Viewing Climate Signals through an AI Lens

Elizabeth A. Barnes, Associate Professor, Dept. of Atmospheric Science, CSU

Collaborators:

Benjamin Toms, PhD student, Dept. of Atmospheric Science, CSU

James W. Hurrell, Faculty, Dept. of Atmospheric Science, CSU

Imme Ebert-Uphoff, Research Faculty, CIRA and Dept of Elect. and Comp. Eng., CSU

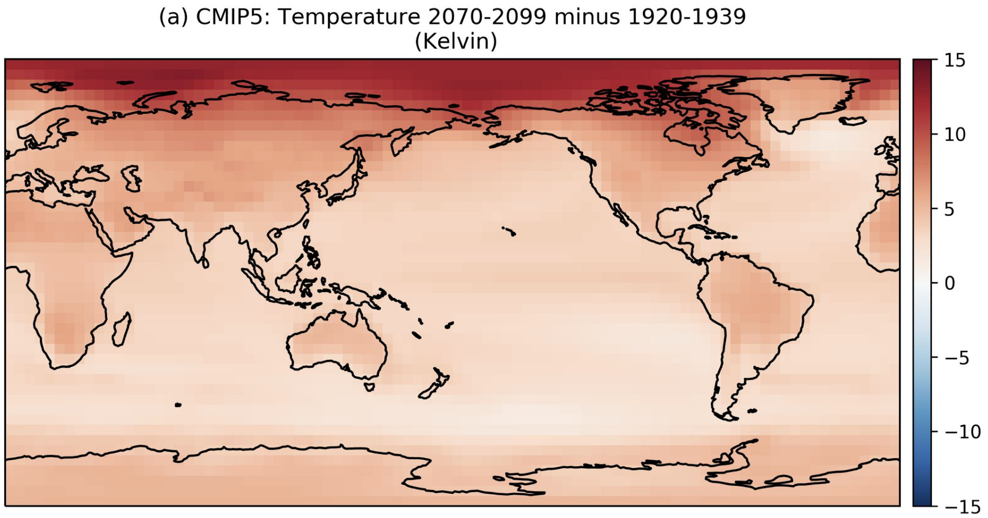
Chuck Anderson, Faculty, Computer Science, CSU

David Anderson, Pattern Exploration LLC, Fort Collins, CO

NOAA AI Conference
August 27, 2020

Climate Change in the 21st Century: a signal-to-noise problem

Surface temperature change under the RCP8.5 future climate scenario between 2070-2099 and 1920-1939 averaged over 29 different climate models

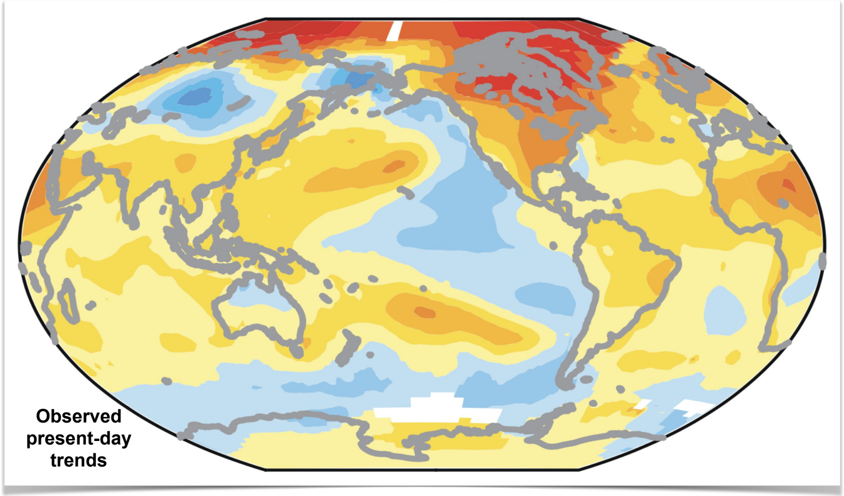


Two Sources of Uncertainty

- structural model uncertainty/disagreement (i.e. simulating the physics)
- internal variability (i.e. climate noise)

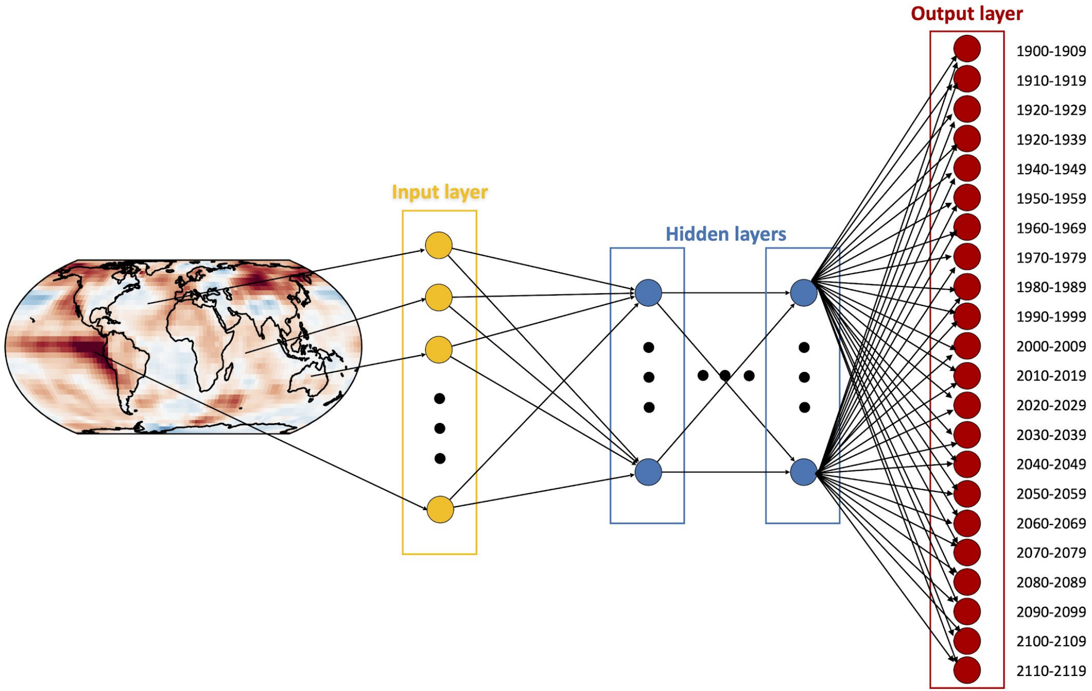
Colorado State University

Climate Change in the 21st Century: a signal-to-noise problem



How can we tell which changes are the SIGNAL and which are the NOISE in our one observed earth?

