# CLIMATE MODEL EVALUATION WITH EXPLAINABLE NEURAL NETWORKS

Zachary M. Labe[1*] & Elizabeth A. Barnes[1]

[1]Department of Atmospheric Science, Colorado State University

*Contact: zmlabe@rams.colostate.edu

@ZLabe

## The opportunity

**Global climate models are now being generated at higher and higher resolutions and thus simulating more complex Earth system interactions. To assess this vast amount of data, scientists are turning to an increasing number of advanced statistical methods, such as neural networks.**

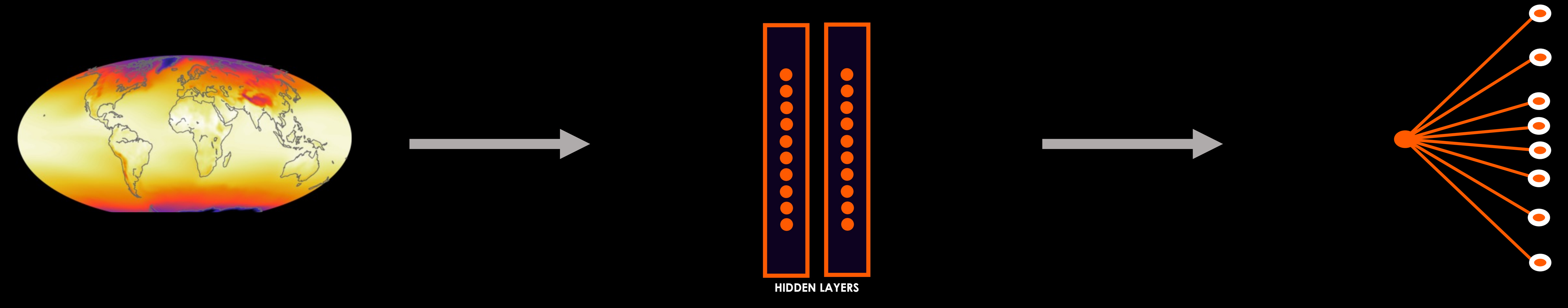
**Previous studies have found that neural networks use regional patterns to predict simple characteristics of climate data (e.g., identify the year of a map)**

**Explainable neural networks** can be used to **identify unique differences** in climate patterns simulated between global climate model large ensembles

**INPUT**
Temperature map

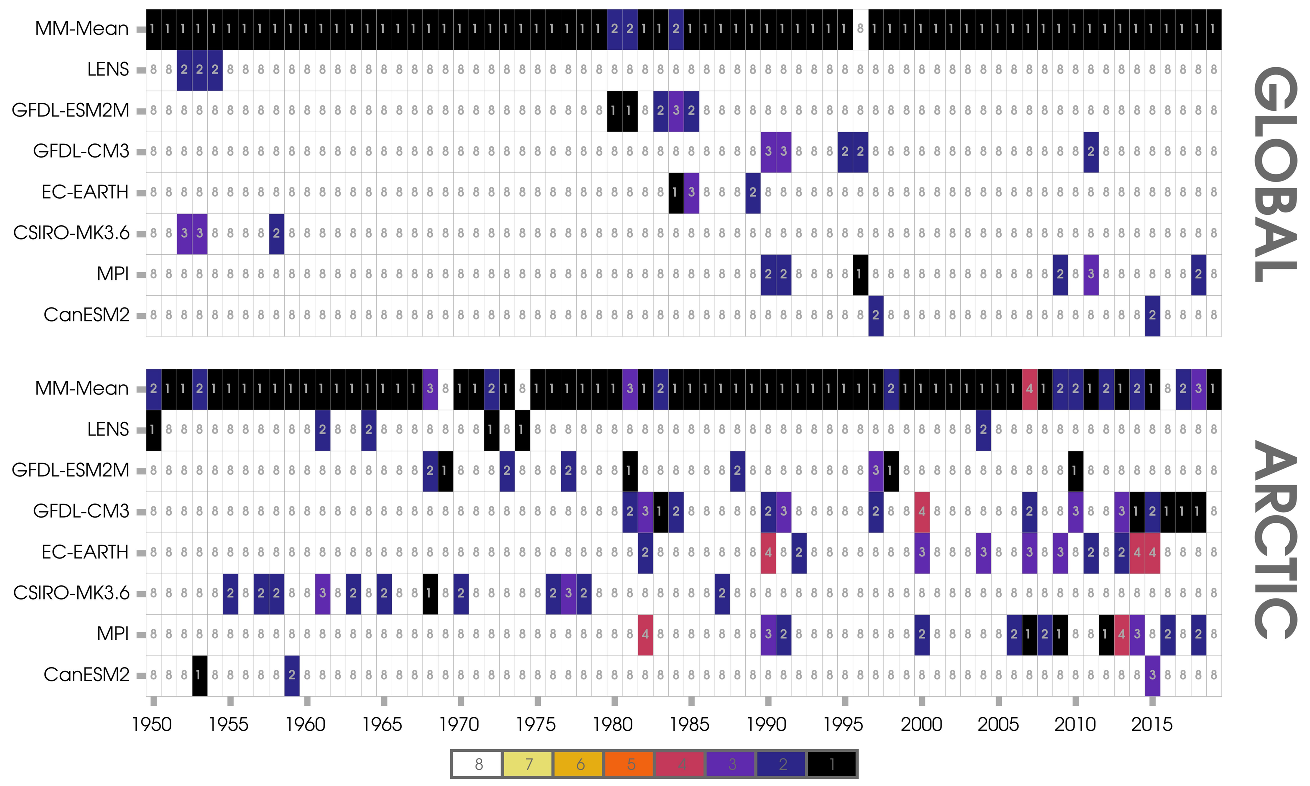
**TASK**
Classify which global climate model produced the temperature map

