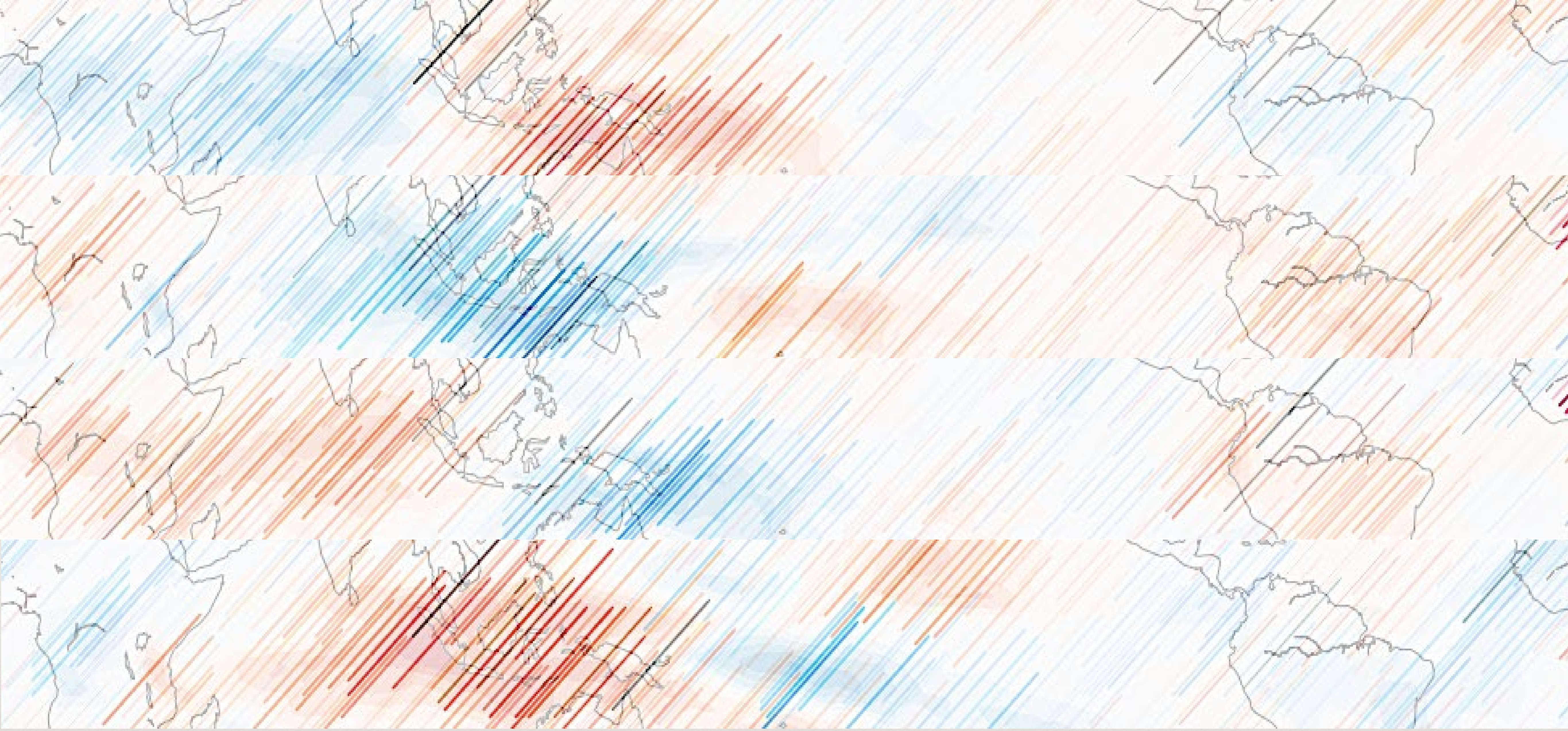


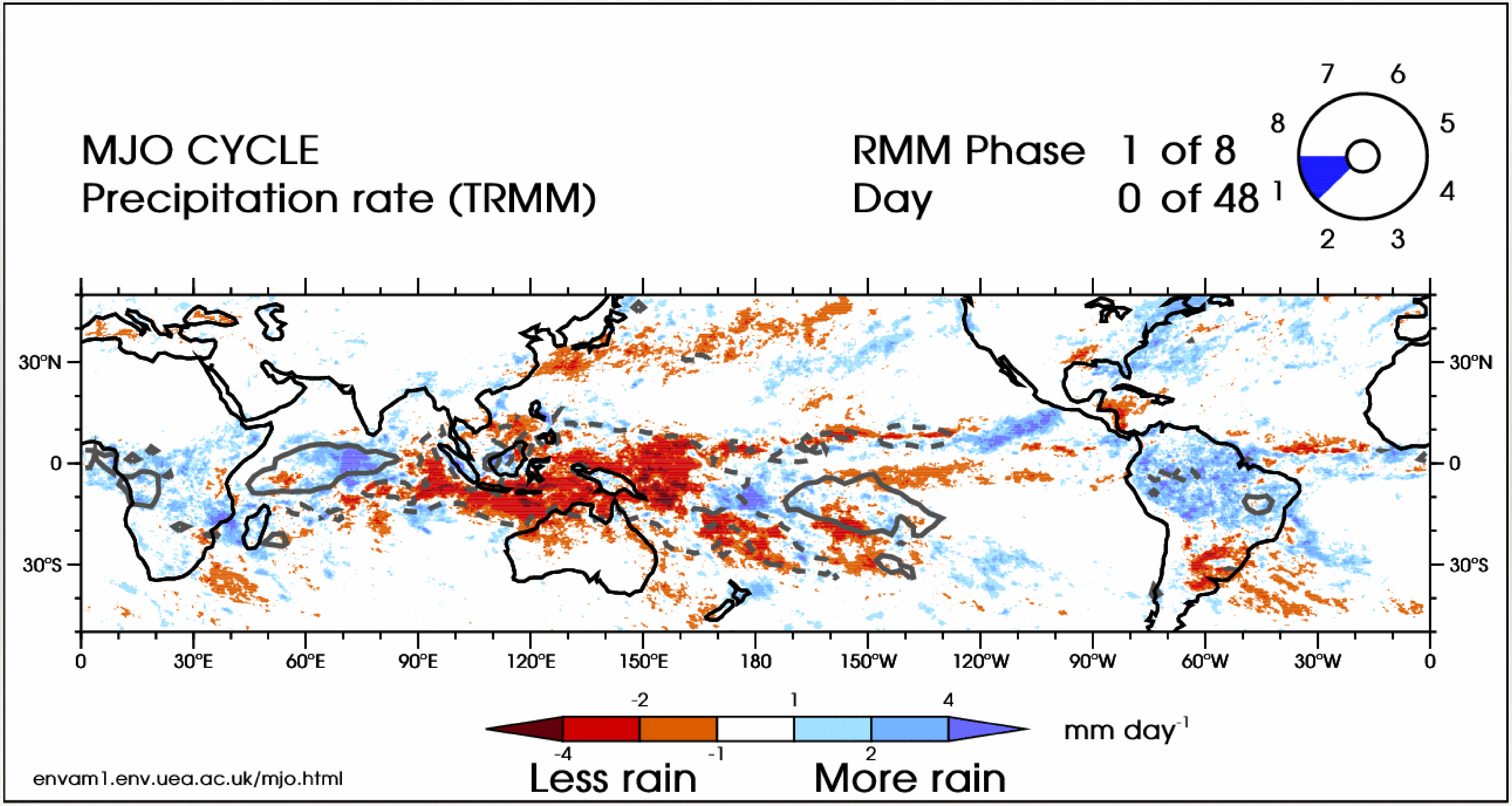
NOAA Weeks 3-4 & S2S Webinar

October 4, 2021

Using simple, explainable neural networks to predict the Madden-Julian oscillation

Zane Martin
Elizabeth Barnes
Eric Maloney

The Madden-Julian Oscillation



Summary of MJO Prediction

From 2010 up to ~now …

Statistical Models

Empirical models which use statistical techniques to predict future MJO behavior given past relationship

- ~2 weeks of skill
- Wide range of approaches, but nearly all methods linear

Dynamical Models

(Global) models which solve fluid dynamic & related equations, initialized from observations