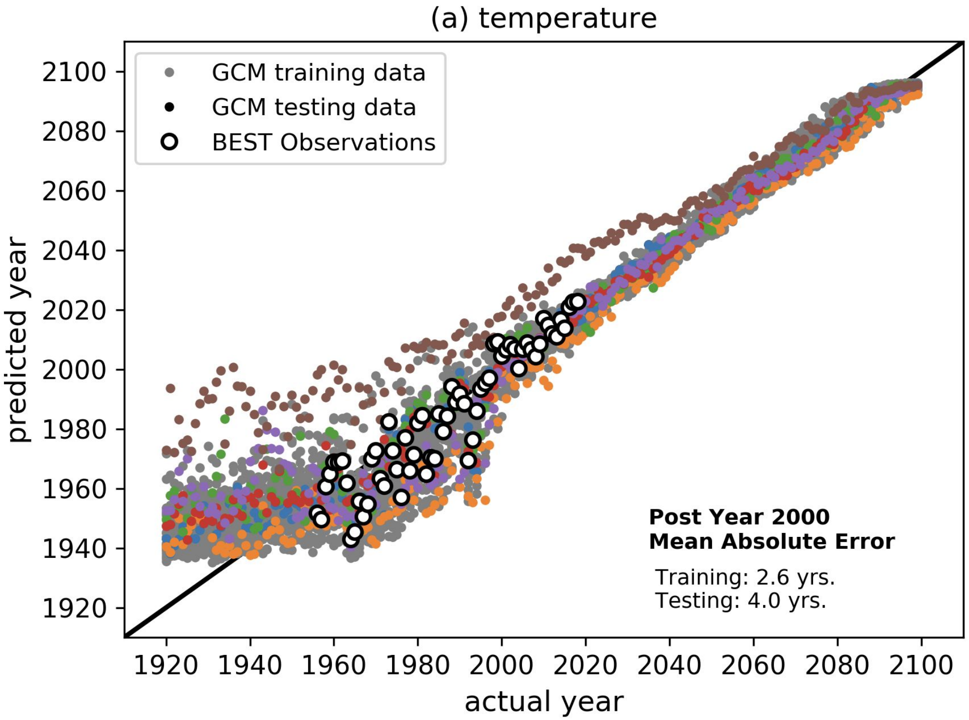
Train ANN to predict the year of a map



*Training and testing
on CMIP5 climate
model output

et al. (2019; GRL)
Barnes et al. (2020; JAMES)

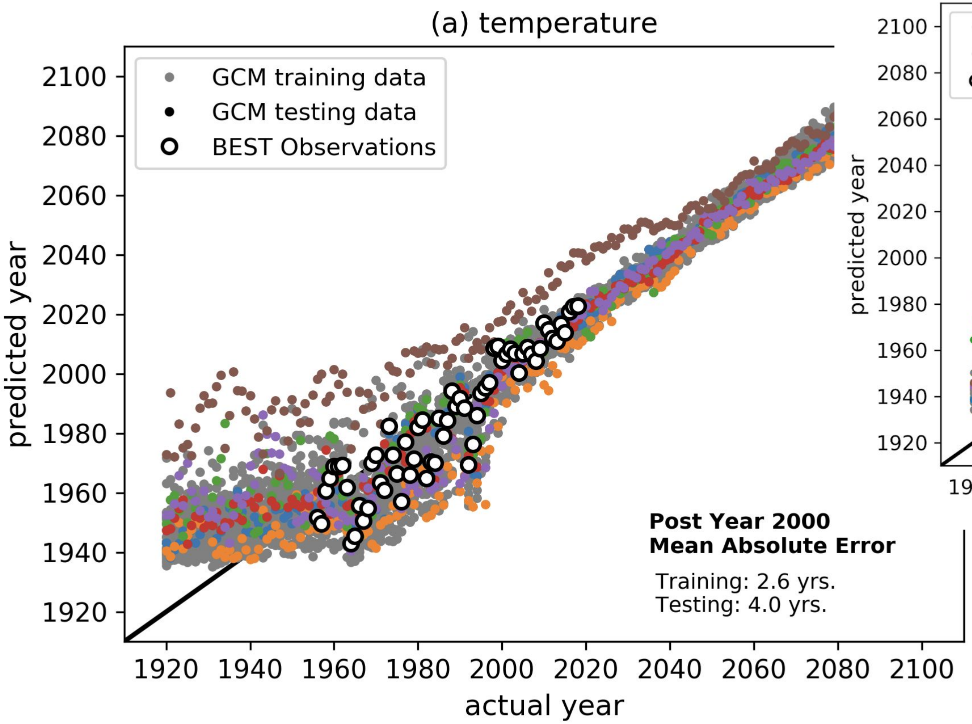
Train ANN to predict the year of a map

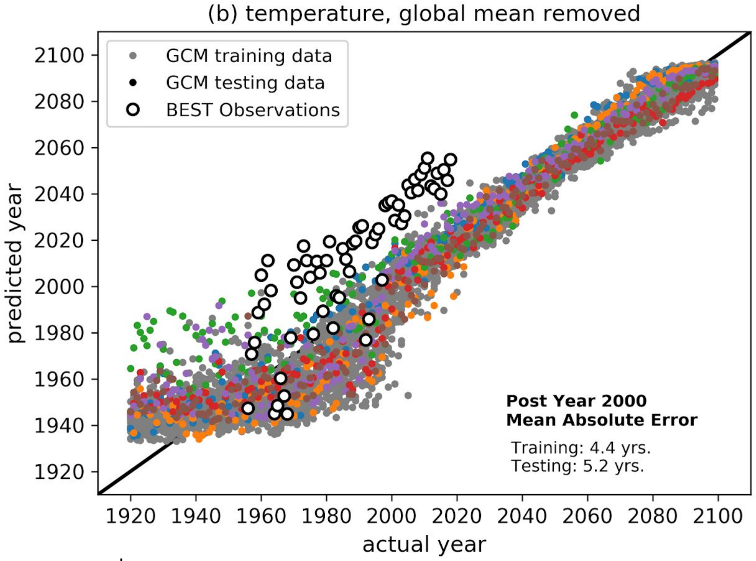


*Training and testing on CMIP5 climate model output

Barnes et al. (2019; GRL)
Barnes et al. (2020; JAMES)

