocking. As a result, they have to be "detoured" around the blocking system. This effect leads to the development of a "heat dome" where hot air is effectively locked under the block, exacerbating the heating conditions.

Therefore, insights into the atmospheric patterns that initiate the dynamical evolution of atmospheric blocking are crucial

for enhancing our understanding of heat wave dynamics and improving the predictability of these events. However, the physical processes of the dynamical evolution of blocking are still poorly understood, and whether precursor patterns of atmospheric blocking actually exist is still an open question. Taking a step closer to this unaddressed question, in our study, we focus on leveraging Interpretable Physics-Informed CNN to test CNN's predictability of atmospheric blocking and investigate the existence of precursor patterns.

The contributions of this paper are as follows: **(1)** We propose the synergy between Convolutional Neural Networks and the Two-Layer QG Model to capture potential precursor patterns of atmospheric blocking. **(2)** We utilize XAI techniques on trained CNN to highlight highly important patterns for blocking prediction at different lead days. **(3)** The composite highlighted patterns are physically consistent with the composite observations. Based on these preliminary results, we formulate a hypothesis about the possible existence of blocking's precursor patterns.

Two-Layer QG Model is an idealized model of the mid-latitude dynamics of Earth's atmosphere, thus it represents an ideal condition to analyze and search for precursor patterns. Although this model is idealized, it captures the most important features in the atmospheric dynamics such as annular modes and eddy-driven jets' behaviors. The details of the two-layer QG model are described in Section 3.1. To the best of our knowledge, our work is the first study to integrate the two-layer QG model with XAI to study the evolution of extreme heat waves and atmospheric blocking.

2. Related Works

Our physics-AI synergistic approach is greatly inspired by the successes of many prior works in applying Neural Networks and eXplainable Artificial Intelligence (XAI) frameworks for scientific discoveries in Earth and Atmospheric Science. Tom et al. (2020) is a notable work that used Layer-wise Relevance Propagation and Backward Optimization with fully-connected Neural Networks to predict ENSO phases. The composite detected patterns outputted by XAI frameworks agree with the composite observations, validating the reliability and robustness of XAI and Deep Learning in Geoscience research. Many other works also achieved physically consistent results in various geoscience topics such as wildfire prediction (Kondylatos et al., 2022), climate warming slowdowns (Labe & Barnes, 2022), Madden-Julian Oscillation (Martin et al., 2021; Delaunay & Christensen, 2022), forced change detection (Barnes et al., 2020; Rader et al., 2022) and extreme precipitation (Davenport & Diffenbaugh, 2021). However, previous studies primarily concentrated on assessing the credibility of XAI by contrasting their findings with known patterns. Therefore, the integration of physical models with ML frameworks to investigate and formulate hypotheses about unknown phenomena was overlooked. Nevertheless, the achievements demonstrated in existing research lay a strong foundation for us to delve into the collaborative approach between Two-Layer QG Model and XAI, thereby uncovering potential precursor patterns associated with atmospheric blocking.

3. Physical Model and Data

3.1. Two-layer Quasi-Geostrophic model

Two-Layer QG model is an idealized physical model of the mid-latitude dynamics of the Earth's atmosphere and is commonly used to investigate large-scale atmospheric phenomena such as atmospheric blocking, jet stream dynamics, and wave propagation. This model consists of two horizontally homogeneous layers of fluid, with each layer having different density and temperature characteristics. The layers are separated by a boundary layer, which acts as an interface between the two layers. This interface layer initially has some equilibrium slope, as shown in Figure 1. In this model, the upper layer is assumed to have a lower density than the lower layer, and both layers have a constant depth.

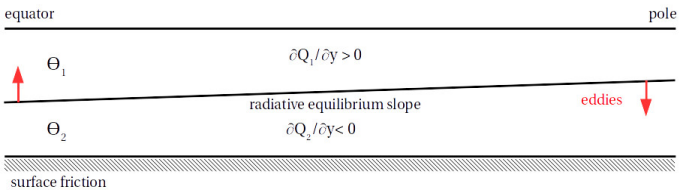

Figure 1. Equilibrium slope of the upper and lower layer, where θ and Q represent the potential temperature and the zonal-mean potential vorticity, respectively (Lutsko, 2018).

