XAI4Extremes: An interpretable machine learning framework for understanding extreme-weather precursors under climate change

Jiawen Wei[1], Aniruddha Bora[2], Vivek Oommen[2], Chenyu Dong[1], Juntao Yang[3], Jeff Adie[3], Chen Chen[4], Simon See[3], George Karniadakis[2], Gianmarco Mengaldo[1]

[1]National University of Singapore [2]Brown University [3]NVIDIA AI Technology Centre, Singapore [4]Centre for Climate Research Singapore

Paper

Motivation

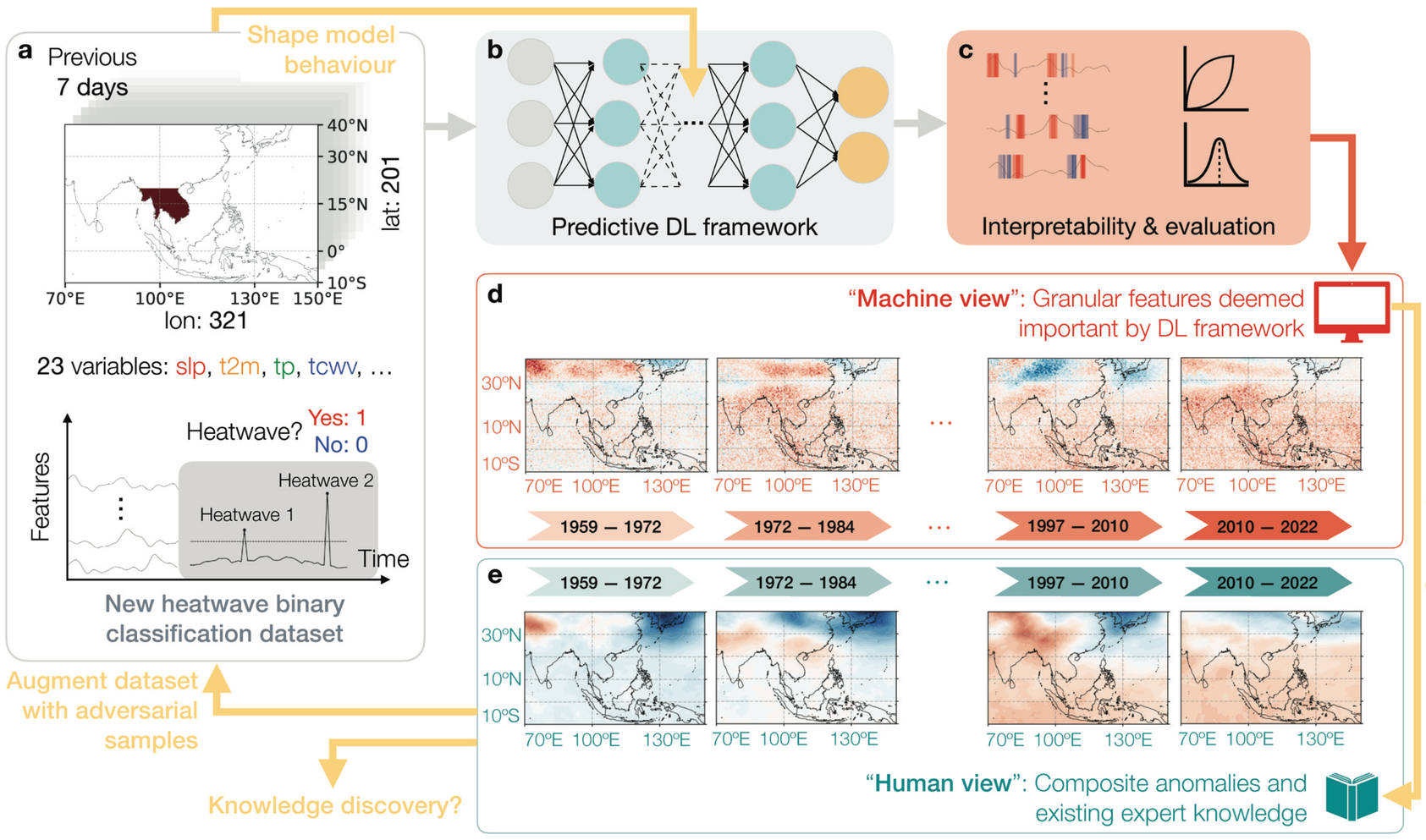
- Extreme weather events are becoming more frequent and intense due to climate change.
- Predicting extreme weather remains challenging.
- Identifying the precursors of extreme weather events and how these precursors may evolve under climate change remain unclear.
- There is a critical need to interpret AI predictions to understand the extreme-weather precursors and the role that climate change is having on these precursors

(i) What are the key precursors of extreme weather events?
(ii) Is climate change influencing these precursors?

Contributions

- We create a **novel binary classification dataset** for heatwaves in Indochina (same dataset type can be created for other weather extreme events in other regions worldwide)
- We develop the explainable AI framework, namely **XAI4Extremes**, that couples the predictive DL model with post-hoc interpretability methods to understand ***what*** data the machine deemed important for its predictive performance of true positive samples ("*machine view*").
- We propose to compare this machine view to existing human expert knowledge ("*human view*"), to responde the question ***why*** the machine used those data.