Train ANN to predict the year of a map

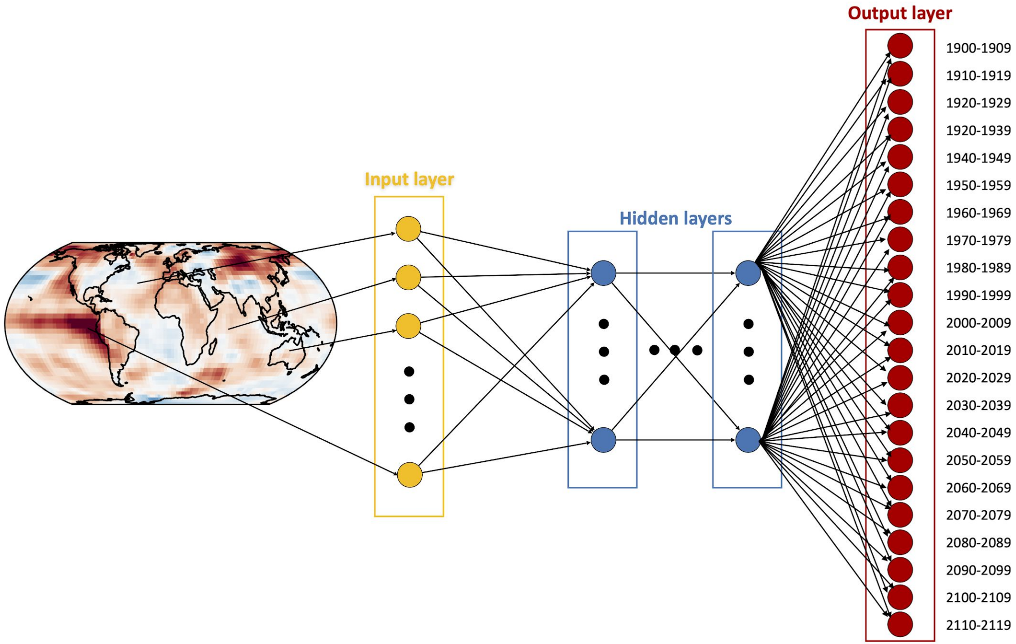




*Training and testing on CMIP5 climate model output

Barnes et al. (2019; GRL)
Barnes et al. (2020; JAMES)

What did the ANN learn?



*Training and testing on CMIP5 climate model output

ANN must learn regional signals that are "reliable" indicators of the year

Barnes et al. (2019; GRL)
Barnes et al. (2020; JAMES)

What to expect from ANN visualization

Put backpack into
X ray scanner

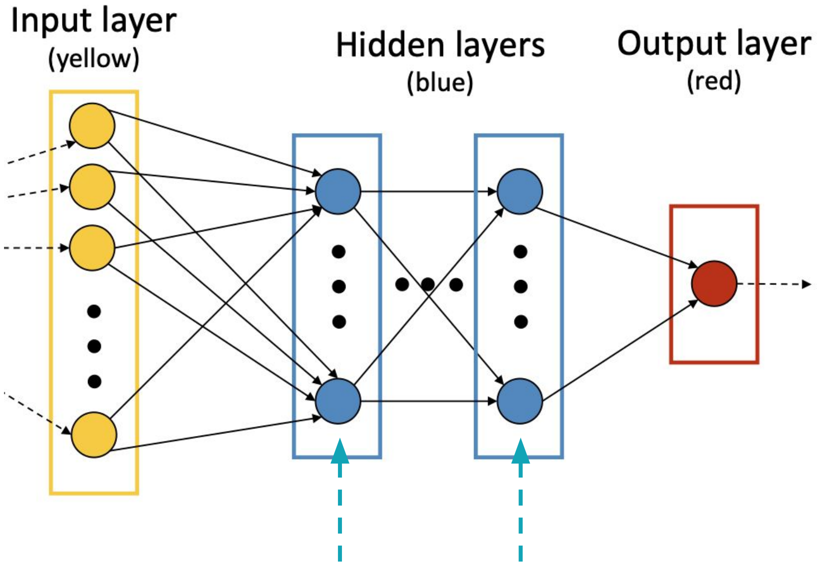
Inside
view

Not a perfect view, but better than
the "black box".

Two types of visualization tools

Type A: Feature Visualization

Philosophy: Seek to understand all internal components of ANN.