The Two-Layer QG model for Potential Vorticity (QGPV) is an extension of the standard Two-layer QG model that incorporates the effect of Potential Vorticity into the model. Potential vorticity is a conserved quantity in the atmosphere that is related to the rotation of air masses, and it is an important quantity for understanding atmospheric dynamics. The partial differential equations for the upper and lower layers are:

$$\frac{\partial Q_k}{\partial t} + J(\psi_k, Q_k) = -\frac{1}{\tau_d}(-1)^k(\psi_1 - \psi_2 - \psi_R) - \frac{1}{\tau_f}\delta_{k,2}\nabla^2\psi_k \quad (1)$$

Considering $k = 1$ denotes the upper layer and $k = 2$ denotes the lower layer, Q_k is the Potential Vorticity, ψ_k is the stream function, and ψ_R is the equilibrium slope in each corresponding layer. ψ_R is selected so that the entire system represented by this model is baroclinically unstable.

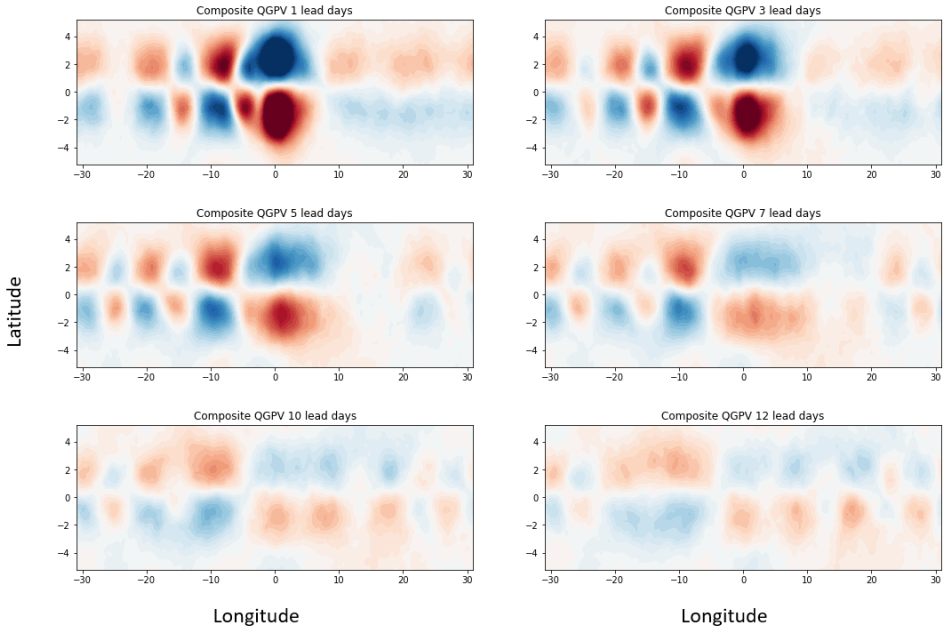

Figure 2. Composite observations of Quasi-Geostrophic Potential Vorticity snapshots at various lead days prior to atmospheric blocking events.

3.2. Data

Using the Two-Layer QG model, we generate multiple independent Quasi-Geostrophic Potential Vorticity dynamical sequences. Among these sequences, 7079 independent blocking cases are manually identified. The blocking cases are formatted as 40-day sequences of Quasi-Geostrophic Potential Vorticity (QGPV) snapshots where atmospheric blocking events first occur on 20^{th} snapshot in each sequence. The composite QGPV observations, which are simply the 2D average of all QGPV snapshots, of several different lead days prior to the first occurrence of blocking can be visualized in Figure 2.

In addition to positive cases, we also include 66856 non-blocking cases. These non-blocking cases are selected such that no atmospheric blocking exists within the 120-day window of each case. In summary, 7079 positive snapshots (blocking) and 66856 negative snapshots (non-blocking) are randomly split into train/validation/test sets with the 80%-10%-10% ratio to train the CNN in our study.

4. Training Pipeline

4.1. Convolutional Neural Network

Our model consists of 3 main convolutional blocks as spatial extractors followed by a fully-connect network as the classifier. Each convolutional block consists of 3 layers: 5×5 convolutional kernels, tanh activation layer, and 2D Average Pooling layer. These 3 convolutional blocks are responsible for learning the most relevant spatial feature in a hierarchical manner. The extracted feature maps outputted by these layers are then learned by a fully-connected neural network consisting of two stacked layers with 84 and 2 neurons, respectively. In the final layer, we use the softmax activation function to convert the outputs into a probability distribution. The final output of the whole CNN model is two neurons representing the probability of a blocking event $p_{blocking}$ and the probability of a non-blocking event $p_{nonblocking}$, which is technically a classification problem. If $p_{blocking} > 0.5$, the CNN predicts the future event to be a blocking event (Class 1); otherwise, the prediction is a non-blocking event (Class 0).

