



AGS CLD CAREER Award
Graduate Research Fellowship
Postdoctoral Fellowship



Benefits of saying “I Don’t Know” when analyzing and modeling the climate system with ML



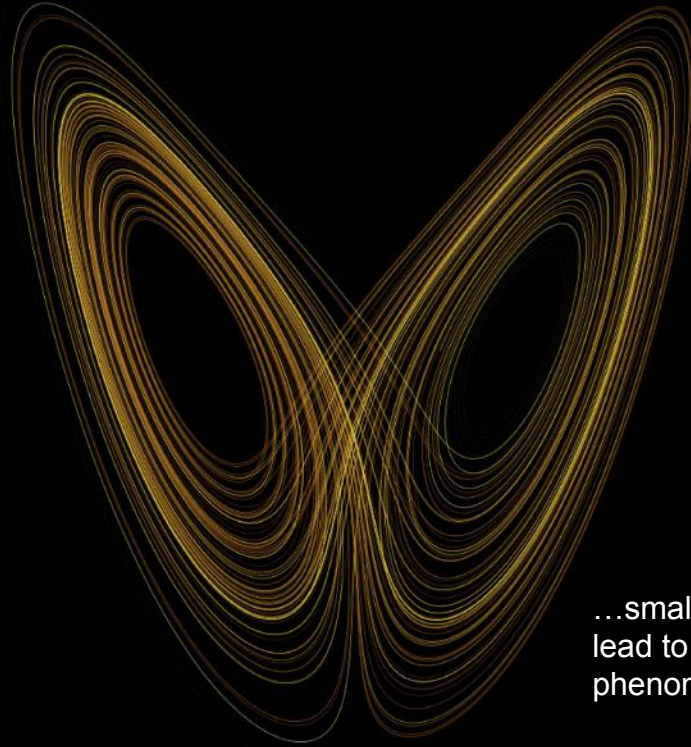
COLORADO STATE
UNIVERSITY

Prof. Elizabeth A. Barnes

STAMPS - Carnegie Mellon University
December 3, 2021

CHAOS

When the flap of a butterfly's wings
in Brazil sets off a tornado in Texas.
- Edward Lorenz (1972)



...small differences in the initial positions may
lead to enormous differences in the final
phenomena. ***Prediction becomes impossible.***
- Henri Poincare (1903)

The earth system is exceedingly complex and often chaotic in nature, making prediction incredibly challenging.

We cannot expect to make perfect predictions all of the time...

Forecasts of Opportunity*

certain conditions lead to more
predictable behaviour than others

Beyond the weather timescale we must look for specific states of the earth system, i.e. “opportunities”, that lead to enhanced predictable behavior.

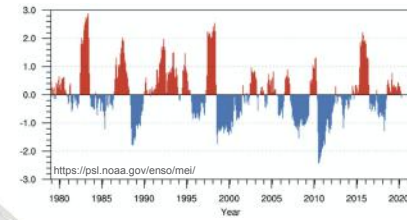
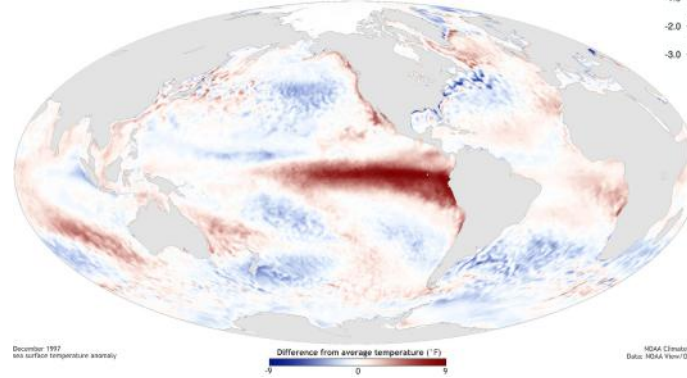
* also known as “State-dependent predictability”

See also Mariotti et al. (2020) and also Albers and Newman (2019)

El Nino Southern Oscillation [ENSO]

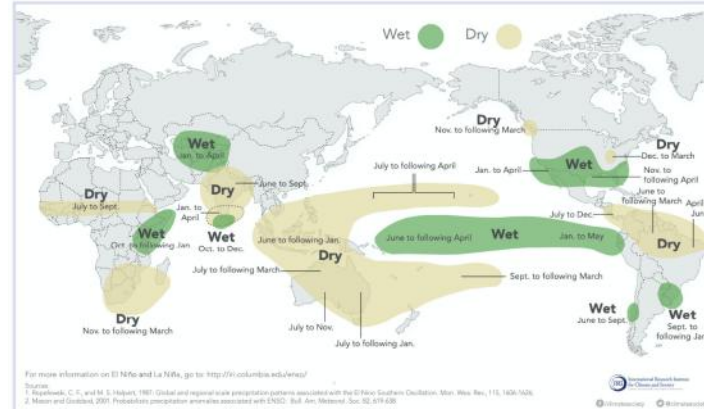
- Timescales of seasons-to-years
- Long-studied tropical phenomenon that, **when active**, impacts weather across the globe

Sea surface temperature anomalies
Dec. 1997 El Nino



El Niño and Rainfall

El Niño conditions in the tropical Pacific are known to shift rainfall patterns in many different parts of the world. Although they vary somewhat from one El Niño to the next, the strongest shifts remain fairly consistent in the regions and seasons shown on the map below.





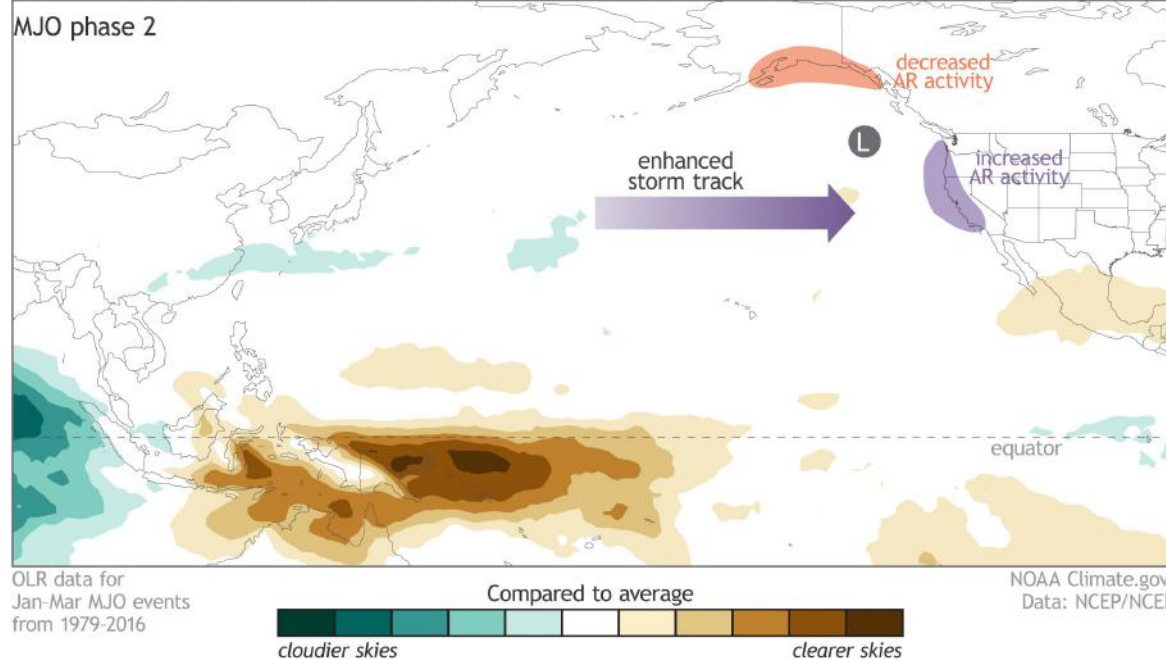
Finding forecasts of opportunity

1. **When?** Under what conditions do we have skillful forecasts of opportunity?
2. **Why?** Where is this predictability coming from?
3. **How** do we leverage these opportunities?

Global teleconnections

Climate phenomena can influence weather across the globe via atmospheric teleconnections

Madden-Julian Oscillation and mid-latitude impacts



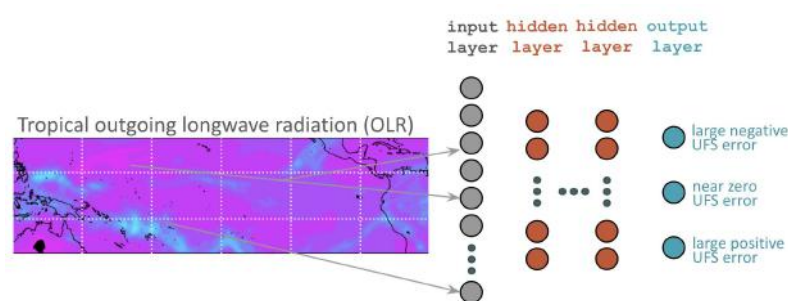
Post-processing by predicting forecast errors

Identifying state-dependent forecast errors in the NOAA Unified Forecast System

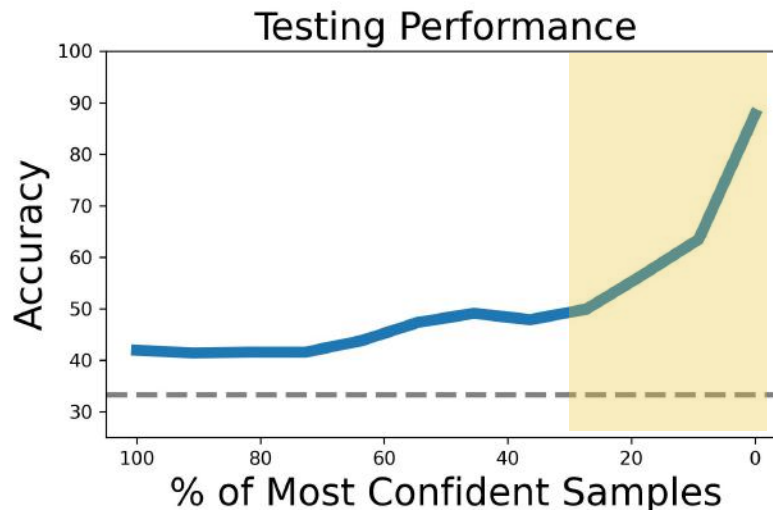
Network Task: Train a neural network to ingest daily maps of outgoing longwave radiation (OLR) to predicting the 5-day averaged precipitation error elsewhere at 10-14 day lead

Forecasts of Opportunity: Confident predictions lead to more accurate predictions = forecasts of opportunity

Explainable AI Approach: Learn tropical patterns of variability that lead to predictable forecast errors



Jack Cahill



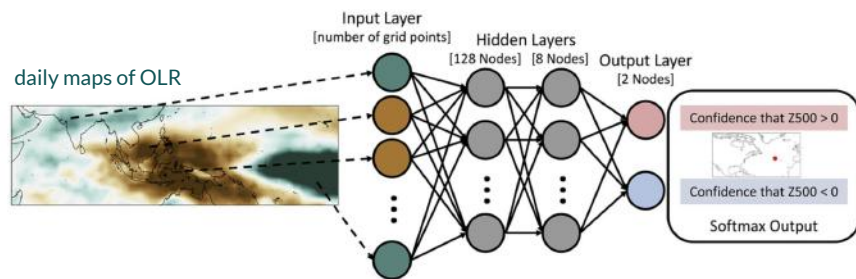
Subseasonal weather prediction

Exploring how tropical information can lend predictability to midlatitude circulation on S2S timescales