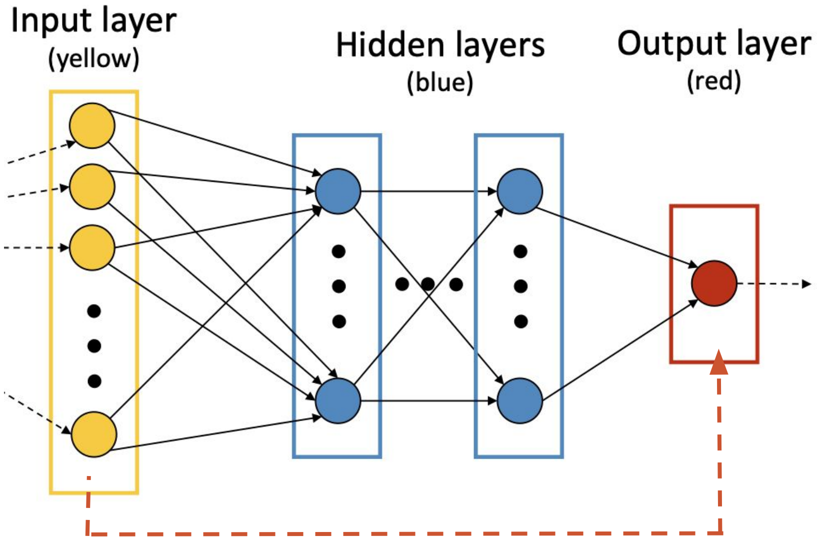


Seek to understand the **meaning of all intermediate (blue) nodes.**

Two types of visualization tools

Type B: Attribution / Explaining Decisions

Philosophy: Understand the ANN's overall decision making for specific input.

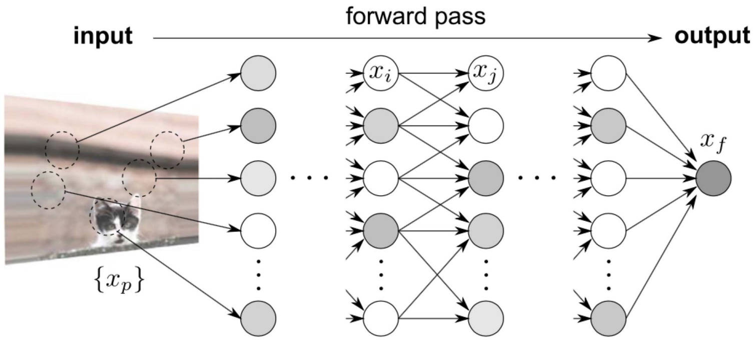


Seek to understand the meaning of the entire algorithm - for a specific input.
Do NOT worry about meaning of intermediate (blue) nodes.

A visualization tool: Layerwise Relevance Propagation

Prediction

of 1 sample



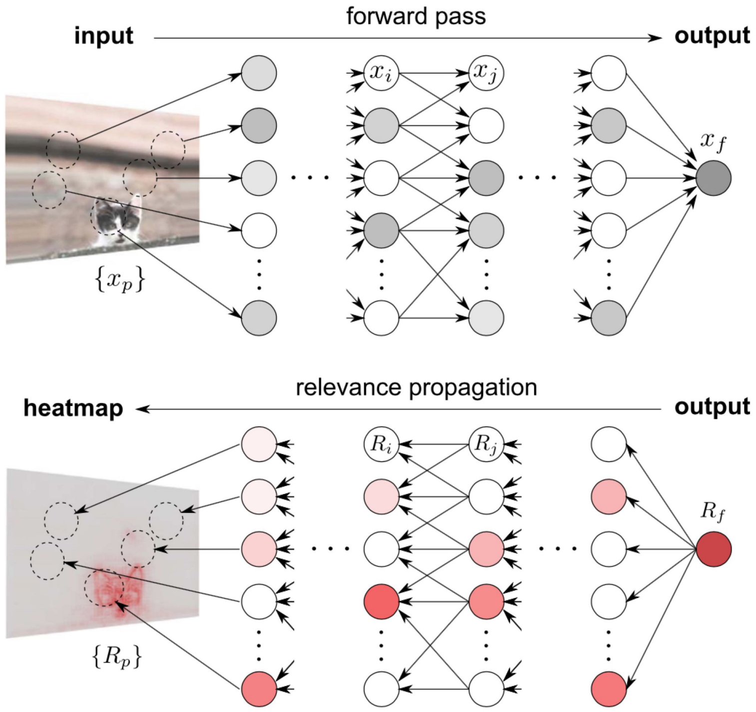
Pr(cat)

Montavon et al. (2017), Pattern Recognition; Montavon et al. (2018), Digital Signal Processing

A visualization tool: Layerwise Relevance Propagation

Prediction
of 1 sample

LRP
of 1 sample



Pr(cat)

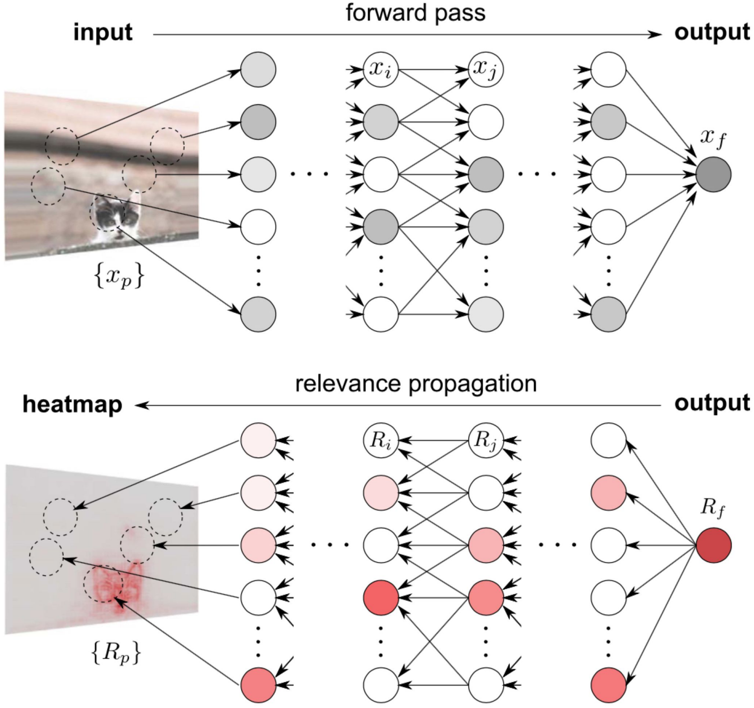
Pr(cat)

Montavon et al. (2017), Pattern Recognition; Montavon et al. (2018), Digital Signal Processing

A visualization tool: Layerwise Relevance Propagation

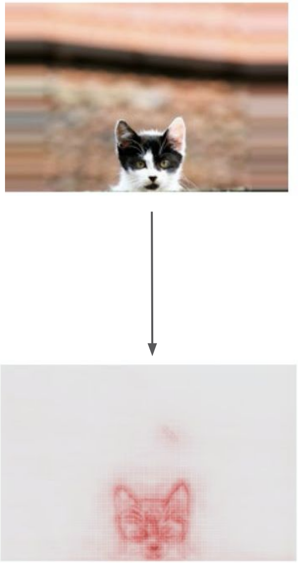
Prediction
of 1 sample

LRP
of 1 sample



Pr(cat)

Pr(cat)



where the network looked to determine it was a "cat"

Montavon et al. (2017), Pattern Recognition; Montavon et al. (2018), Digital Signal Processing

Example use of LRP

Task: Decide whether there is a horse in a given image.