The hyperparameters are reported in Table 1. To train the model, we used a single NVIDIA Tesla V100-PCIE-16GB, accessed via the Purdue Gilbreth High-Performance Computing Clusters.

Hyperparameters	Value
Total Epochs	100
Batch size	256
Optimizer	SGD
Learning rate	10^{-2}
Patience on plateau	10
Learning rate decay factor	0.2
Momentum	0.9

Table 1. Training Hyperparameters

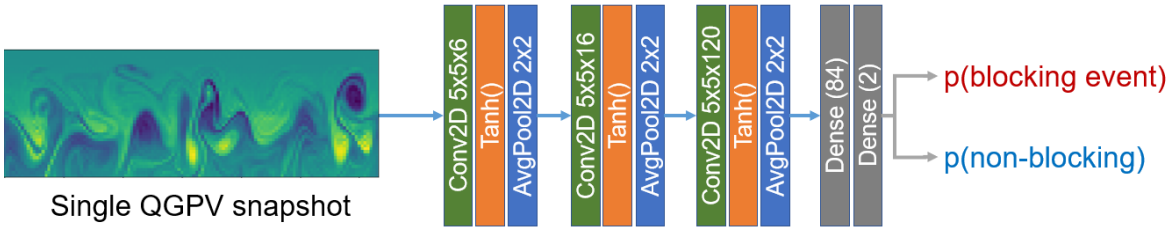

Figure 3. CNN Model Architecture with input as a single-channel QGPV snapshot preceding blocking event. The input snapshot is normalized between -1 and 1. The outputs are two nodes representing probabilities of blocking and non-blocking events, respectively.

4.2. Physical Interpretation with XAI

We freeze CNN's weights and deploy the XGradCAM (aXiom-based Gradient-weighted Class Activation Mapping) (Fu et al., 2020), an improved version of Grad-CAM (Selvaraju et al., 2017), to highlight specific patterns CNN used for blocking event predictions. The physical consistency of XAI's outputs is shown and discussed in Section 5.2.

Suppose the last convolutional layer of the CNN is $\mathbf{A}$ with K feature maps $\mathbf{A^1}$, $\mathbf{A^2}$, ..., $\mathbf{A^K}$ (depth = K). The procedure of the Grad-CAM algorithm is as follows:

1. A snapshot $\mathbf{X}$ is input to the CNN, and the final prediction is class c ("blocking" or "non-blocking").

2. Suppose y^c is the class score of predicted class c and Z is the number of pixels in A_k. Grad-CAM computes the gradients of y^c with respect to each pixel in the final convolutional feature maps: $\frac{\partial y^c}{\partial A^k}$

3. The weight representing the degree of contribution to the final prediction c of feature map k^{th} in $\mathbf{A}$ is:

$$\alpha_k^c = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A^k} \text{ (this is Global Average Pooling)} \tag{2}$$

4. Compute coarse weighted gradient activation map using weights α_k^c:

$$L^c_{coarse} = ReLU(\sum_k \alpha_k^c A^k) \tag{3}$$

   The main purpose of ReLU is to highlight only the features that contribute positively to CNN's predictions, suppressing features with negative contributions.

5. Since the resolution of the final map A^k is much smaller than the original input X due to the feature extraction mechanism of convolutional layers, Grad-CAM interpolates L^c to the size of X:

$$L^c_{final} = bilinear(L^c_{coarse}, X) \tag{4}$$

One drawback of Grad-CAM class algorithms is the coarser resolution of interpreted patterns without fine-grain details. However, one advantage is that the heatmap is much less noisy, facilitating straightforward verification of the CNN's physical consistency.

To implement XGrad-CAM for the discovery of atmospheric blocking patterns, we use the Torch-CAM package (Fernandez, 2020).