Network Task: Train a neural network to ingest daily maps of outgoing longwave radiation (OLR) to predict the sign of the subseasonal circulation anomalies over the North Atlantic 22 days in advance

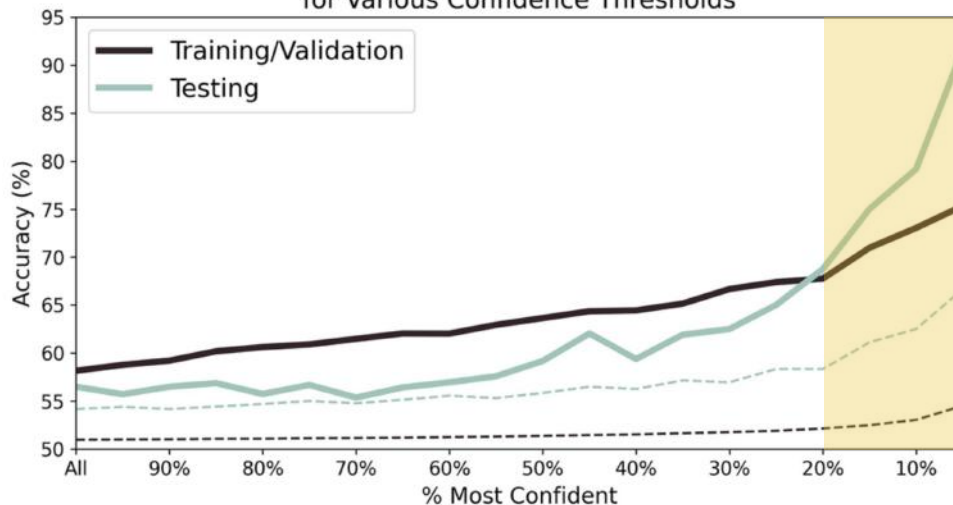
Forecasts of Opportunity: Confident predictions lead to more accurate predictions = forecasts of opportunity

Explainable AI Approach: Learn tropical patterns of variability that lead to enhanced predictability of midlatitude weather on subseasonal timescales



Kirsten Mayer

b) Single Model Prediction Accuracy for Various Confidence Thresholds



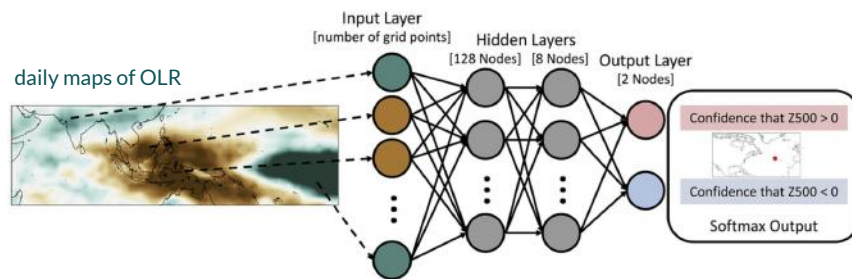
Subseasonal weather prediction

Exploring how tropical information can lend predictability to midlatitude circulation on S2S timescales

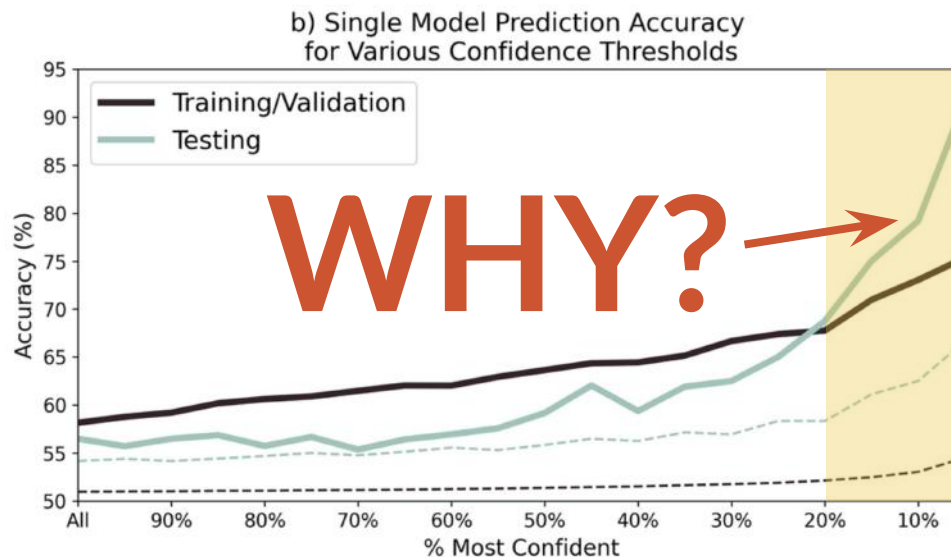
Network Task: Train a neural network to ingest daily maps of outgoing longwave radiation (OLR) to predict the sign of the subseasonal circulation anomalies over the North Atlantic 22 days in advance

Forecasts of Opportunity: Confident predictions lead to more accurate predictions = forecasts of opportunity

Explainable AI Approach: Learn tropical patterns of variability that lead to enhanced predictability of midlatitude weather on subseasonal timescales



Kirsten Mayer



Opening the Black Box with XAI

In the past few years multiple papers have come out demonstrating the use of explainable AI (XAI) methods for geoscience

MAKING THE BLACK BOX MORE TRANSPARENT

Understanding the Physical Implications of Machine Learning

AMY MCGOVERN, RYAN LAGERQUIST, DAVID JOHN GAGNE II, G. EU JENSENSEN, KIMBERLY L. ELMORE, CAMERON R. HOFMEYER, AND TRAVIS SMITH

Machine learning model interpretation and visualization focusing on meteorological domains are introduced and analyzed.

Machine learning (ML) and deep learning (DL; LeCun et al. 2015) have recently achieved breakthroughs across a variety of fields, including the world's best Go player (Silver et al. 2016, 2017), medical diagnosis (Rakhtin et al. 2018), and galaxy

classification (Dieleman et al. 2015). Simple forms of ML (e.g., linear regression) since ML has been used hazards since we use linear regression, large

AFFILIATIONS: McGovern and Jensen—University of Oklahoma, Norman, Oklahoma, LEICHTEN—Cooperative Institute for Mesoscale Meteorological Studies, and University of Oklahoma, Norman, Oklahoma, GAGNE—National Center for Atmospheric Research, Boulder, Colorado, ELMORE and Smith—Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma, HOFMEYER—School of Meteorology, University of Oklahoma, Norman, Oklahoma

CORRESPONDING AUTHOR: Amy McGovern, amcgovern@ou.edu

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-D-18-01915.1

A supplement to this article is available online (10.1175/BAMS-D-18-01915.2).

In final form 20 June 2019

©2019 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy.

AMERICAN METEOROLOGICAL SOCIETY

BAMS

ISSUES EARLY ONLINE RELEASE COLLECTIONS FOR AUTHORS

Article Contents

Capable:

Abstract:

Footnote:

RESEARCH ARTICLE | 31 AUGUST 2020

Evaluation, Tuning and Interpretation of Neural Networks for Working with Images in Meteorological Applications

RYAN LAGERQUIST AND G. EU JENSEN

10.1175/BAMS-D-20-00977.1

Split-Screen PDF Share Cite Get Permissions

Capable:

This article discusses strategies for the development of neural networks (aka deep learning) for meteorological applications. Topics include evaluation, tuning and interpretation of neural networks for working with meteorological images.

Abstract

The method of neural networks (aka deep learning) has opened up many new opportunities to utilize remotely sensed images in meteorology. Common applications include image classification, e.g., to determine whether an image contains a tropical cyclone, and image-to-image translation, e.g., to emulate radar imagery for satellites that only have passive channels. However, there are yet many open questions regarding the use of neural networks for working with meteorological images, such as best practices for evaluation, tuning and interpretation. This article highlights several strategies and practical considerations for neural network development that have not yet received much attention in the meteorological community, such as the concept of receptive fields, underutilized meteorological performance measures, and methods for neural network interpretation, such as synthetic experiments and layer-wise relevance propagation. We also consider the process of neural network interpretation as a whole, recognizing it as an iterative meteorologist-driven discovery process that builds on experimental design and hypothesis generation and testing. Finally, while most work on neural network interpretation in meteorology has so far focused on networks for image classification tasks, we expand the focus to also include networks for image-to-image translation.



JAMES Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE

10.1029/2019MS001800

Key Points

• Interpretable neural networks can identify the coherent spatial patterns of climate model Earth system variability

• The feature relevance propagation and backward optimization methods enable new ways to use neural networks for geoscientific research

• We propose that the interpretation of climate neural networks has moved on to the other scientific domains of a neural network

Supporting Information

• Supporting Information S1

Correspondence to: A. T. Tsonis, at.tsonis@colorado.edu

Citation: Tsonis, A. T., Barnes, S. A., A. T. Tsonis, S. A. (2020). Understanding the physical implications of machine learning for the geosciences: Applications to Earth system variability. *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001800. <https://doi.org/10.1029/2019MS001800>

Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability

Benjamin A. Tsonis¹, Elizabeth A. Barnes², and Imme Thiele-Ehrenfeld³

¹Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO, USA, ²Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO, USA, ³Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA

Abstract Neural networks have become increasingly prevalent within the geosciences, although a common limitation of their usage has been a lack of methods to interpret what the networks learn and how they make decisions. As such, neural networks have often been used within the geosciences to most accurately identify a desired output given a set of inputs, with the interpretation of what the network learns used as a secondary metric to ensure the network is making the right decision for the right reason. Neural network interpretation techniques have become more advanced in recent years, however, and we therefore propose that the ultimate objective of using a neural network can also be the interpretation of what the network has learned rather than the output itself. We show that the interpretation of neural networks can enable the discovery of scientifically meaningful connections within geoscientific data. In particular, we use two methods for neural network interpretation called backward optimization and layerwise relevance propagation, both of which project the decision pathways of a network back onto the original input dimensions. To the best of our knowledge, LRP has not yet been applied to geoscientific research, and we believe it has great potential in this area. We show how these interpretation techniques can be used to reliably extract scientifically meaningful information from neural networks by applying them to common climate patterns. These results suggest that combining interpretable neural networks with novel scientific hypotheses will open the door to many new avenues in neural network related geoscientific research.

Plain Language Summary Neural networks, a form of machine learning, have become a common tool in the geosciences. However, a common limitation of neural networks in geoscience has been that their decision-making process is uninterpretable. This has meant that neural networks, since a understanding of how and why they make decisions. Methods for interpreting neural networks have become more advanced in recent years, however, and we therefore propose that the ultimate objective of using a neural network can also be the interpretation of what the network has learned rather than the output itself. We show that the interpretation of neural networks can enable the discovery of scientifically meaningful connections within geoscientific data. In particular, we use two methods for neural network interpretation called backward optimization and layerwise relevance propagation, both of which project the decision pathways of a network back onto the original input dimensions. To the best of our knowledge, LRP has not yet been applied to geoscientific research, and we believe it has great potential in this area. We show how these interpretation techniques can be used to reliably extract scientifically meaningful information from neural networks by applying them to common climate patterns. These results suggest that combining interpretable neural networks with novel scientific hypotheses will open the door to many new avenues in neural network related geoscientific research.