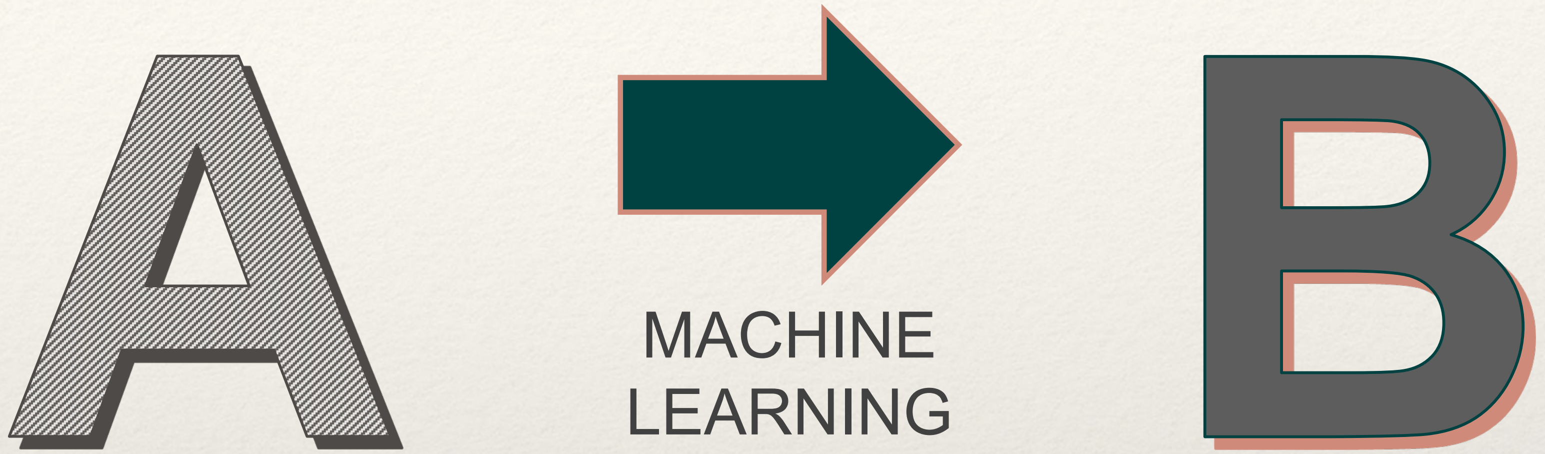
- 3-5 weeks of skill
- Since 2000s, model processes improved, ensembles grew, intercomparisons developed

Summary of MJO Prediction

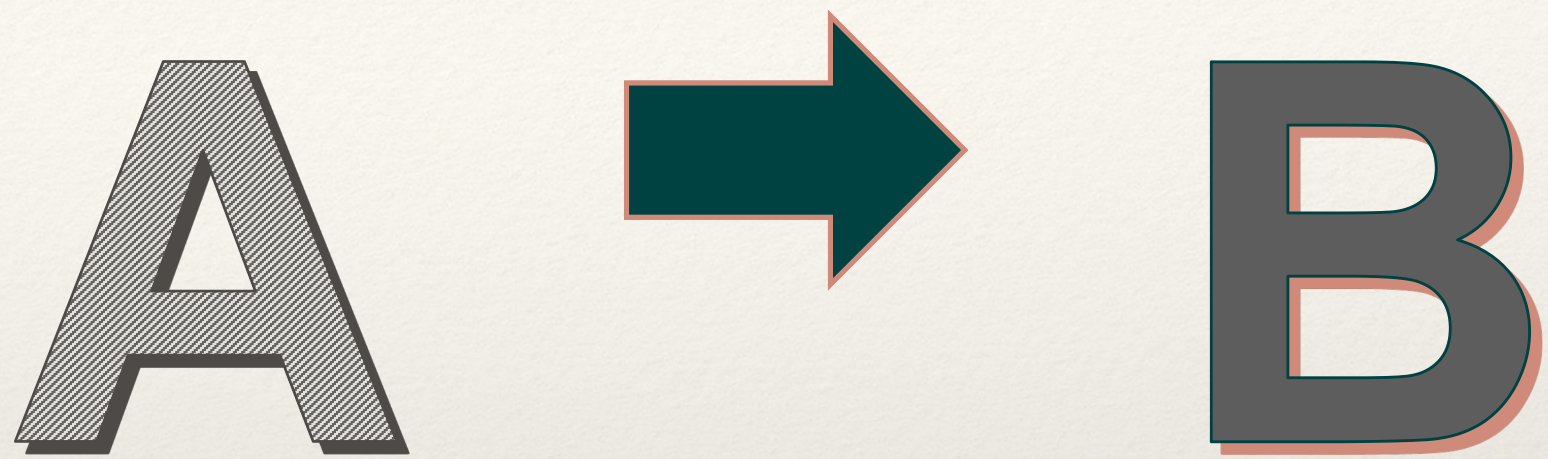
The future of statistical MJO prediction…