5. Experimental Results

5.1. Predictability of Atmospheric Blocking

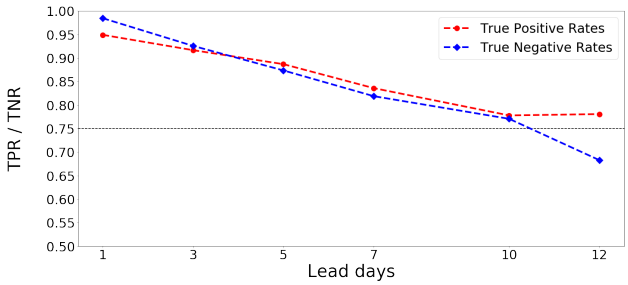

Figure 4. CNN's atmospheric blocking predictability by Number of Lead Days, measured in True Positive Rate and True Negative Rate. The predictability decreases with longer Lead days.

The initial stage in hypothesizing the presence of precursor patterns for atmospheric blocking involves investigating the model's predictability. If the model's blocking predictability is high, it signals the utilization of some precursor patterns within the model.

Figure 4 shows the CNN's predictability of blocking events for different lead days. At 1 lead day, CNN correctly predicts 95% of all blocking events and 98% of non-blocking events. Most significantly, its True Positive Rate remains above 75% even at 12 lead days. The high predictability of CNN based on a single input QGPV snapshot indicates the existence of some underlying precursor patterns. To further ensure that these patterns are signals rather than noise, we use XAI to investigate their physical fidelity.

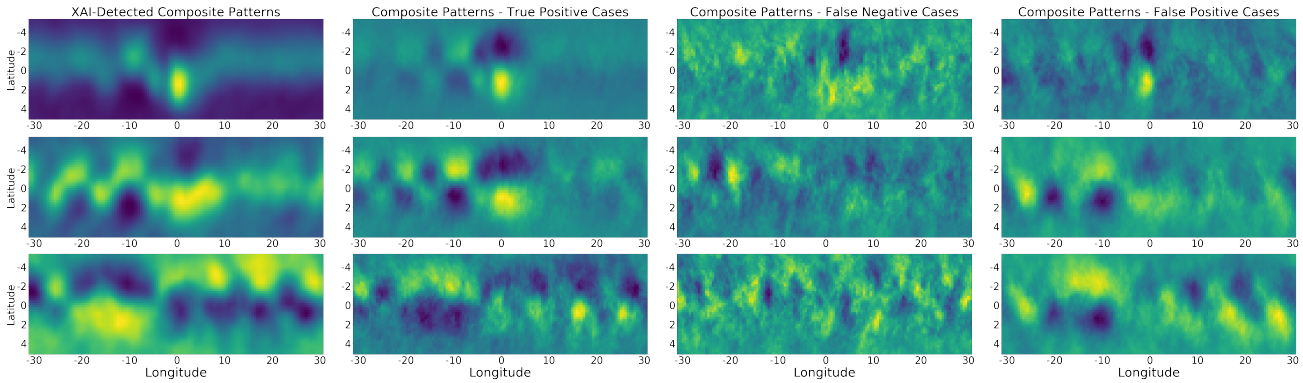

Figure 5. Composite Precursor Patterns Detected by XGrad-CAM and Ground-Truth Composite Patterns at 1, 5, and 12 lead days (each row from top to bottom order). For each lead day, there are 4 graphs from left to right: **(1)** Composite Pattern highlighted by XGrad-CAM. The brightness indicates a higher relevance to blockings; **(2)** Ground-Truth Composite (GT) QGPV Pattern of True Positive Cases; **(3)** GT QGPV Pattern of False Negative Cases; and **(4)** GT QGPV Pattern of False Positive Cases.

5.2. Physical Consistency of Precursor Patterns

A pattern, in the context of our study, refers to a clustered spatial configuration that exhibits visual distinctiveness from the background (for e.g: very bright or very dark block of pixels). We use XGrad-CAM to detect spatial QGPV patterns that CNN leverages for blocking predictions. Examining whether these patterns are spatially coherent with our Earth system understanding indicates whether such patterns are signals rather than noise. We compare the composite XAI-detected patterns with the composite ground-truth observations. The k-lead-day composite pattern is simply defined as the average of all QGPV snapshots that precede atmospheric blocking by k day(s):

$$M^k_{composite} = \frac{1}{N} \sum_{allM} M^k_i$$

The motivation for using composite instead of case-by-case patterns is that case-by-case precursor patterns are unknown and thus cannot be validated. On the other hand, composite patterns, shown in Figure 5, reflect our current knowledge of where precursor patterns of atmospheric blocking generally form. For the patterns to be considered physically consistent, there are 3 pattern-matching criteria that need to be met:

1. Composite XAI-explained precursor patterns should reflect at least an overlapping area with Composite Ground Truth patterns True Positives cases.
2. Composite Ground Truth patterns of False Negative cases should show no patterns (random noise), explaining why CNN fails to predict blocking events.
3. Composite XAI-explained precursor pattern should have at least one overlap with Ground-Truth Composite patterns of False Positive cases, explaining the source of confusion for false positive predictions.