A powerful tool in scientific applications across all areas of geoscience, including marine science (e.g., Mahe et al., 2018), and atmospheric science (e.g., Barnes et al., 2019; Tsonis et al., 2019). This revolution in machine learning within climate science of novel algorithms, as well as the high computational power for processing immense quantities of data to the application of machine learning methods within

1 of 20

Layerwise Relevance Propagation

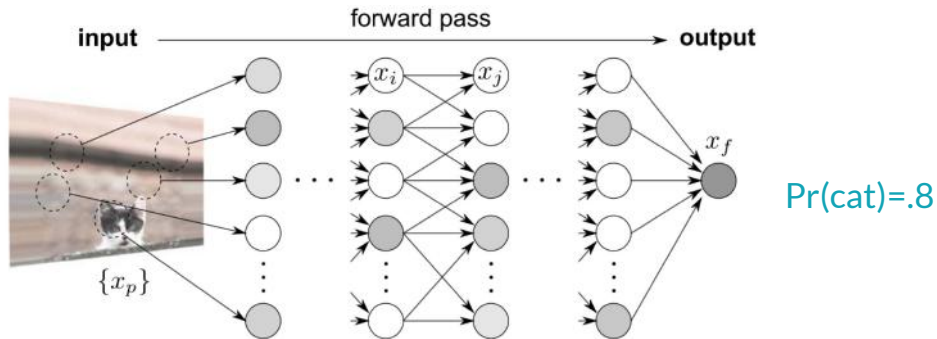
LRP is an XAI method that produces a heatmap of the most relevant regions of the input for each prediction

LRP is largely consistent with how many climate scientists analyze and interpret data methods

While many visualization tools are coming out of the computer science community, LRP has been most useful for our group thus far

Prediction

of 1 sample



Layerwise Relevance Propagation

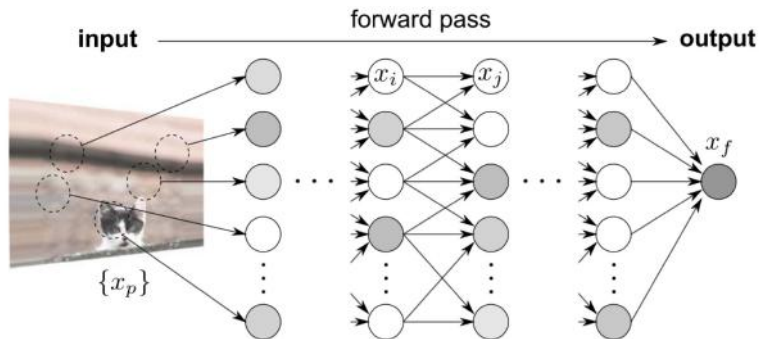
LRP is an XAI method that produces a heatmap of the most relevant regions of the input for each prediction

LRP is largely consistent with how many climate scientists analyze and interpret data methods

While many visualization tools are coming out of the computer science community, LRP has been most useful for our group thus far

Prediction

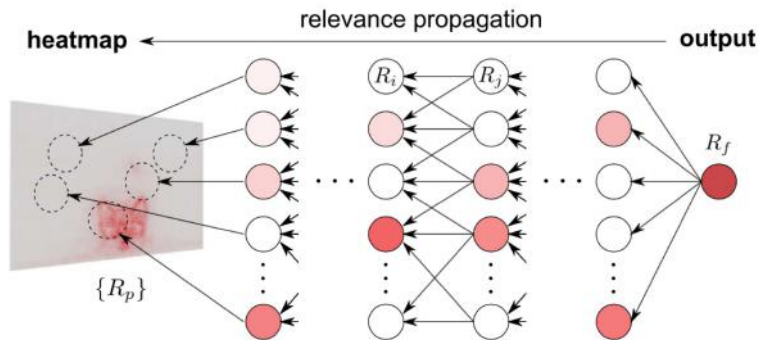
of 1 sample



$\Pr(\text{cat})=.8$

LRP

of 1 sample



$\Pr(\text{cat})=.8$

Why use XAI?

One use of LRP is to ensure the right answers for the right reasons

Example Task:

Decide whether there is a horse in a given image.

Interpretable AI Approach:

What strategy did the network use? Is it focusing on the right things?



Why use XAI?

One use of LRP is to ensure the right answers for the right reasons



Example Task:

Decide whether there is a horse in a given image.

Interpretable AI Approach:

What strategy did the network use? Is it focusing on the right things?

Why use XAI?

#1 Identify problematic strategies

#2 Evaluate trust

#3 Choose the approach

#4 Learn something new

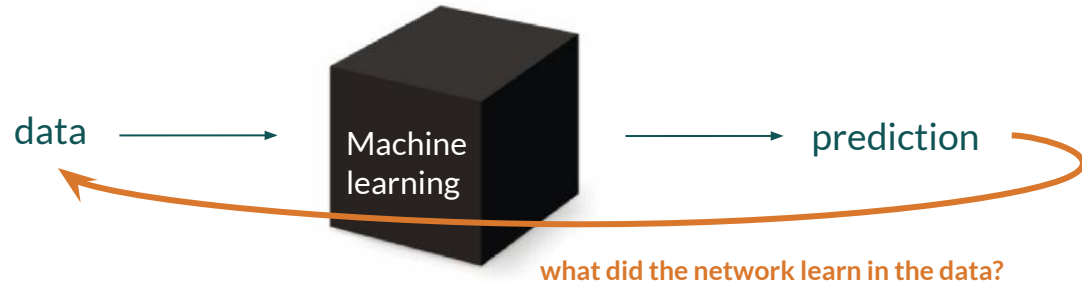
Why use XAI?

#1 Identify problematic strategies

#2 Evaluate trust

#3 Choose the approach

#4 Learn something new



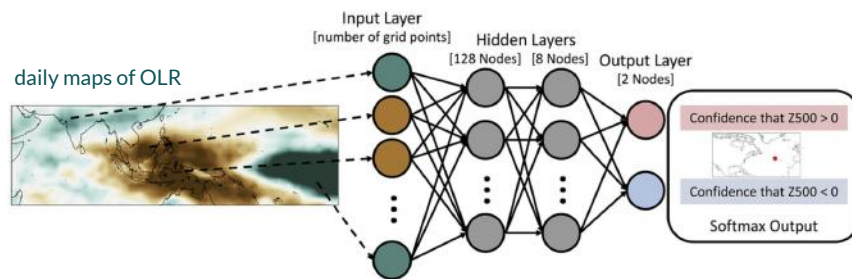
Subseasonal weather prediction

Exploring how tropical information can lend predictability to midlatitude circulation on S2S timescales

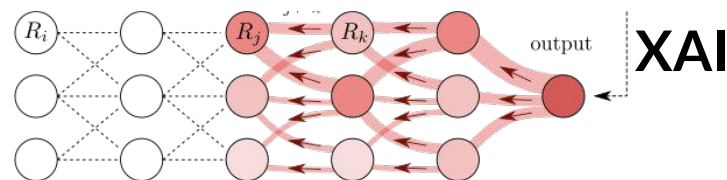
Network Task: Train a neural network to ingest daily maps of outgoing longwave radiation (OLR) to predict the sign of the subseasonal circulation anomalies over the North Atlantic 22 days in advance

Forecasts of Opportunity: Confident predictions lead to more accurate predictions = forecasts of opportunity

Explainable AI Approach: Learn tropical patterns of variability that lead to enhanced predictability of midlatitude weather on subseasonal timescales



Kirsten Mayer



apply explainable AI methods (e.g. post-hoc attribution approaches) to create a heatmap of relevant regions in the input for the network's prediction



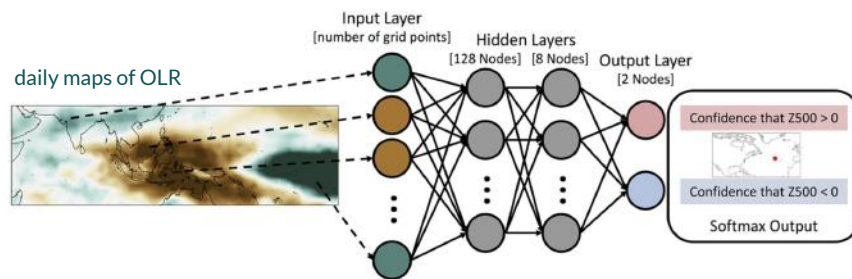
Subseasonal weather prediction

Exploring how tropical information can lend predictability to midlatitude circulation on S2S timescales

Network Task: Train a neural network to ingest daily maps of outgoing longwave radiation (OLR) to predict the sign of the subseasonal circulation anomalies over the North Atlantic 22 days in advance

Forecasts of Opportunity: Confident predictions lead to more accurate predictions = forecasts of opportunity

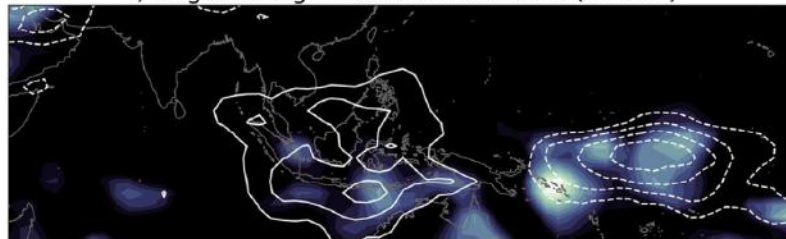
Explainable AI Approach: Learn tropical patterns of variability that lead to enhanced predictability of midlatitude weather on subseasonal timescales



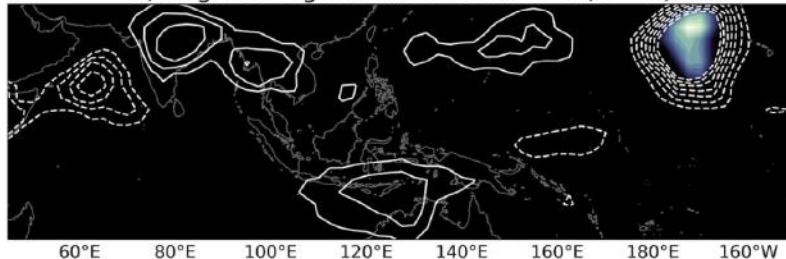
Kirsten Mayer

LRP: OLR patterns that lead to accurate & confident Z500<0 predictions over the North Atlantic

f) Negative Sign Prediction Cluster 1 (N=127)

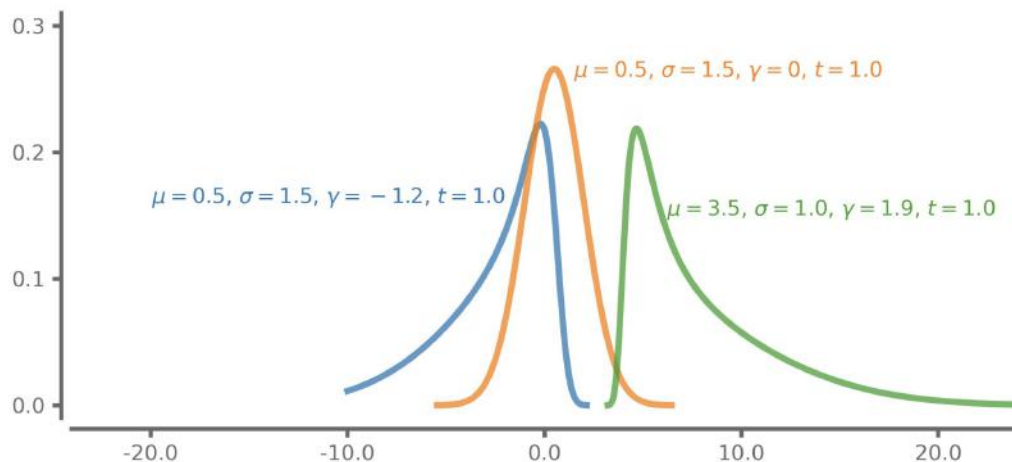
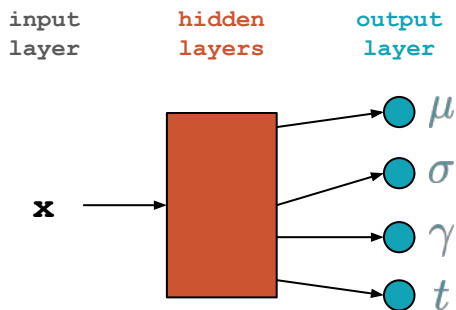


h) Negative Sign Prediction Cluster 2 (N=48)



