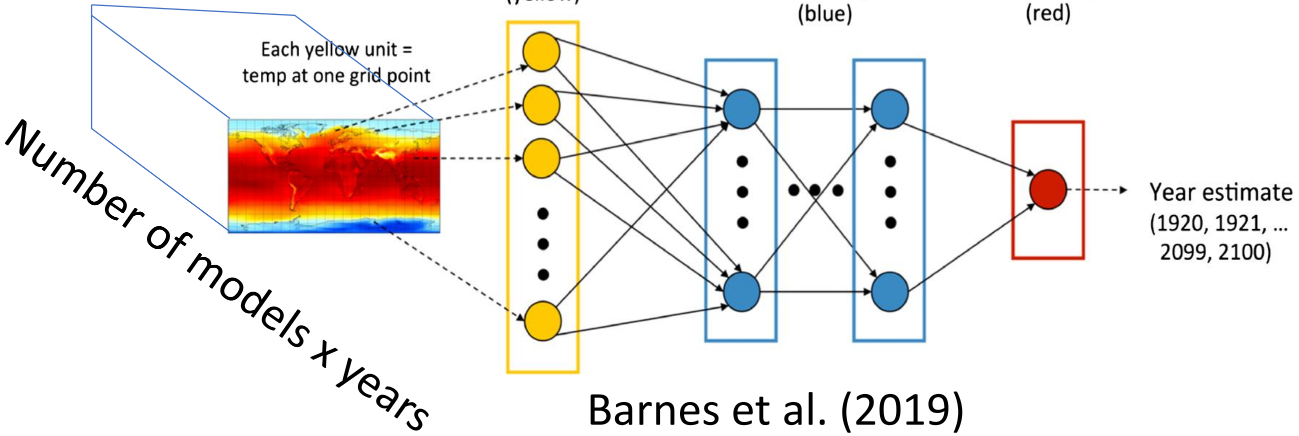
Anthropogenic influence on historical extreme precipitation over global land areas detected using an ‘explainable’ artificial neural network

Gavin D. Madakumbura
with
Chad Thackeray, Jesse Norris, Naomi Goldenson and Alex Hall

Training an Artificial Neural Network (ANN) to predict/model the forced signal of Rx1day

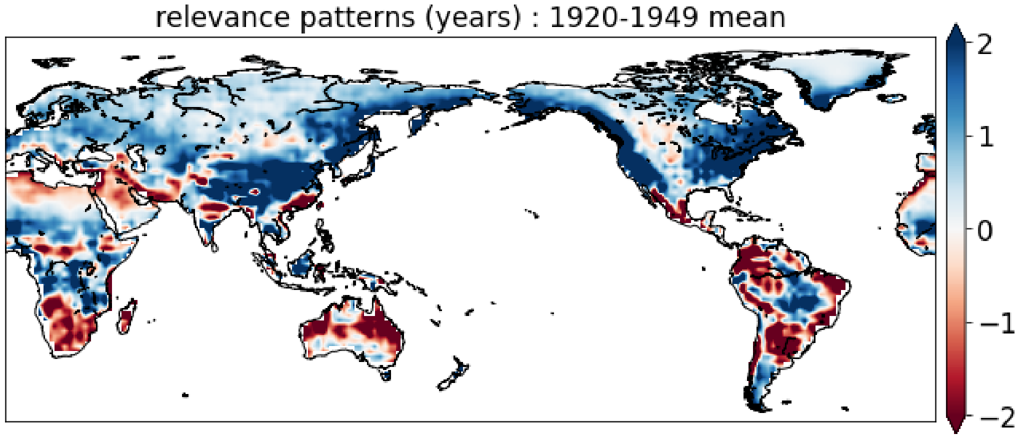
The ANN should learn to separate the forced signal from
1) Model uncertainty
2) Internal variability

Motivations/Strengths :
- Applicable for shorter observational records (unlike traditional D&A methods which are trend based and require a long data record)
- Emerging techniques to peer inside the black box allow us to obtain the ANN learned fingerprints