As shown in Figure 5, XAI-detected composite precursor patterns, highlighted in bright clustered pixels, exhibit a high degree of overlapping patterns with the GT composites of True Positive cases and of False Positive cases at all lead days, satisfying criteria (1) and (3). Furthermore, the GT composite of False Negative cases displays random noises without any clear pattern (criteria (2)), which can be the cause of CNN's incorrect prediction as non-blocking events. The satisfaction of the 3 aforementioned criteria validates the physical consistency of detected precursor patterns, supporting our hypothesis that precursor patterns of atmospheric blocking do exist.

6. Conclusions and Future Work

In this paper, we have presented a synergistic approach between Two-Layer Quasi-Geostrophic Model and Explainable Convolutional Neural Network to form a hypothesis regarding the existence of atmospheric blocking's precursor patterns. The two-Layer QG model represents an idealized condition to extract and capture the most important signals for atmospheric blocking, while CNN is trained to capture such precursor patterns. The high performance of CNN indicates the presence of precursor patterns, while the composite patterns detected by XAI validate the physical consistency of such patterns. These two layers of validation provide a ground for the hypothesis that atmospheric blocking precursor patterns do exist.

In future work, we will develop causal inference models to further identify the causal relationship between such precursor patterns and blocking events. Based on the case-by-case patterns detected in this study, we will also focus on fundamental studies to derive the PDE governing the dynamical evolution of heatwaves, which is strongly linked to atmospheric blocking.

Broader impact

This paper employs a synergistic approach by integrating an idealized physical model and Explainable Machine Learning (XAI) to explore the existence of precursor patterns of blocking, taking a step closer to an open question in Earth and Atmospheric Science: do precursor patterns of blocking exist? This hypothesis about the existence of such patterns serves as a foundational contribution in guiding future research directions on dynamic evolution of large-scale extreme weather events, such as heatwaves. Such studies will improve our understanding of how heatwaves evolve, allowing early warning for informed decision-making that can save lives and minimize societal costs. Furthermore, the results from this paper are a proof of concept that Physics-Informed Explainable Machine Learning can be efficiently applied to form novel hypotheses, accelerating research in the field of Earth and Atmospheric Science.

References

Bao, M., Tan, X., Hartmann, D. L., and Ceppi, P. Classifying the tropospheric precursor patterns of sudden stratospheric warmings. *Geophysical Research Letters*, 44(15):8011–8016, August 2017. doi: 10.1002/2017gl074611. URL https://doi.org/10.1002/2017gl074611.

Barnes, E. A., Toms, B., Hurrell, J. W., Ebert-Uphoff, I., Anderson, C., and Anderson, D. Indicator patterns of forced change learned by an artificial neural network. *Journal of Advances in Modeling Earth Systems*, 12(9), September 2020. doi: 10.1029/2020ms002195. URL https://doi.org/10.1029/2020ms002195.

Davenport, F. V. and Diffenbaugh, N. S. Using machine learning to analyze physical causes of climate change: A case study of u.s. midwest extreme precipitation. *Geophysical Research Letters*, 48(15), July 2021. doi: 10.1029/2021gl093787. URL https://doi.org/10.1029/2021gl093787.

Delaunay, A. and Christensen, H. M. Interpretable deep learning for probabilistic MJO prediction. *Geophysical Research Letters*, 49(16), August 2022. doi: 10.1029/2022gl098566. URL https://doi.org/10.1029/2022gl098566.

Dole, R., Hoerling, M., Perlwitz, J., Eischeid, J., Pegion, P., Zhang, T., Quan, X.-W., Xu, T., and Murray, D. Was there a basis for anticipating the 2010 russian heat wave? *Geophysical Research Letters*, 38(6):n/a–n/a, March 2011. doi: 10.1029/2010gl046582. URL https://doi.org/10.1029/2010gl046582.