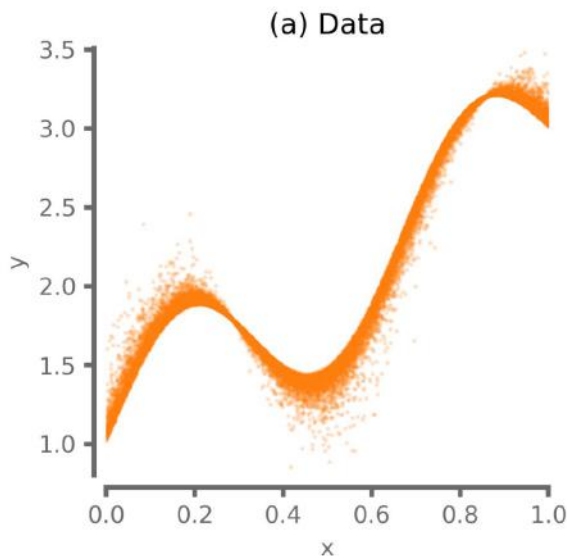
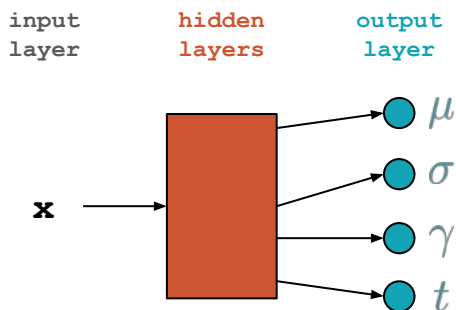
What about regression problems?

Simple uncertainty for neural network regression tasks



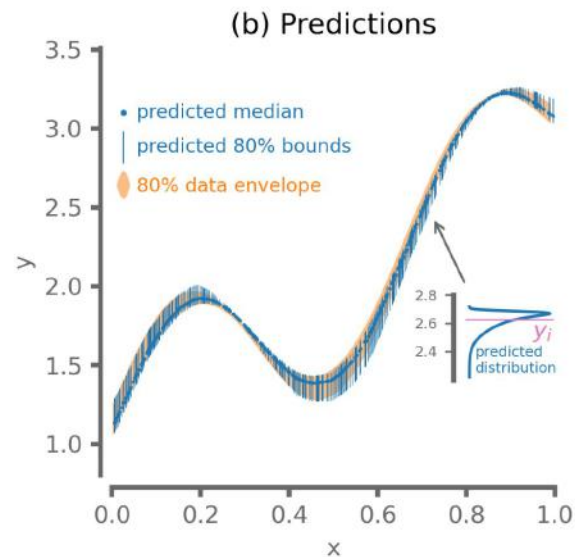
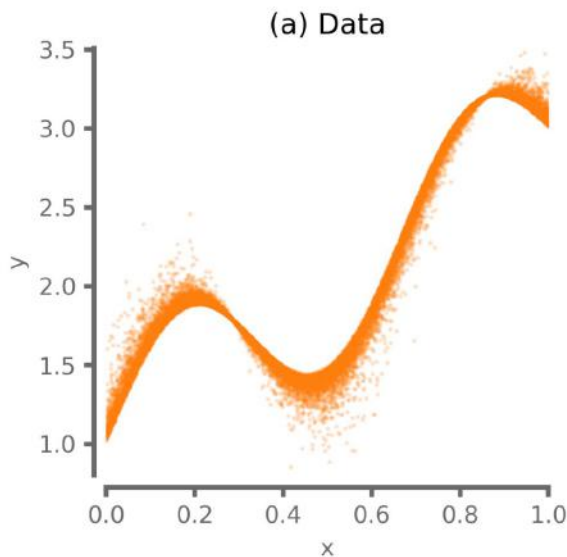
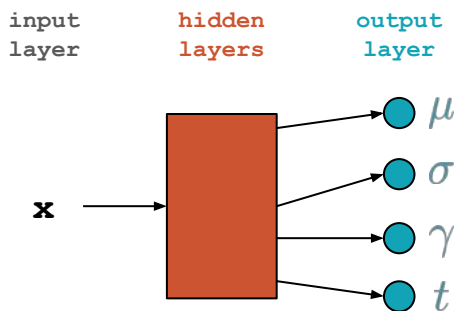
trained with $-\log(p)$, based on maximum likelihood estimation
Barnes, Barnes and Gordillo (2021)

Simple uncertainty for neural network regression tasks



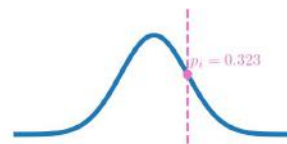
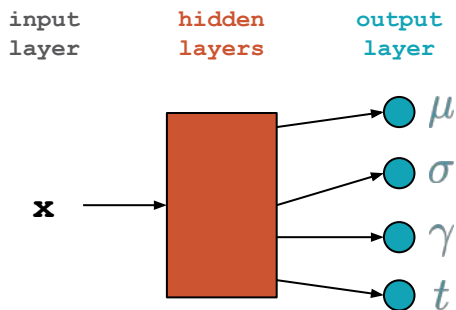
trained with $-\log(p)$, based on maximum likelihood estimation
Barnes, Barnes and Gordillo (2021)

Simple uncertainty for neural network regression tasks



trained with $-\log(p)$, based on maximum likelihood estimation
Barnes, Barnes and Gordillo (2021)

The loss function

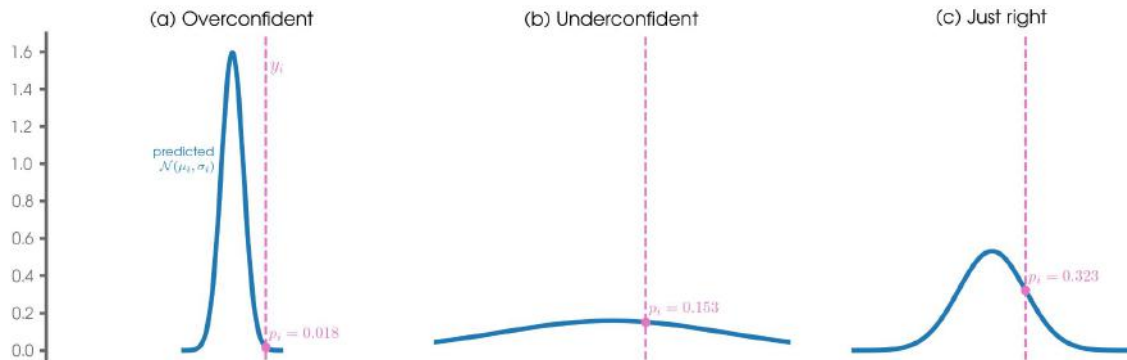
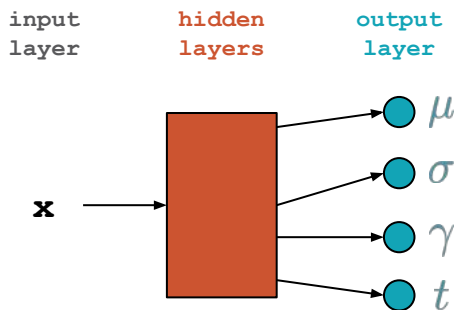


$$p_i = \mathcal{P}(y_i | \mu, \sigma, \gamma, \tau)$$

$$\mathcal{L}(x_i) = -\log p_i$$

An orange arrow points from the p_i in the equation above to the p_i in this equation.

The loss function

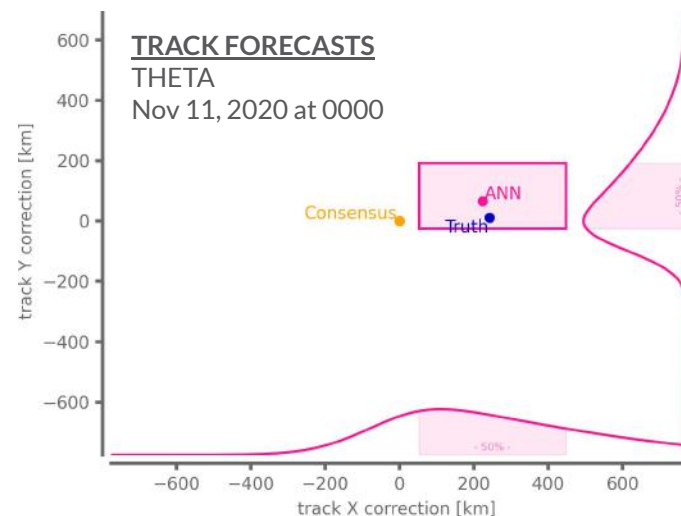
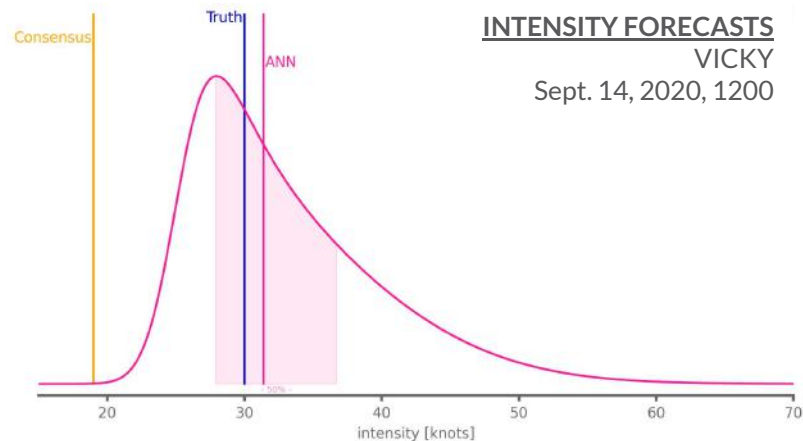


$$p_i = \mathcal{P}(y_i | \mu, \sigma, \gamma, \tau)$$

$$\mathcal{L}(x_i) = -\log p_i$$

Operational hurricane forecasts

- **Network Task:** Train neural networks to predict error of the “Consensus” forecast of physics-based models (used by the National Hurricane Center)
- **Predicted PDF:** Allows us to update the forecast as well as understand the uncertainty of the ANN

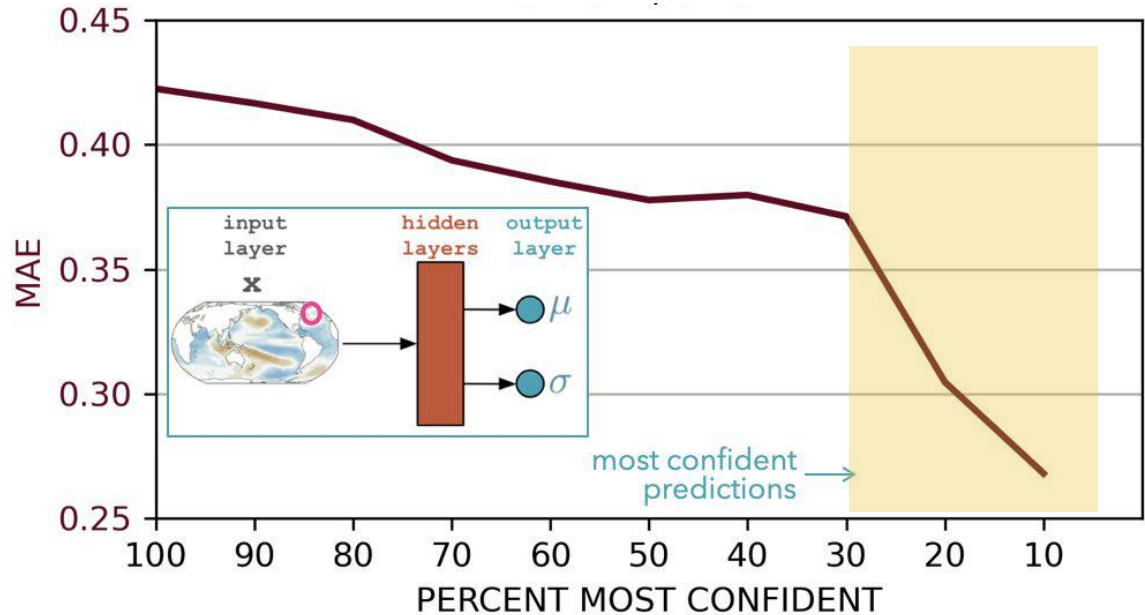


Decadal Prediction

- **Network Task:** Train a neural network to ingest maps of Ocean Heat Content and predict future sea-surface temperatures averaged over the next 1-5 years
- **State-Dependent Predictability:** Confident predictions lead to more smaller errors
- **Explainable AI Approach:** Learn decadal patterns of variability that lead to predictability



Emily Gordon



Controlled Abstention Networks (CAN)

The abstention loss works by incorporating uncertainty in the network's prediction to identify the more confident samples and abstain (say "I don't know") on the less confident samples.