Methodology

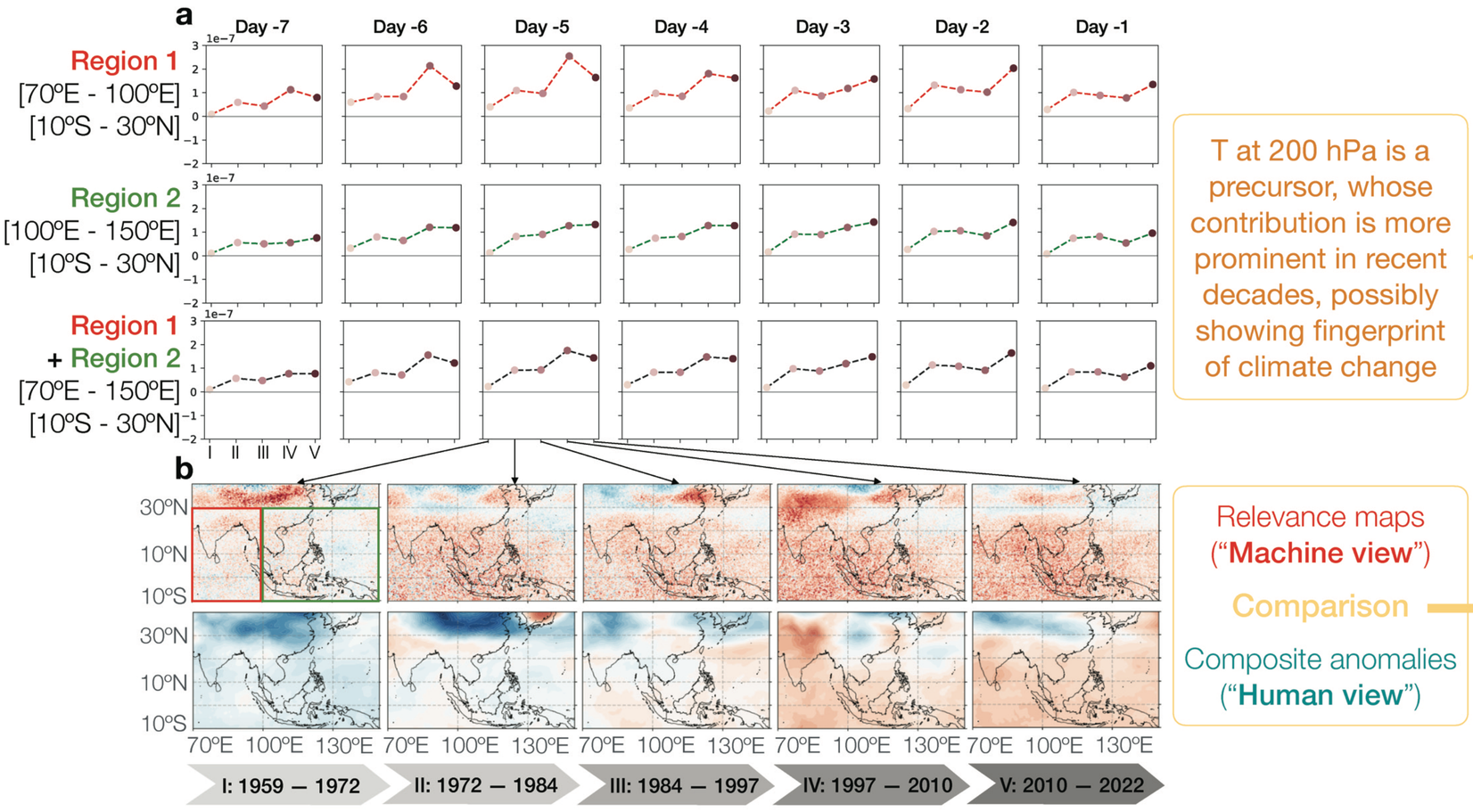


XAI4Extremes

(a) a novel extreme weather dataset
(b) a DL predictive model
(c) an interpretability block along with its evaluation
(d) relevance maps ("*machine view*")
(e) existing human expert knowledge ("*human view*")

Comparison of "machine view" and "human view" may lead to knowledge discovery in terms of heatwave precursors and role of climate change in heatwave precursors.

Results



T at 200 hPa is a precursor, whose contribution is more prominent in recent decades, possibly showing fingerprint of climate change

Relevance maps ("**Machine view**")

Comparison

Composite anomalies ("**Human view**")

- Temperature at 200 hPa (upper troposphere) is deemed more important by relevance maps ("*machine view*") for heatwaves in Indochina in more recent decades.
- This aligns with composite anomalies ("*human view*") showing a warming of the upper troposphere.

The finding indicates that the temperature at 200 hPa is becoming a key precursor of Indochina heatwaves in recent decades, possibly reflecting the fingerprint of climate change.

Conclusions: *We incorporate interpretability methods to uncover what drives model predictions ("machine view") and compare this with expert knowledge ("human view"). XAI4Extremes enriches our understanding of extreme-weather precursors and may lead to knowledge discovery, while highlighting challenges and open opportunities for the AI and broader scientific research communities.*

Wei, J., Bora, A., Oommen, V., Dong, C., Yang, J., Adie, J., Chen, C., See, S., Karniadakis, G. and Mengaldo, G., 2025. XAI4Extremes: An interpretable machine learning framework for understanding extreme-weather precursors under climate change. arXiv preprint arXiv:2503.08163.