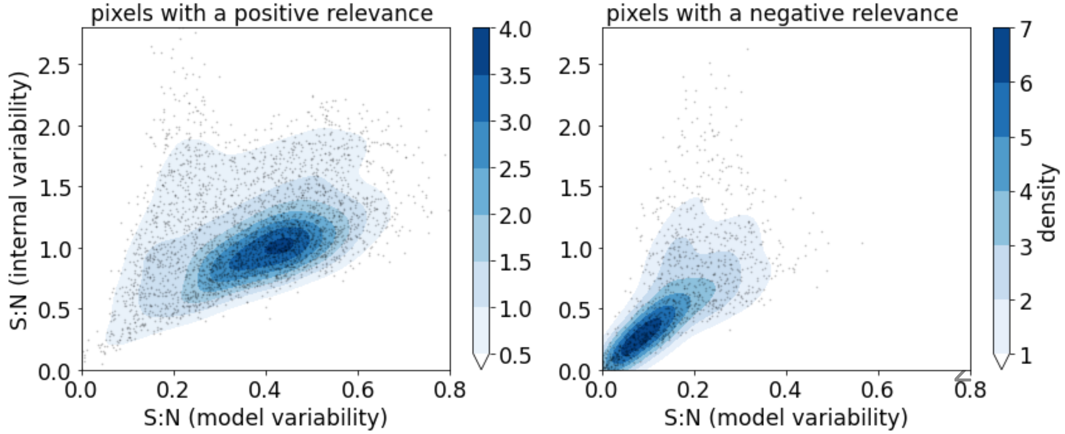


Barnes et al. (2019)

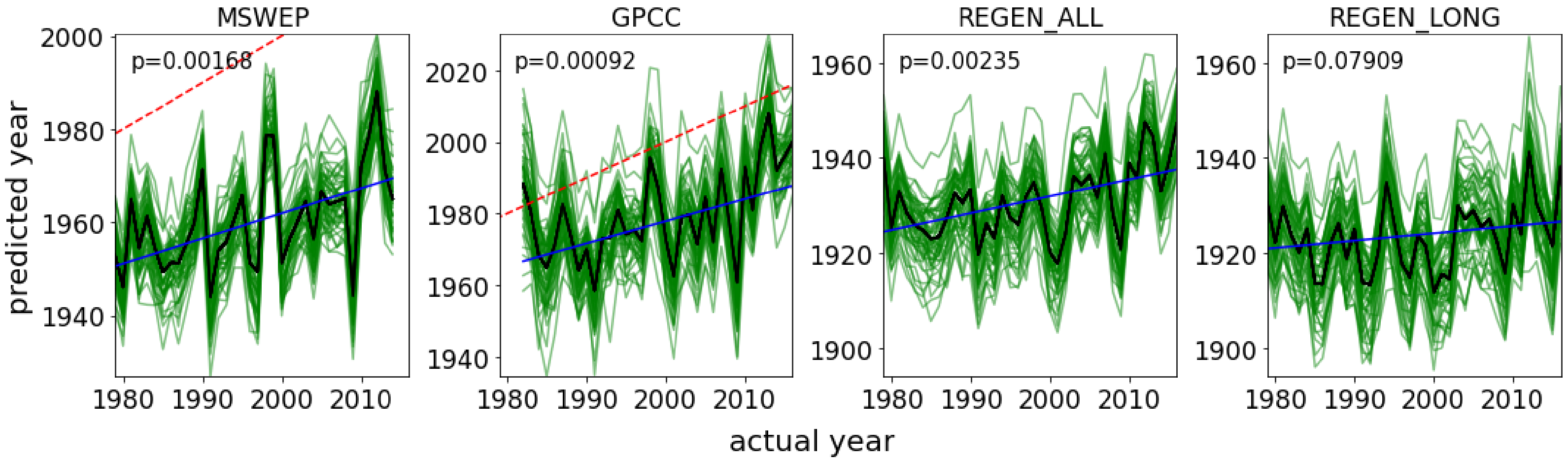
Fingerprint of external forcing in Rx1day learned by the ANN