...the abstention loss is applied *during training* to preferentially learn from the more confident samples.

Our work has heavily informed
by the dissertation of Dr. Sunil
Thulasidasan (2020):

- <https://arxiv.org/abs/1905.10964>
- https://github.com/thulas/dac-label-noise/blob/master/dac_loss.py
- <https://digital.lib.washington.edu/researchworks/handle/1773/45781>

General Idea of CANs

1. Estimate uncertainty of each prediction during training
 - **Classification:** simple - just use the softmax output
 - **Regression:** we need a way to predict uncertainty [more later]
2. Implement a loss function that learns to identify more confident predictions and learn them better
 - **Classification:** we introduce the *NotWrong Loss*
 - **Regression:** we introduce a modified *negative log likelihood*
3. Compare to baseline methods that filter out samples post training
 - While the baseline methods perform very well, we find that the **abstention method outperforms the baseline** for a variety of tasks



Our work has heavily informed
by the dissertation of Dr. Sunil
Thulasidasan (2020):

- <https://arxiv.org/abs/1905.10964>
- https://github.com/thulas/dac-label-noise/blob/master/dac_loss.py
- <https://digital.lib.washington.edu/researchworks/handle/1773/45781>

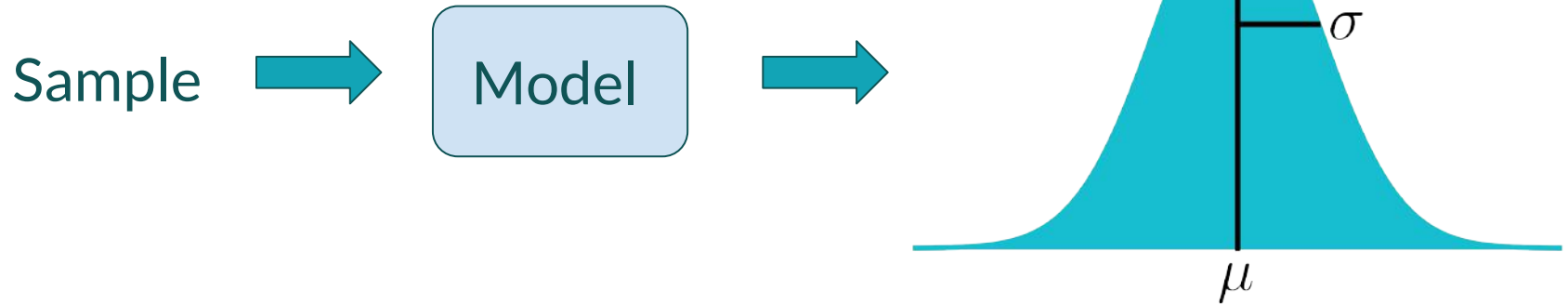
General Idea of CANs

1. Estimate uncertainty of each prediction during training
 - **Classification:** simple - just use the softmax output
 - **Regression:** we need a way to predict uncertainty (more later)
2. Implement a loss function that learns to identify more confident predictions and learn them better
 - **Classification:** we introduce the *NotWrong Loss*
 - **Regression:** we introduce a modified *negative log likelihood*
3. Compare to baseline methods that filter out samples post training
 - While the baseline methods perform very well, we find that the **abstention method outperforms the baseline** for a variety of tasks



—

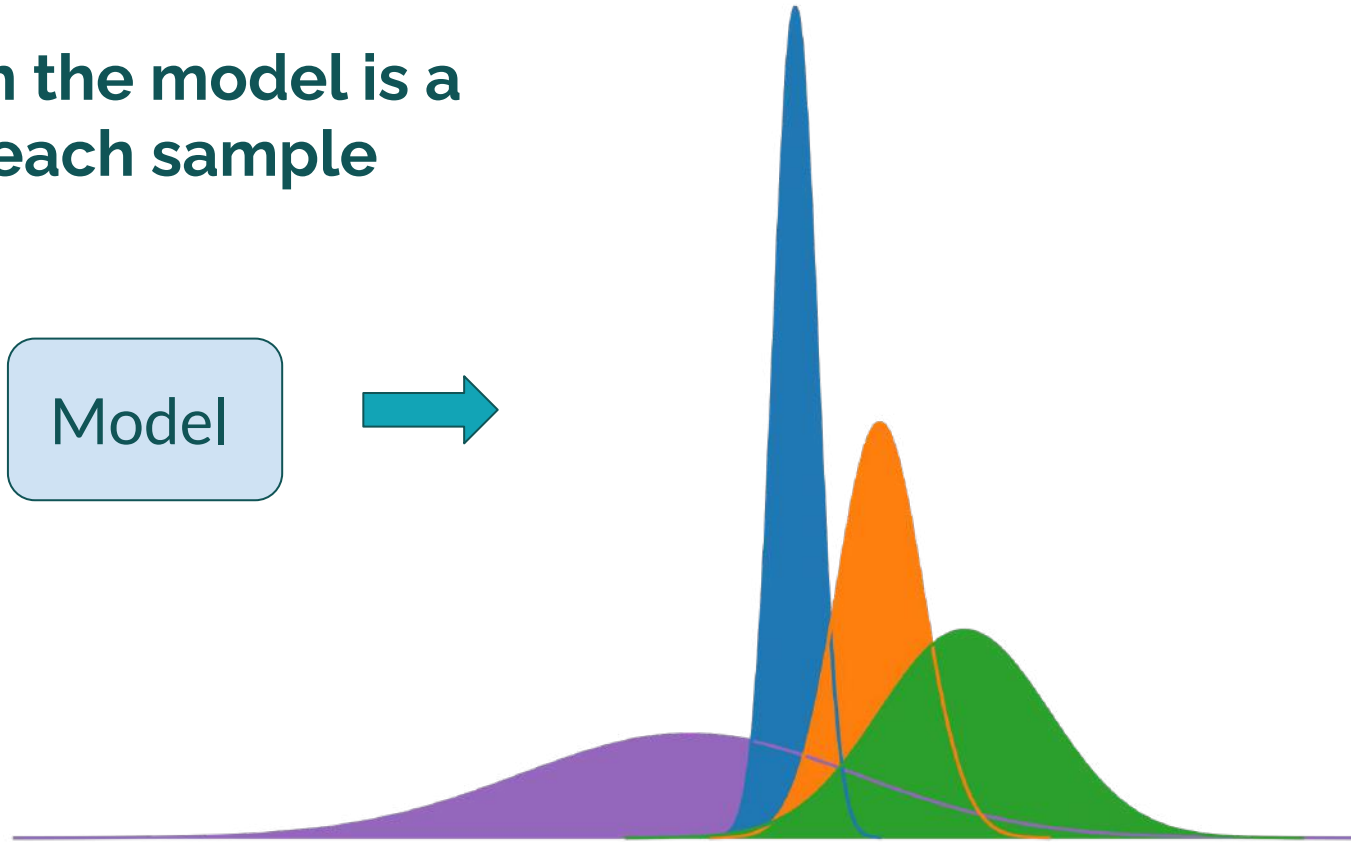
The output from the model is a distribution for each sample



—

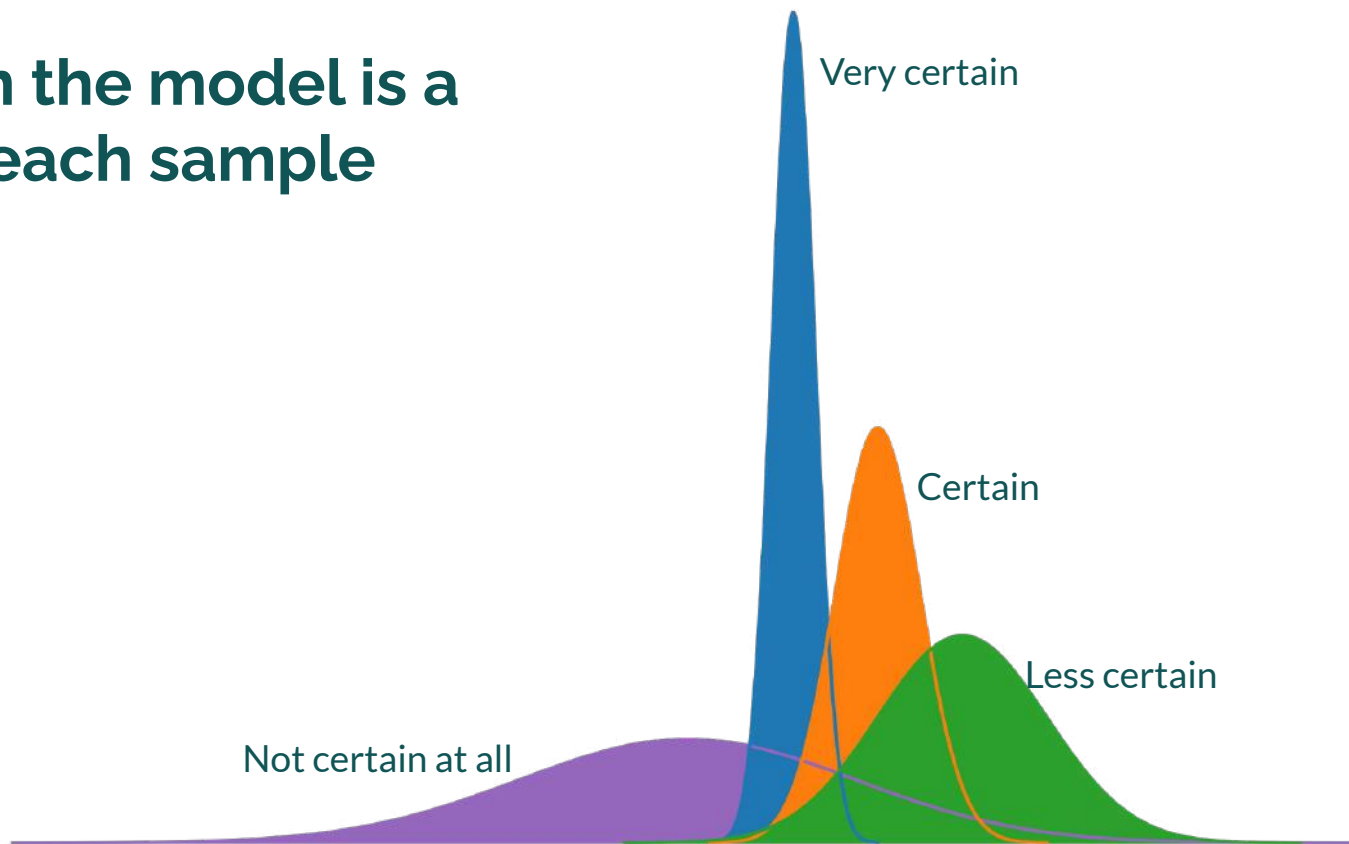
**The output from the model is a
distribution for each sample**