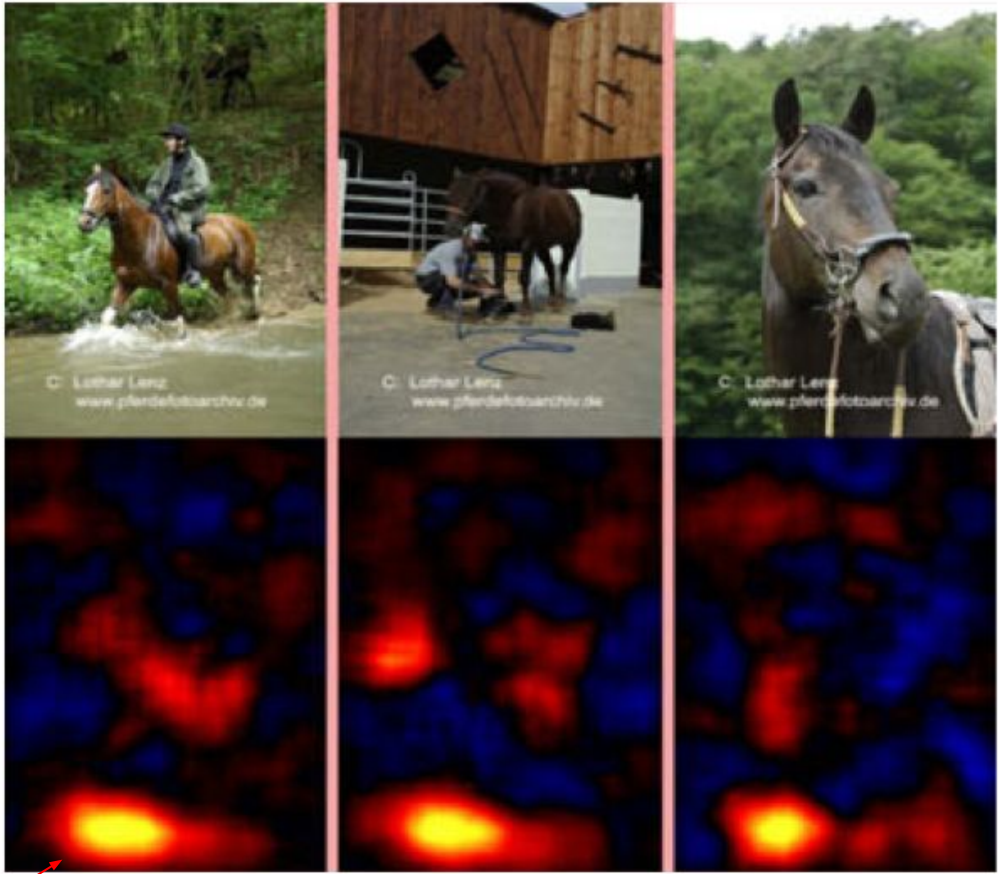
Decision making strategy: use visualization tools to determine the strategy the network used to make a decision

Lapuschkin et al. (2019)

Example use of LRP

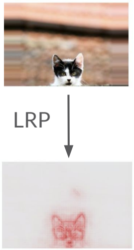
Task: Decide whether there is a horse in a given image.

Decision making strategy: use visualization tools to determine the strategy the network used to make a decision



regions relevant to the network's decision

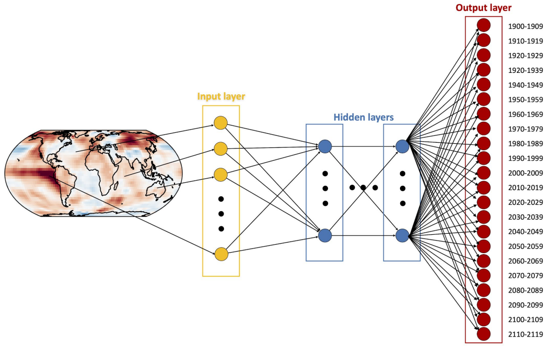
Lapuschkin et al. (2019)

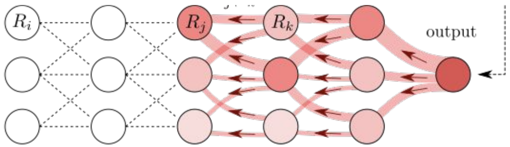


What does this mean for earth science research?

1. Identifying problematic strategies (i.e. right answer for the wrong reasons)
2. Designing the machine learning methodology
3. Building trust
4. **Discovering new science!**
    - **When** our machine learning method is capable of making a correct prediction we can explore **why**

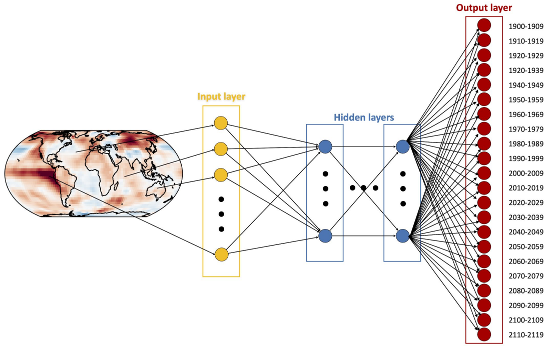
Indicators of climate change: temperature

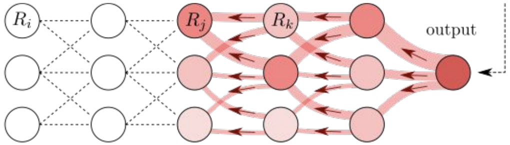




Which regions are **relevant** for correctly predicting a specific year?