**OUTPUT**
Climate model label




**Train 2-layer artificial neural network on maps of temperature from climate model large ensemble data**

- **MM-Mean**
- **LENS**
- **GFDL-ESM2M**
- **GFDL-CM3**
- **EC-EARTH**
- **CSIRO-MK3.6**
- **MPI**
- **CanESM2**



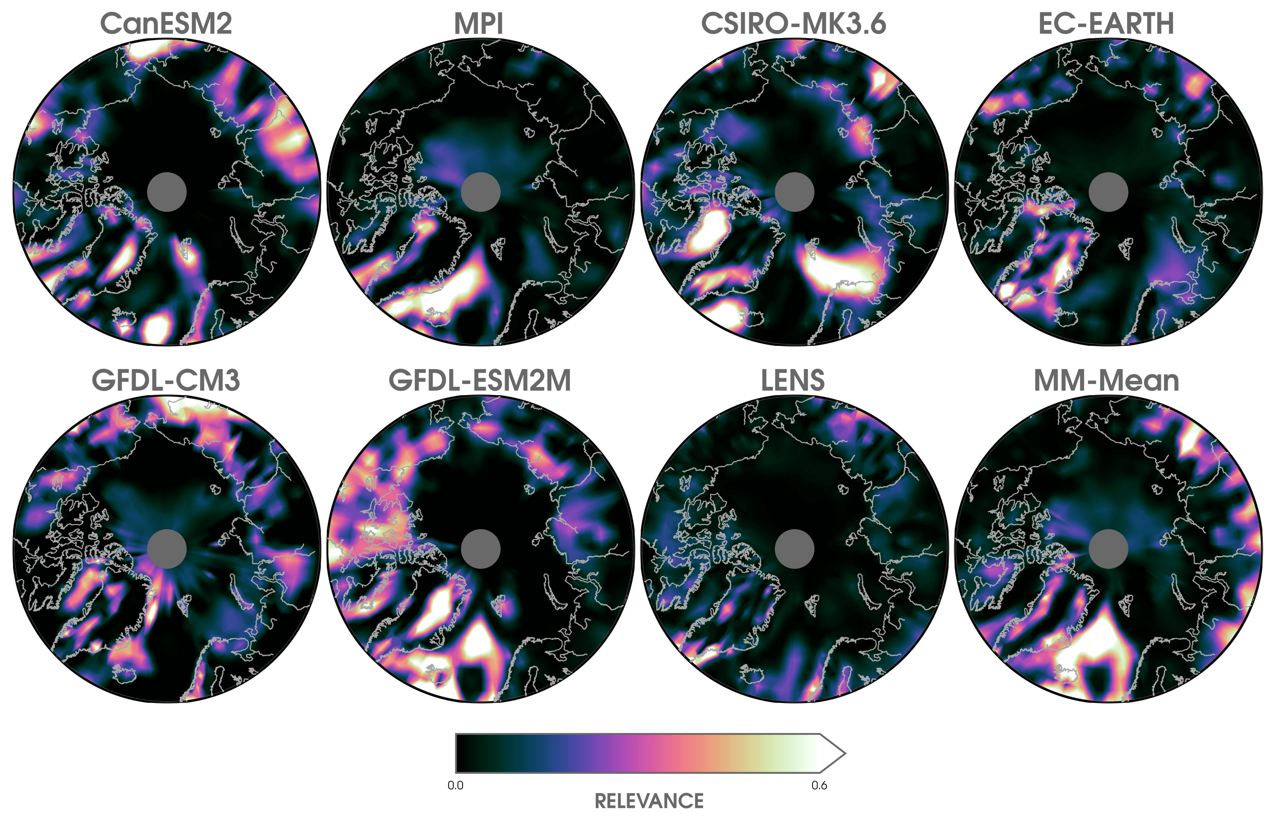

**RELATIVE MODEL CHOICE**

A softmax activation function is applied to the final layer of the neural network in order to transform the class probabilities so that they sum to one.

We then use the order of these likelihoods to rank the choice of climate model for each map after testing observations (where rank 1 = largest predicted probability).

Layer-wise relevance propagation (LRP) heatmaps for an artificial neural network that is given temperature maps for the Arctic from 1950-2019.

Each LRP map is separately composited according to the different climate model classifications from testing data. Higher LRP values indicate greater relevance for the network prediction.




## The conclusions

**After training on climate model data, we input annual-mean (global) maps from observations (ERA5-BE/20CRv3 reanalysis). The neural network mostly predicts them as the multi-model mean ensemble class for each year.**

**In addition to using other climate variables, the neural network architecture can be adapted for climate model comparison/evaluation in regions with known large biases, such as over the Arctic or the Southern Ocean**