Many
Samples



—

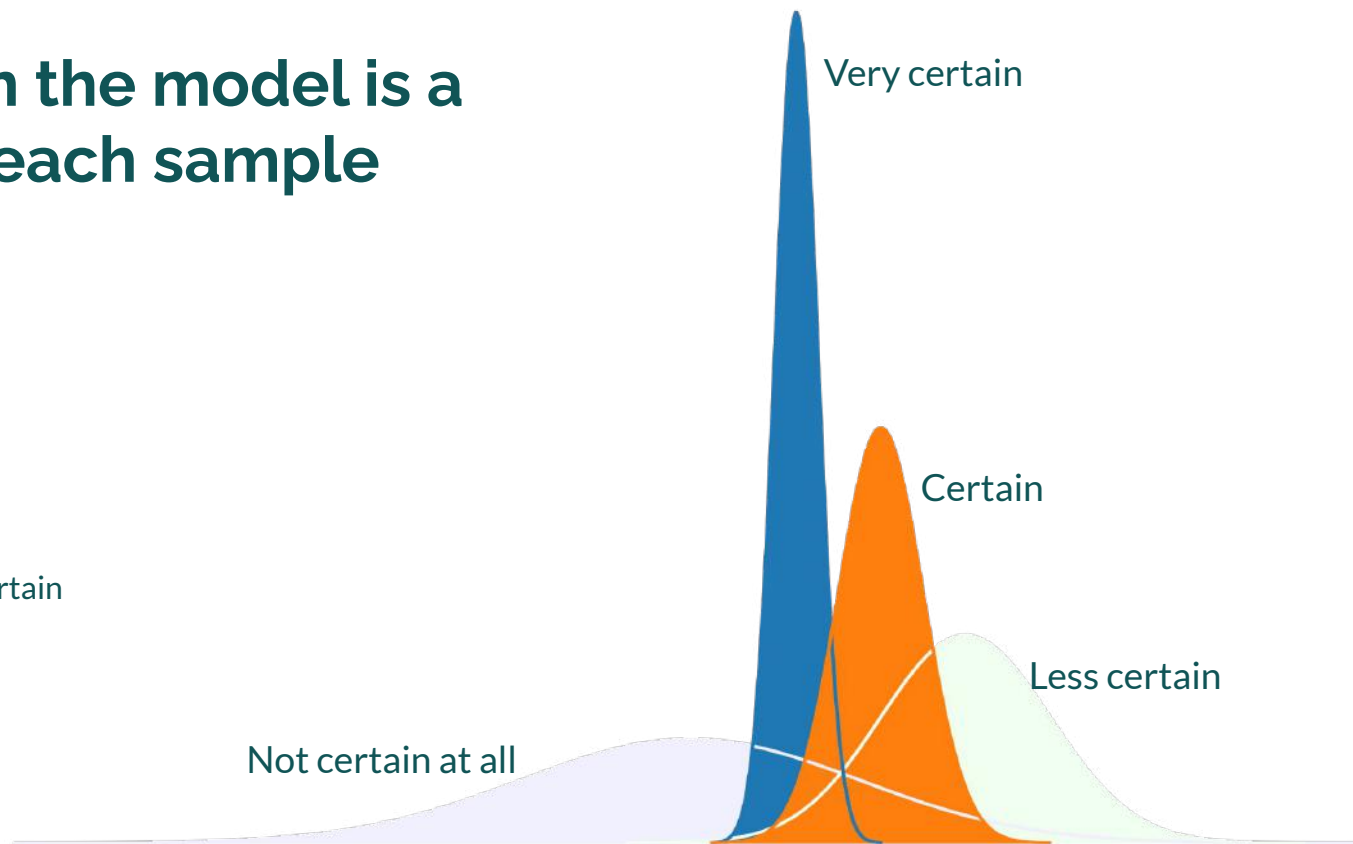
The output from the model is a distribution for each sample



The output from the model is a distribution for each sample

Baseline:

- Train the model
- Sort the results by sigma
- Keep only the x% most certain distributions.



—

**Now for abstention
during neural network
training**

Abstention During Training

- Abstention loss is very similar for both classification and regression
- The abstention regression loss is a modified log loss, weighted by the “prediction weight” determined by the uncertainty sigma
- An additional term penalizes abstention

3.2.1 Abstention loss

Unlike the baseline ANN, the CAN loss is designed to identify the less confident predictions so as to preferentially learn from the more confident predictions. The CAN loss for sample x_i is defined as

$$\mathcal{L}(x_i) = -q_i \log p_i - \alpha \log q_i. \quad (4)$$

where α controls the amount of abstention (see next subsection) and q_i represents the prediction weight defined as

$$q_i = \min \left(1.0, \left[\frac{\kappa}{\sigma_i} \right]^2 \right). \quad (5)$$

Abstention During Training

- Abstention loss is very similar for both classification and regression
- The abstention regression loss is a modified log loss, weighted by the “prediction weight” determined by the uncertainty sigma
- An additional term penalizes abstention

3.2.1 Abstention loss

Unlike the baseline ANN, the CAN loss is designed to identify the less confident predictions so as to preferentially learn from the more confident predictions. The CAN loss for sample x_i is defined as

$$\mathcal{L}(x_i) = -q_i \log p_i - \alpha \log q_i. \quad (4)$$

Diagram annotations for equation (4):
- A yellow arrow points from the label "baseline -log(p)" to the term $-\log p_i$.
- A red arrow points from the label "controls amount of abstention" to the parameter α .

where α controls the amount of abstention (see next subsection) and q_i represents the prediction weight defined as

$$q_i = \min \left(1.0, \left[\frac{\kappa}{\sigma_i} \right]^2 \right). \quad (5)$$

Diagram annotations for equation (5):
- A blue arrow points from the label "prediction weight" to the variable q_i .
- A blue arrow points from the label "data-specific scale" to the term $\left[\frac{\kappa}{\sigma_i} \right]^2$.

Abstention During Training

- Abstention loss is very similar for both classification and regression
- The abstention regression loss is a modified log loss, weighted by the “prediction weight” determined by the uncertainty sigma
- An additional term penalizes abstention
- **alpha**: abstention fraction can be set by a PID controller or user can have network predict the best abstention fraction

3.2.1 Abstention loss

Unlike the baseline ANN, the CAN loss is designed to identify the less confident predictions so as to preferentially learn from the more confident predictions. The CAN loss for sample x_i is defined as

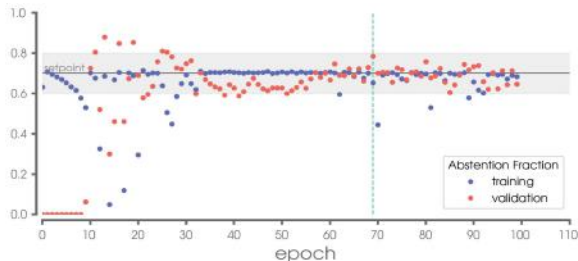
$$\mathcal{L}(x_i) = -q_i \log p_i - \alpha \log q_i. \quad (4)$$

baseline-log(p) controls amount of abstention

where α controls the amount of abstention (see next subsection) and q_i represents the prediction weight defined as

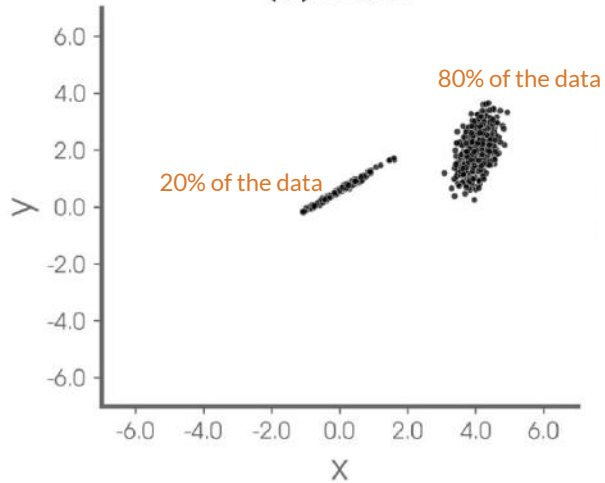
$$q_i = \min \left(1.0, \left[\frac{\kappa}{\sigma_i} \right]^2 \right). \quad (5)$$