Relevance maps (aka 'heatmaps') obtained using **Layer-wise Relevance Propagation** show regions which are positively and negatively related to the prediction (i.e. the year)



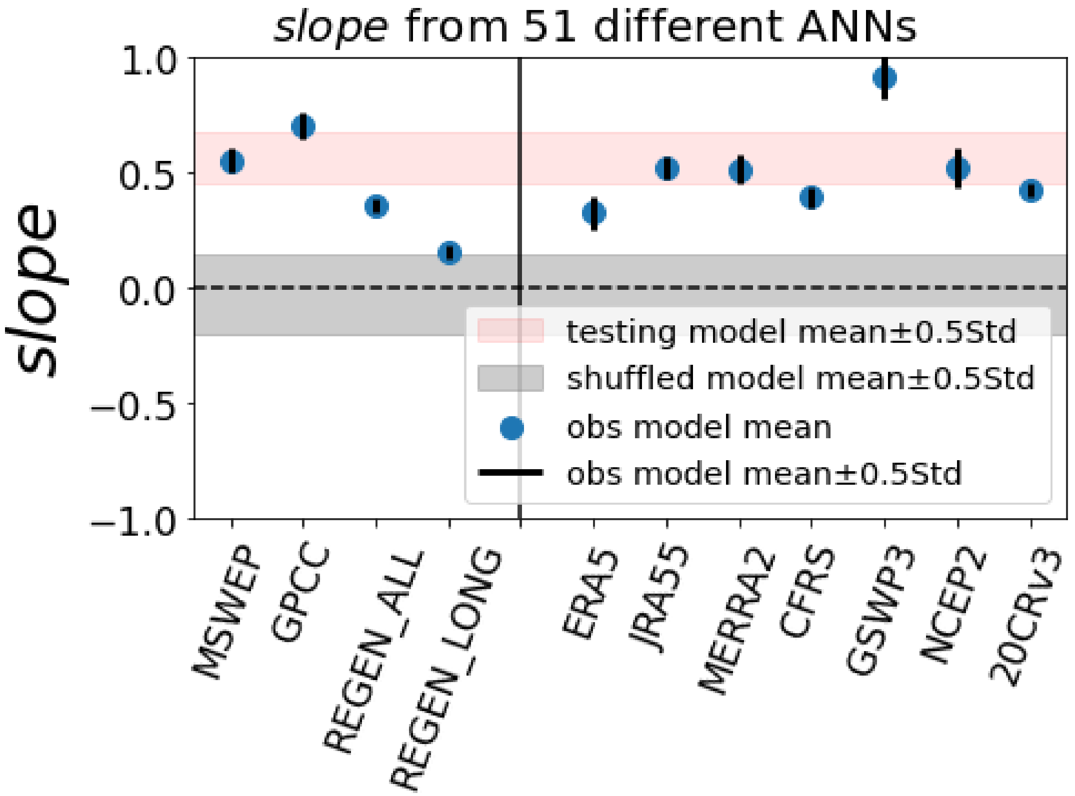
Rx1day : Forced signal in observations



Can think of this as projecting the observations on to the 'fingerprints' identified by the ANN

A metric for calculating the forced response :

- **Slope** of the regression line between predicted and actual year



Two out of four observations (MSWEP,GPCC) show a strong anthropogenic signal

A large observational uncertainty exists!

Future Directions

Short-term goals :

- Comparisons with traditional D&A methods
- Apply ANN D&A to different variables (e.g. fire-weather)
- Use interpretable ANNs to identify systematic model variability and potential emergent constraints

Long-term goals :

- Incorporate physics (i.e. physics guided ANNs)
  - A multivariate ANN D&A framework will allow us to impose physical constraints (e.g. conservation laws)

Content related to white paper

- A carefully designed ANN can be used to extract the forced response from the noise.
- ANN D&A method uses pattern-based learning, without relying on the trend.
- By applying AI interpretation techniques, we can check if the results are physically consistent or not.