Indicators of climate change: temperature

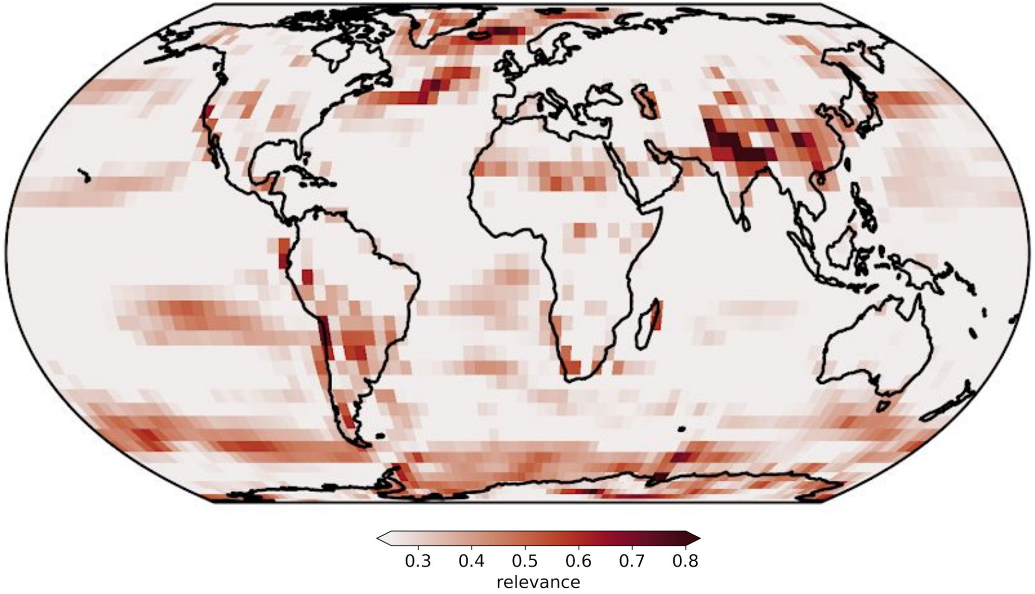




Which regions are **relevant** for correctly predicting a specific year?