prediction weight data-specific scale



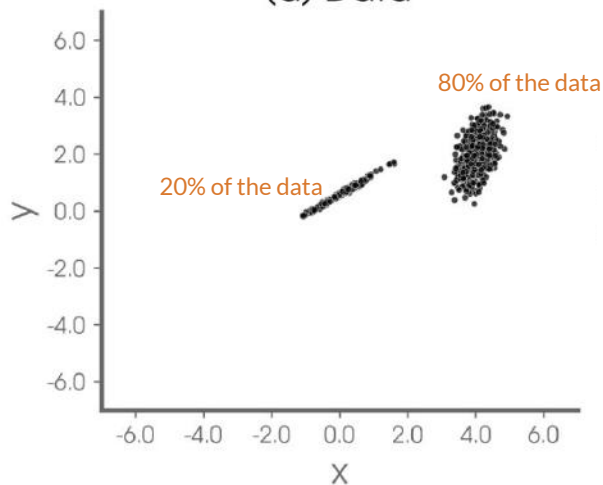
A simple 1D example

(a) Data

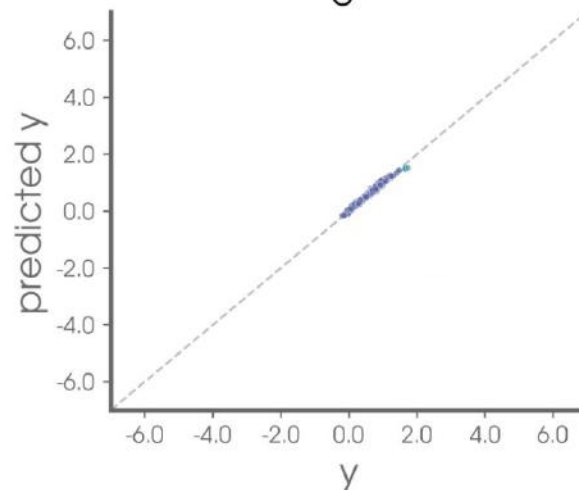


A simple 1D example

(a) Data

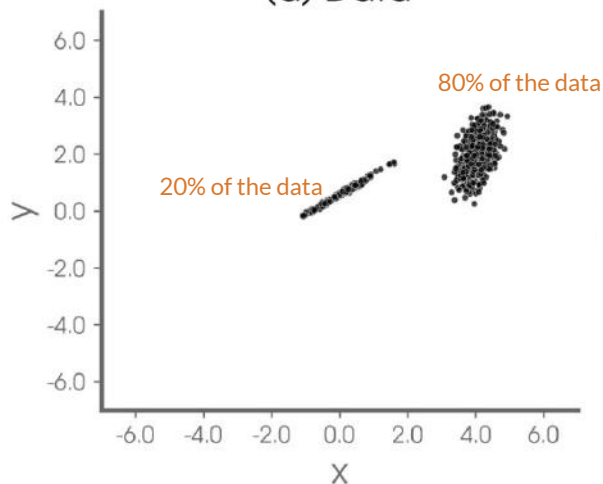


(c) CAN Predictions
coverage = 19%

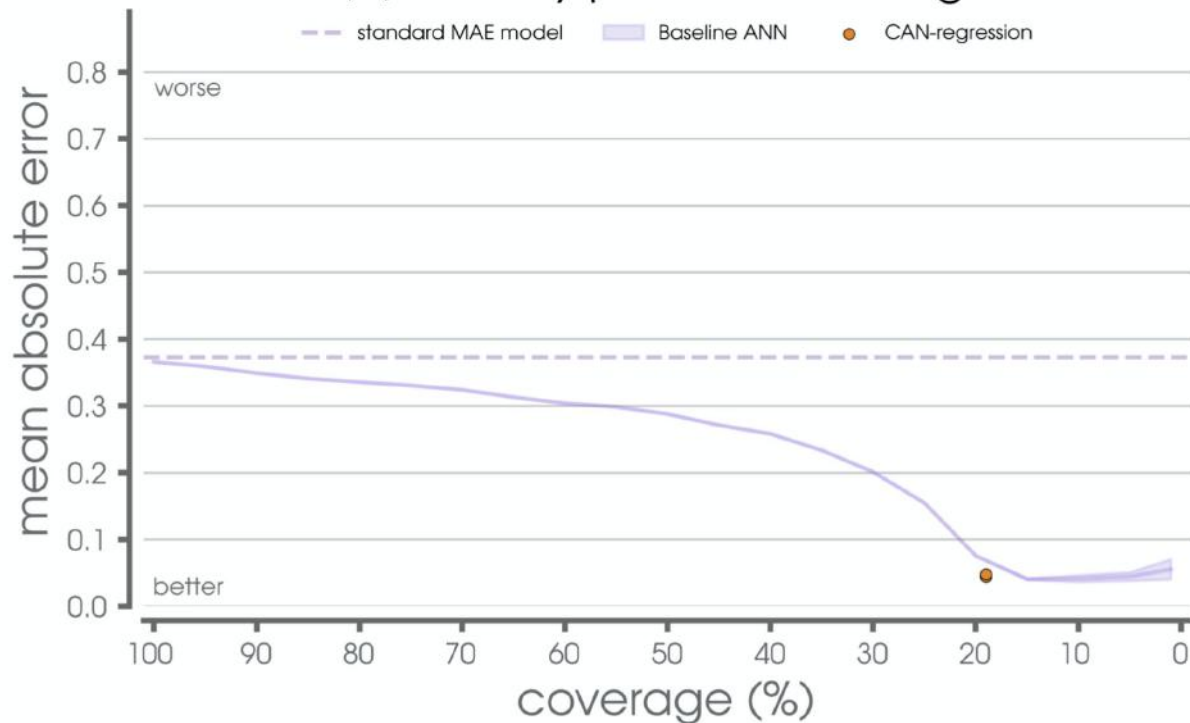


A simple 1D example

(a) Data

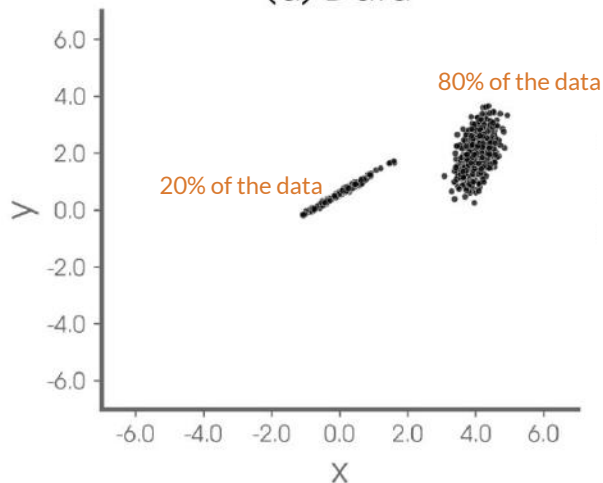


(d) Error by percent coverage

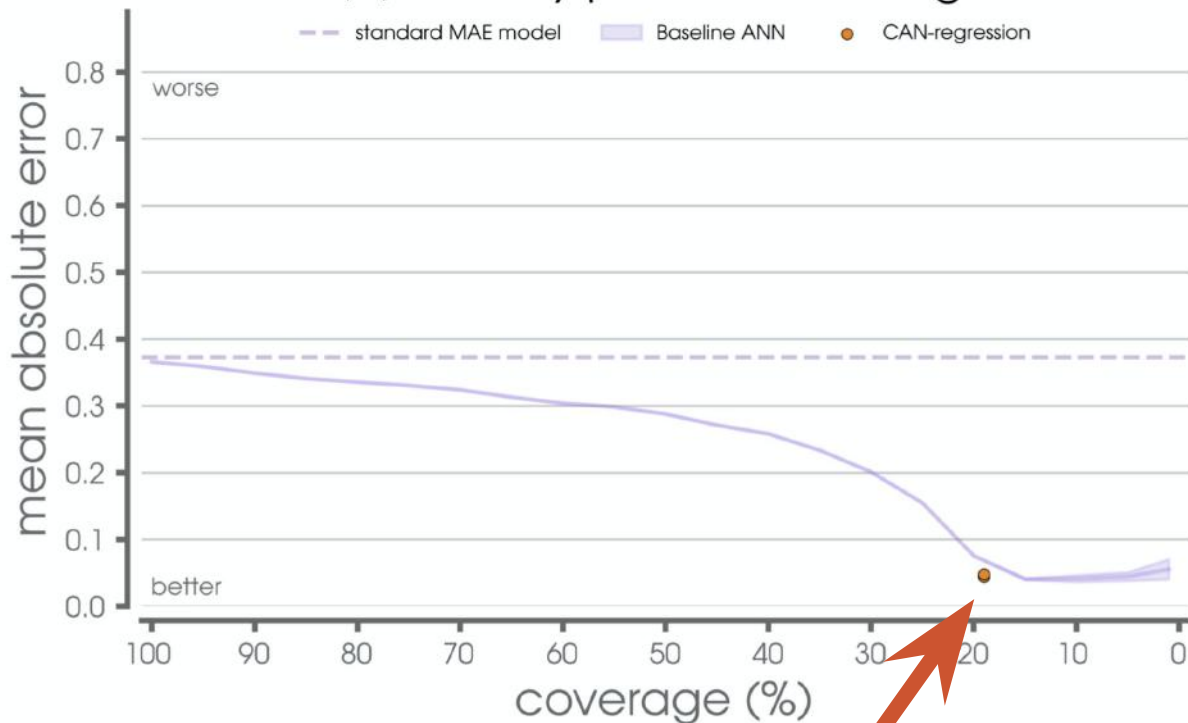


A simple 1D example

(a) Data



(d) Error by percent coverage



Barnes and Barnes (2021)

—

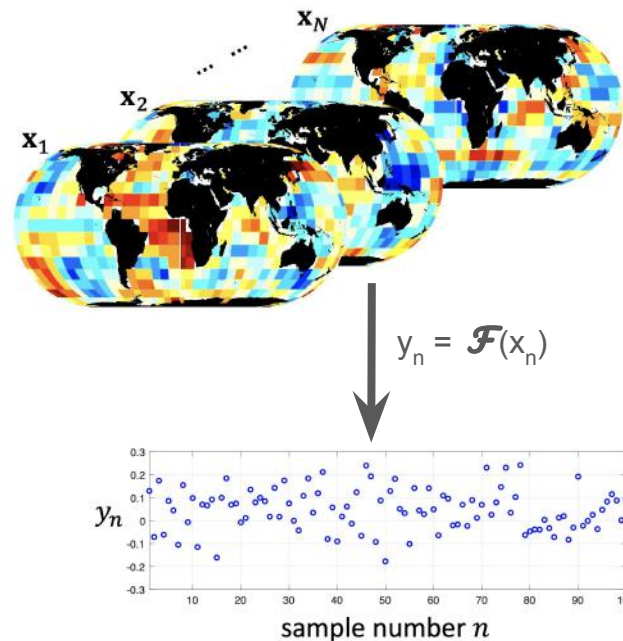
A more complex example

Synthetic Climate Data



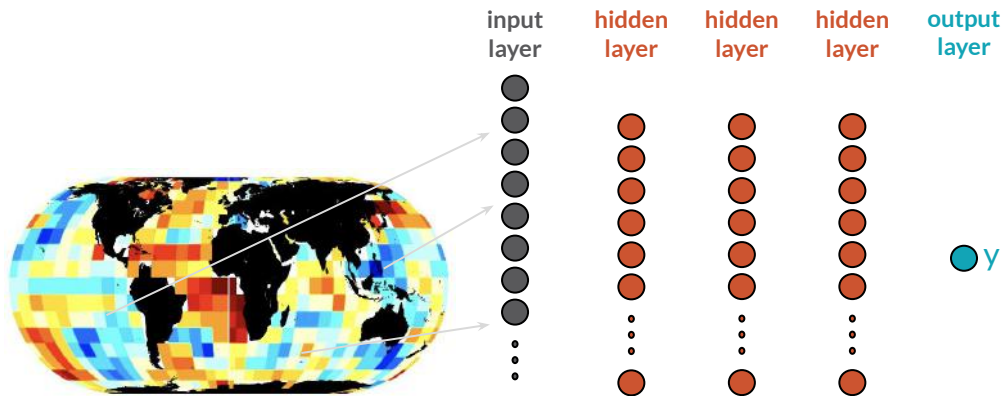
Dr. Antonios
Mamalakis

- Created by CSU postdoc Dr. Antonios Mamalakis
- Each sample is one global map of “SSTs” computed from real-world spatial covariances
- Use a known nonlinear function \mathcal{F} to map each map x_n to a scalar y_n



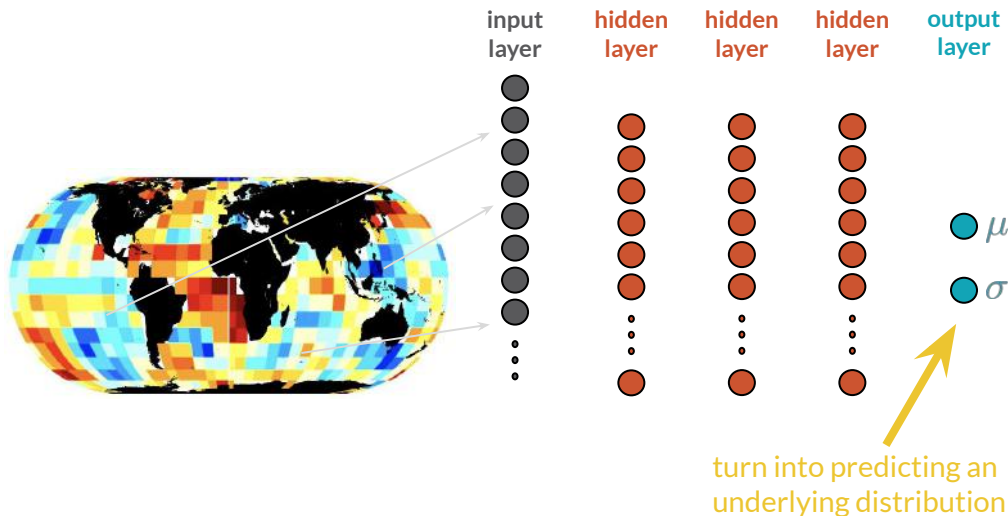
Synthetic Climate Data

- Created by CSU postdoc Dr. Antonios Mamalakis
- Each sample is one global map of “SSTs” computed from real-world spatial covariances
- Use a known nonlinear function \mathcal{F} to map each map x_n to a scalar y_n
- **Network Task: predict the value “y” for each input map**