Dong, L., Mitra, C., Greer, S., and Burt, E. The dynamical linkage of atmospheric blocking to drought, heatwave and urban heat island in southeastern US: A multi-scale case study. *Atmosphere*, 9(1):33, January 2018. doi: 10.3390/atmos9010033. URL https://doi.org/10.3390/atmos9010033.

Fernandez, F.-G. Torchcam: class activation explorer. https://github.com/frgfm/torch-cam, March 2020.

Fu, R., Hu, Q., Dong, X., Guo, Y., Gao, Y., and Li, B. Axiom-based grad-cam: Towards accurate visualization and explanation of cnns, 2020.

Kautz, L.-A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., and Woollings, T. Atmospheric blocking and weather extremes over the euro-atlantic sector – a review. *Weather and Climate Dynamics*, 3(1):305–336, 2022. doi: 10.5194/wcd-3-305-2022. URL https://wcd.copernicus.org/articles/3/305/2022/.

Kondylatos, S., Prapas, I., Ronco, M., Papoutsis, I., Camps-Valls, G., Piles, M., Fernández-Torres, M.-Á., and Carvalhais, N. Wildfire danger prediction and understanding with deep learning. *Geophysical Research Letters*, 49 (17), September 2022. doi: 10.1029/2022gl099368. URL https://doi.org/10.1029/2022gl099368.

Labe, Z. M. and Barnes, E. A. Predicting slowdowns in decadal climate warming trends with explainable neural networks. *Geophysical Research Letters*, 49(9), May 2022. doi: 10.1029/2022gl098173. URL https://doi.org/10.1029/2022gl098173.

Lutsko, N. The Two Climates of the Two-Layer QG Model — nicklutsko.github.io. https://nicklutsko.github.io/blog/2018/06/28/Two-Climates-of-the-Two-Layer-QG-Model, 2018. [Accessed 24-May-2023].

Marengo, J. A., Ambrizzi, T., Barreto, N., Cunha, A. P., Ramos, A. M., Skansi, M., Carpio, J. M., and Salinas, R. The heat wave of october 2020 in central south america. *International Journal of Climatology*, 42(4):2281–2298, September 2021. doi: 10.1002/joc.7365. URL https://doi.org/10.1002/joc.7365.

Martin, Z., Barnes, E., and Maloney, E. Predicting the MJO using interpretable machine-learning models. March 2021. doi: 10.1002/essoar.10506356.1. URL https://doi.org/10.1002/essoar.10506356.1.

Martius, O., Polvani, L. M., and Davies, H. C. Blocking precursors to stratospheric sudden warming events. *Geophysical Research Letters*, 36(14), July 2009. doi: 10.1029/2009gl038776. URL https://doi.org/10.1029/2009gl038776.

Rader, J. K., Barnes, E. A., Ebert-Uphoff, I., and Anderson, C. Detection of forced change within combined climate fields using explainable neural networks. *Journal of Advances in Modeling Earth Systems*, 14 (7), July 2022. doi: 10.1029/2021ms002941. URL https://doi.org/10.1029/2021ms002941.

Robinson, P. J. On the definition of a heat wave. *Journal of Applied Meteorology*, 40(4):762–775, April 2001. doi: 10.1175/1520-0450(2001)040⟨0762:otdoah⟩2.0.co;2. URL https://doi.org/10.1175/1520-0450(2001)040<0762:otdoah>2.0.co;2.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626, 2017. doi: 10.1109/ICCV.2017.74.

Tan, J., Zheng, Y., Song, G., Kalkstein, L. S., Kalkstein, A. J., and Tang, X. Heat wave impacts on mortality in shanghai, 1998 and 2003. *International Journal of Biometeorology*, 51(3):193–200, October 2006. doi: 10.1007/s00484-006-0058-3. URL https://doi.org/10.1007/s00484-006-0058-3.

Toms, B. A., Barnes, E. A., and Ebert-Uphoff, I. Physically interpretable neural networks for the geosciences: Applications to earth system variability. *Journal of Advances in Modeling Earth Systems*, 12(9), August 2020. doi: 10.1029/2019ms002002. URL https://doi.org/10.1029/2019ms002002.

Zschenderlein, P., Pfahl, S., Wernli, H., and Fink, A. H. A lagrangian analysis of upper-tropospheric anticyclones associated with heat waves in europe. *Weather and Climate Dynamics*, 1(1):191–206, April 2020. doi: 10.5194/wcd-1-191-2020. URL https://doi.org/10.5194/wcd-1-191-2020.