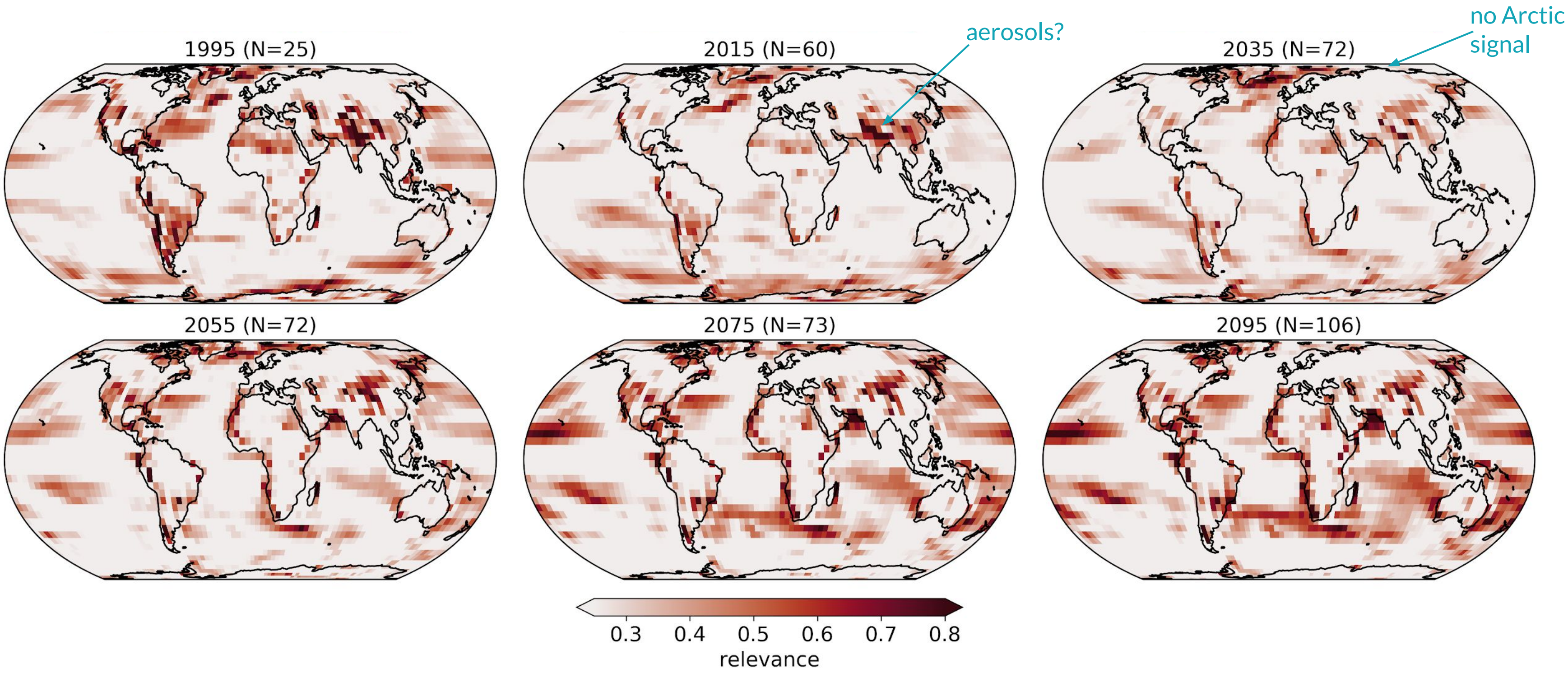
Year = 2015

Relevant Regions for Predicting Year from Temperature Map



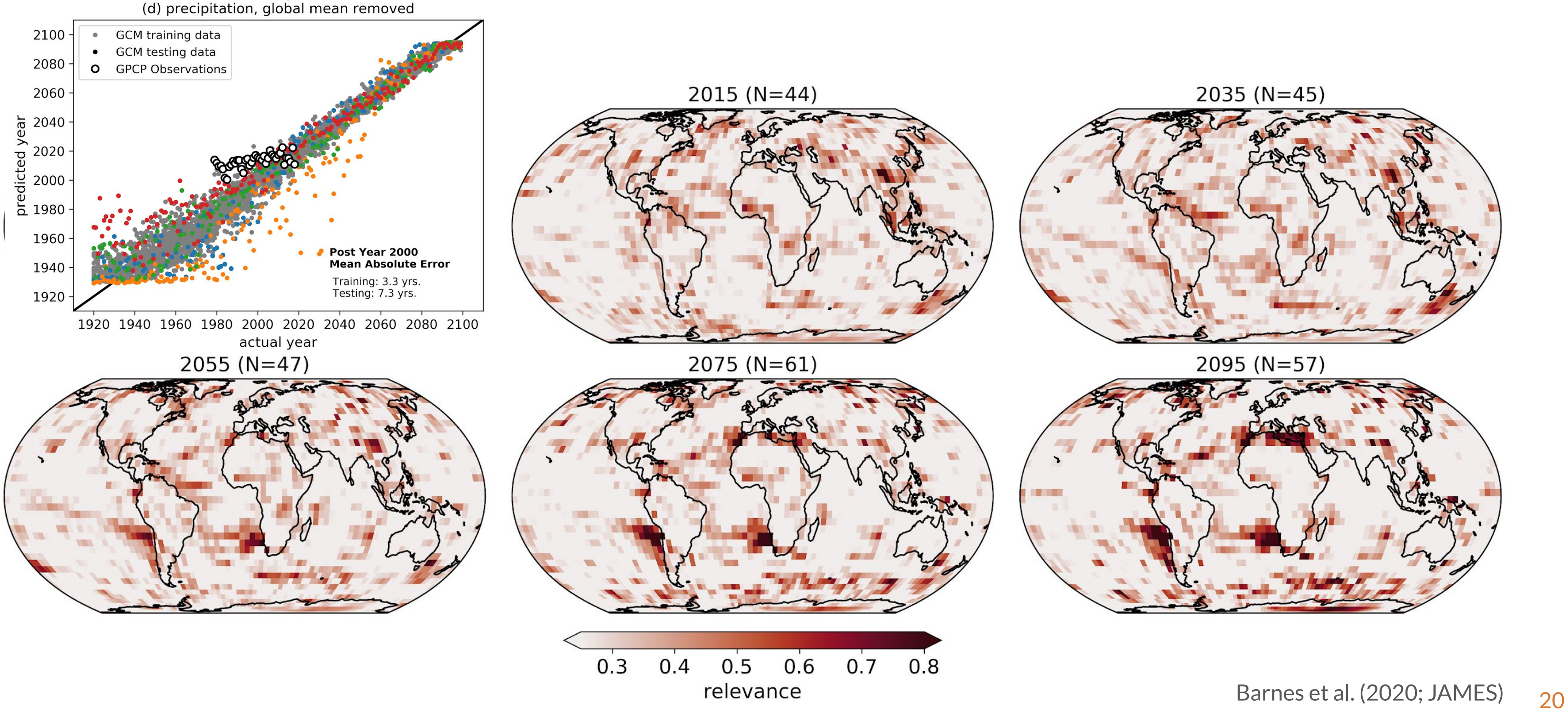
Barnes et al. (2020; JAMES)

Indicators of climate change: temperature



Barnes et al. (2020; JAMES)

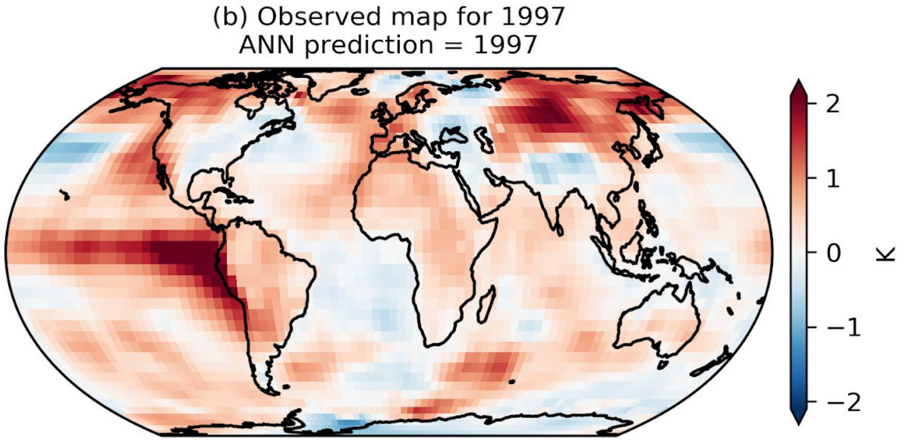
Indicators of climate change: precipitation

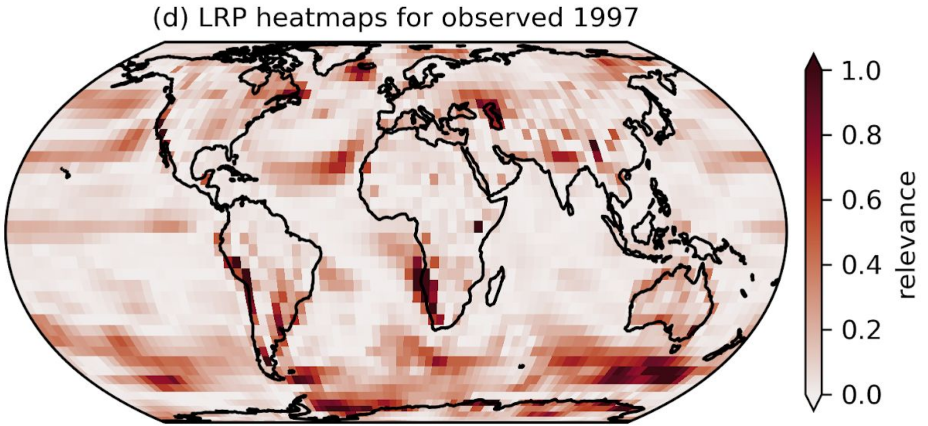


Barnes et al. (2020; JAMES)