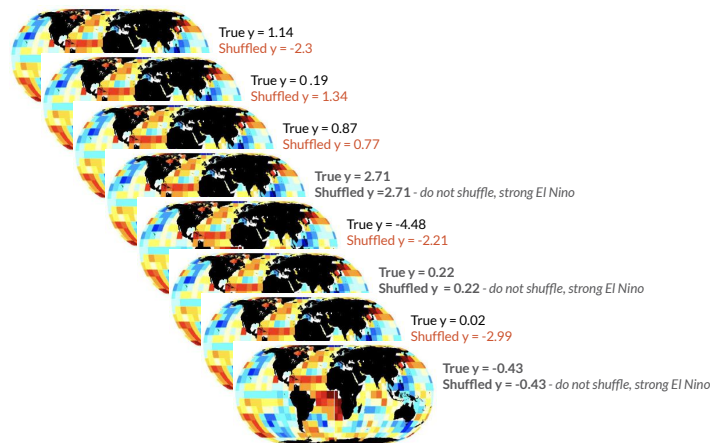
Synthetic Climate Data

- Created by CSU postdoc Dr. Antonios Mamalakis
- Each sample is one global map of “SSTs” computed from real-world spatial covariances
- Use a known nonlinear function \mathcal{F} to map each map x_n to a scalar y_n
- **Network Task: predict the value “y” for each input map**



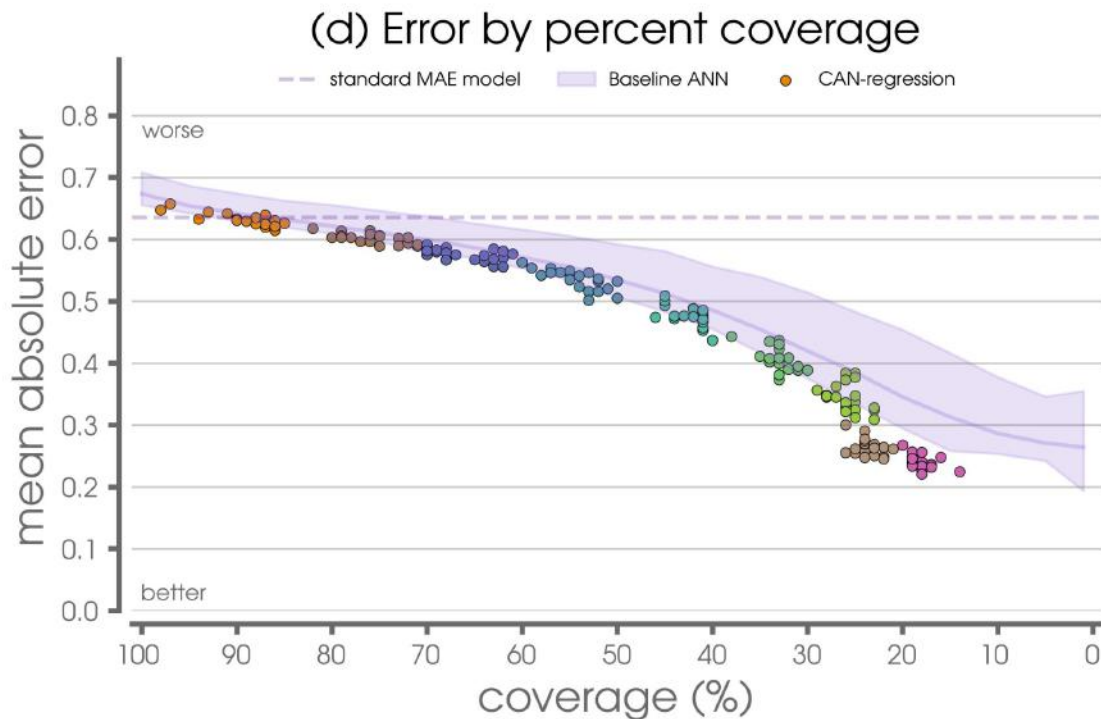
Forecasts of Opportunity Experiment

- All ENSO+ samples (average ENSO region > 0.5) are untouched
- 100% of the other samples are corrupted (shuffled)
- 29% untouched
- 71% corrupt
- **Only samples with strong El Nino signals have a learnable relationship with their labels**



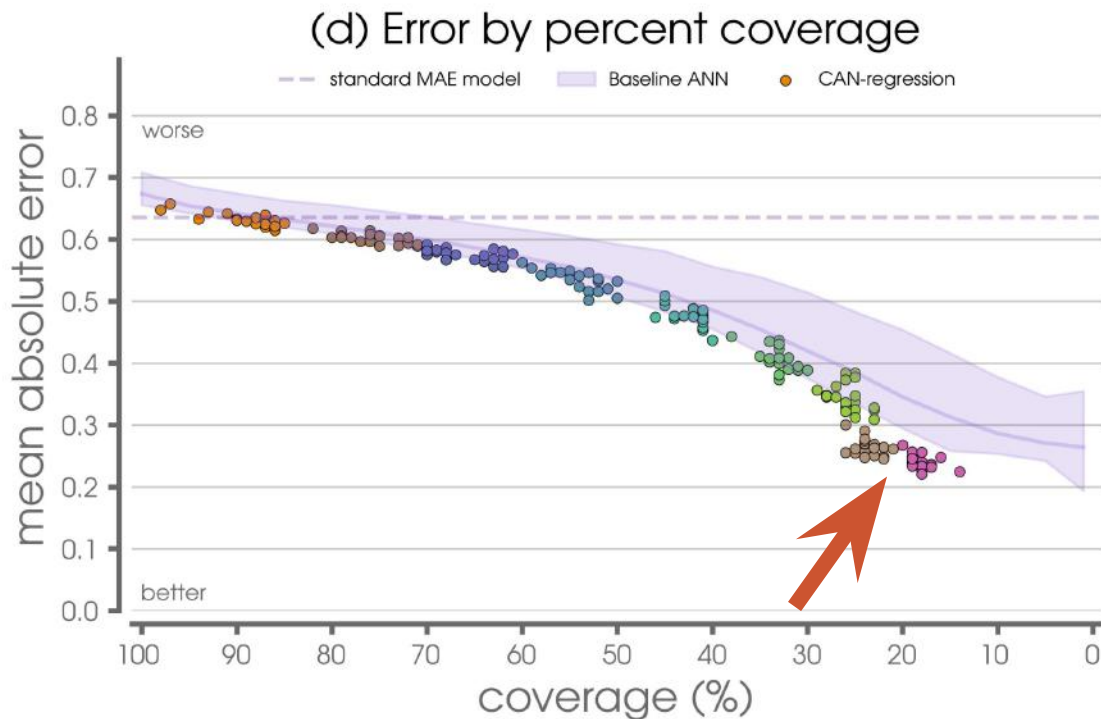
Abstention outperforms baseline

- Train abstention network for different abstention setpoints
- The best CAN models are always better (lower error) than the best baseline ANN



Abstention outperforms baseline

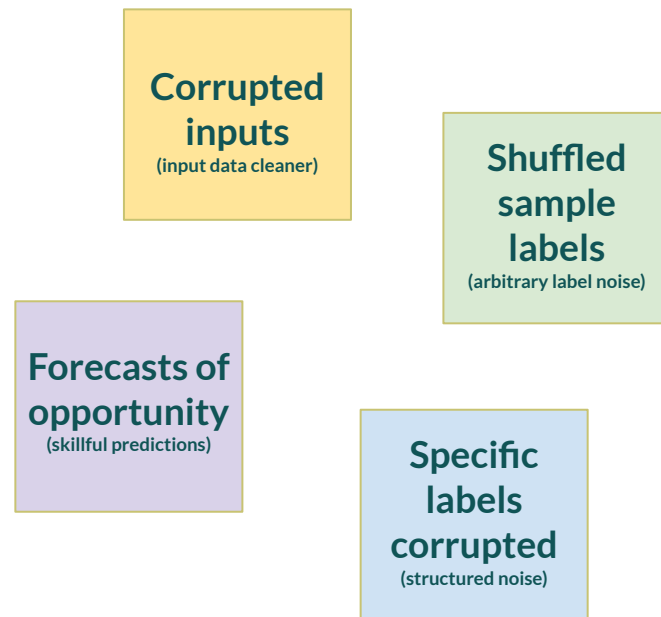
- Train abstention network for different abstention setpoints
- The best CAN models are always better (lower error) than the best baseline ANN



— Abstention outperforms baseline

- Train abstention network for different abstention setpoints
- The best CAN models are always better (lower error) than the best baseline ANN

CAN outperforms baseline networks



Forecasts of Opportunity

*aka "State-Dependent Predictability"

Leveraging “Forecasts of Opportunity”

1. This is more than just uncertainty quantification and more than just a post-processing application.

it is worthwhile for anyone working on AI for climate science to consider taking this mindset

2. Impossible predictions may be hampering learning of predictable behaviour

e.g. predicting climate variables, predicting dynamical forecast errors, etc.

could this be helpful in filtering out the “harder” predictions to train them separately?

3. May support hybrid approach to climate model parameterizations

e.g. use uncertainty measures or abstention to kick predictions to the ML or physics-based parameterizations in real-time; could this be helpful for out-of-sample climate change?

4. Utility of this concept revolves around the fact that we have a “small” amount of data to train on

if we had lots and lots of data, presumably the ML could figure out what to ignore and what to use?



A few reference links

- Mamalakis, Antonios, Imme Ebert-Uphoff and Elizabeth A. Barnes: Neural Network Attribution Methods for Problems in Geoscience: A Novel Synthetic Benchmark Dataset, submitted to IEEE Transactions on Neural Networks and Learning Systems, 03/2021, preprint available <https://arxiv.org/abs/2103.10005>.
- Barnes, Elizabeth A. and Randal J. Barnes: Controlled abstention neural networks for identifying skillful predictions for regression problems, submitted to JAMES, 04/2021, preprint available at <https://arxiv.org/abs/2104.08236>
 - https://github.com/eabarnes1010/controlled_abstention_networks
- Barnes, Elizabeth A. and Randal J. Barnes: Controlled abstention neural networks for identifying skillful predictions for classification problems, submitted to JAMES, 04/2021, preprint available at <https://arxiv.org/abs/2104.08281>
 - https://github.com/eabarnes1010/controlled_abstention_networks
- Barnes, Elizabeth A., Randal J. Barnes and Nicolas Gordillo: Adding Uncertainty to Neural Network Regression Tasks in the Geosciences, 2021: <https://arxiv.org/abs/2109.07250>
- Thulasidasan, Sunil. 2020. "Deep Learning with Abstention: Algorithms for Robust Training and Predictive Uncertainty." <https://digital.lib.washington.edu/researchworks/handle/1773/45781>.
- Thulasidasan, Sunil, Tanmoy Bhattacharya, Jeff Bilmes, Gopinath Chennupati, and Jamal Mohd-Yusof. 2019. "Combating Label Noise in Deep Learning Using Abstention." arXiv [stat.ML]. arXiv. <http://arxiv.org/abs/1905.10964>.
 - https://github.com/thulas/dac-label-noise/blob/master/dac_loss.py



Extra Slides