LRP for Observations

- Largest anomalies are not necessarily the most reliable indicator regions
- ANN focuses on the Southern Ocean and the southern coasts of South America and Africa





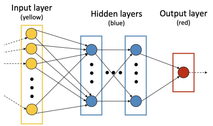
Barnes et al. (2020; JAMES)

Barnes Research Group

Our Current Projects Using LRP

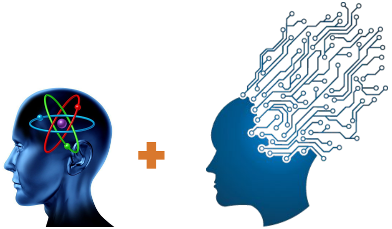
1. Indicator patterns of forced change
2. Multi-year prediction
3. Subseasonal-to-seasonal prediction
4. Eddy-mean flow interactions
5. Human impacts on the land surface from Landsat imagery

Wrap-up



- The most basic of neural networks can be viewed as nonlinear regression - **climate scientists are well-equipped** to think about this architecture

- Artificial neural networks are **no longer black boxes** - tools exist to help **visualize their decisions**. This is a **game changer** for their use in geoscience research.



- ANNs can be used for more than just prediction. The **science can be what the network learns**, rather than the prediction. **Get creative** combining your science with these tools!

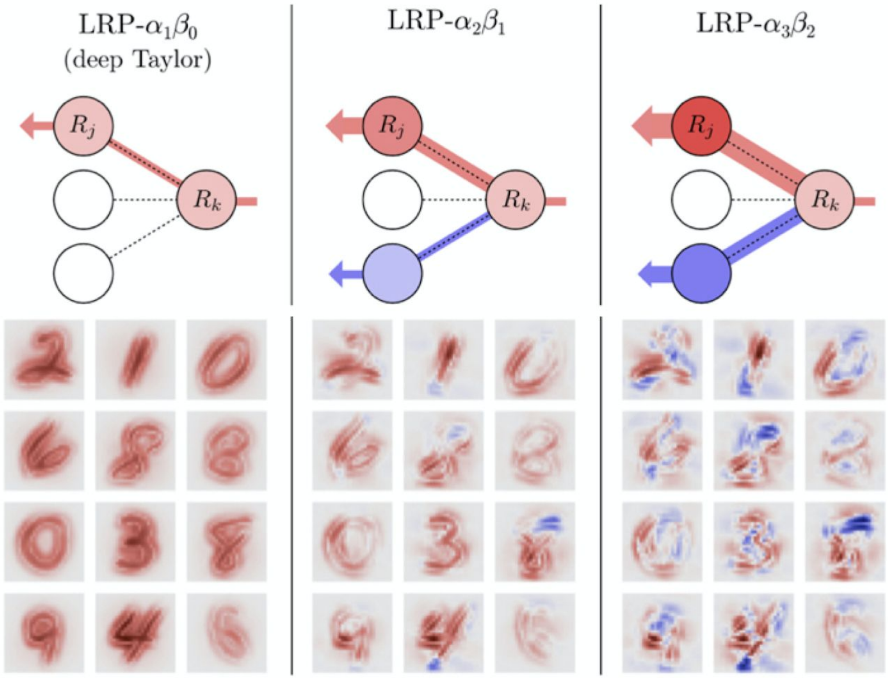
Elizabeth A. Barnes
eabarnes@rams.colostate.edu,
Twitter @atmosbarnes

CSU papers in this area

- Toms, Benjamin A., Elizabeth A. Barnes, and Imme Ebert-Uphoff: Physically interpretable neural networks for the geosciences: Applications to earth system variability, *JAMES*, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002002.
- Barnes, E. A., J. W. Hurrell, I. Ebert-Uphoff, C. Anderson and D. Anderson: Viewing forced climate patterns through an AI Lens, Geophysical Research Letters, doi.org/10.1029/2019GL084944.
- Barnes, Elizabeth A., Benjamin Toms, James Hurrell, Imme Ebert-Uphoff, Chuck Anderson and David Anderson: Indicator patterns of forced change learned by an artificial neural network, JAMES, under review, preprint available at http://arxiv.org/abs/2005.12322.
- Toms, B., K. Kashinath, Prabhat, and D. Yang (2020), Testing the Reliability of Interpretable Neural Networks in Geoscience Using the Madden-Julian Oscillation, Submitted to Geophysical Model Development (GMD), Preprint available: https://arxiv.org/abs/1902.04621.
- Ebert-Uphoff, I., & Hilburn, K. A. (2020). Evaluation, Tuning and Interpretation of Neural Networks for Meteorological Applications. Submitted to Bulletin of the American Meteorological Society (in review). Preprint available: https://arxiv.org/abs/2005.03126
- Lapuschkin et al. “Unmasking Clever Hans Predictors and Assessing What Machines Really Learn.” Nature Communications, vol. 10, no. 1, Mar. 2019, p. 1096, doi:10.1038/s41467-019-08987-4.
- Ebert-Uphoff, Imme, Savini Samarasinghe, and Elizabeth A. Barnes: Thoughtfully Using Artificial Intelligence in Earth Science, EOS, 100, https://doi.org/10.1029/2019EO135235.

Extra slides

LRP Example Propagation Rules



tunable parameters: α, β
fixed parameters: a, w, R

$$R_j = \sum_k \left(\alpha \frac{a_j w_{jk}^+}{\sum_j a_j w_{jk}^+} - \beta \frac{a_j w_{jk}^-}{\sum_j a_j w_{jk}^-} \right) R_k$$

one possible propagation rule
(there are many)