Non-linear methods and machine learning!

Machine Learning



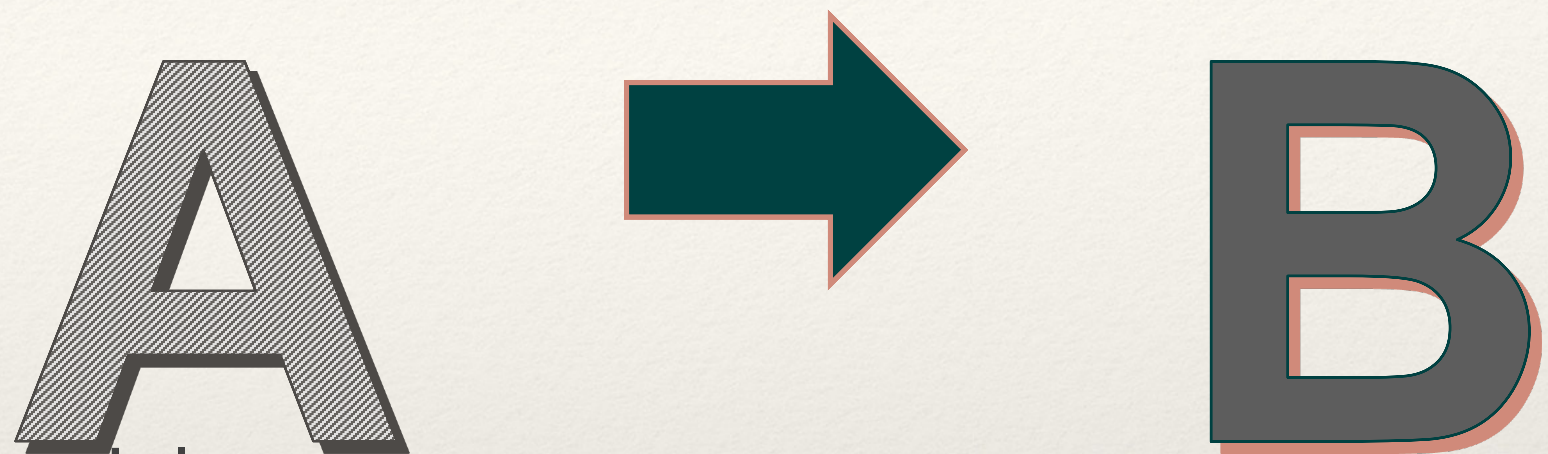
Machine Learning



A *linear* data transformation might take the form:

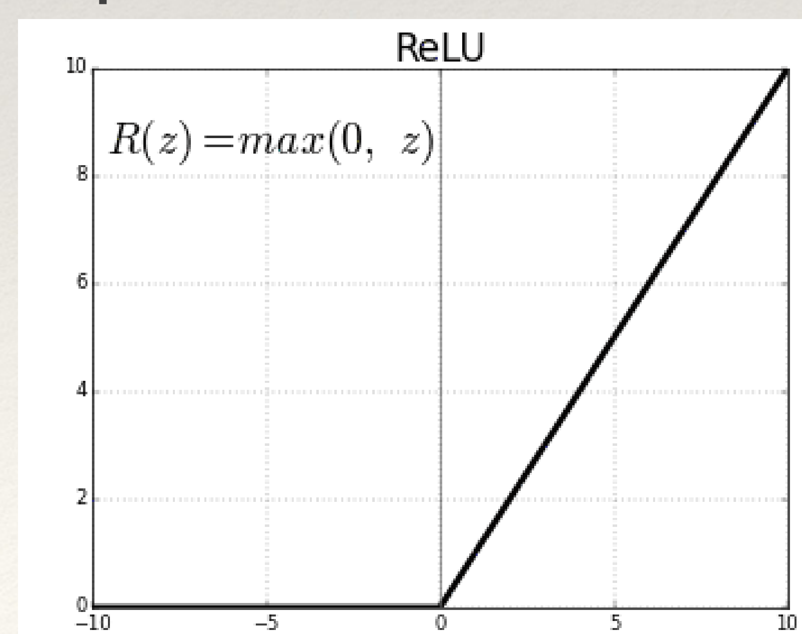
$$X \cdot A + y$$

Machine Learning



Artificial neural network models:

- Incorporate non-linearity into the data transformations
- Iterate over data during “training” data to minimize a *loss function* that describes how skillful the model prediction are



Machine Learning & the MJO

Maps of the key daily tropical variables (pre-processed)

MACHINE LEARNING

Information about an **MJO index** at various leads

Machine Learning & the MJO

Maps of the key daily tropical variables (pre-processed)

MACHINE LEARNING

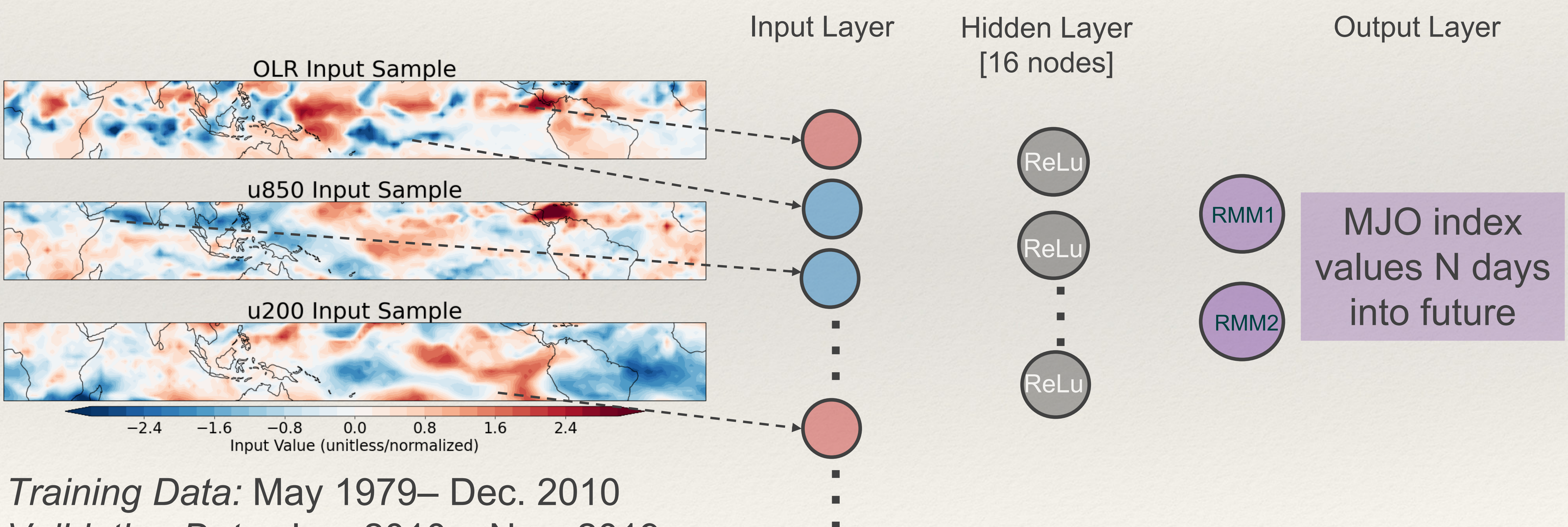
Information about an **MJO index** at various leads

We explore **2 machine learning frameworks** for MJO prediction: one deterministic and one probabilistic

Machine Learning & the MJO

Regression Model

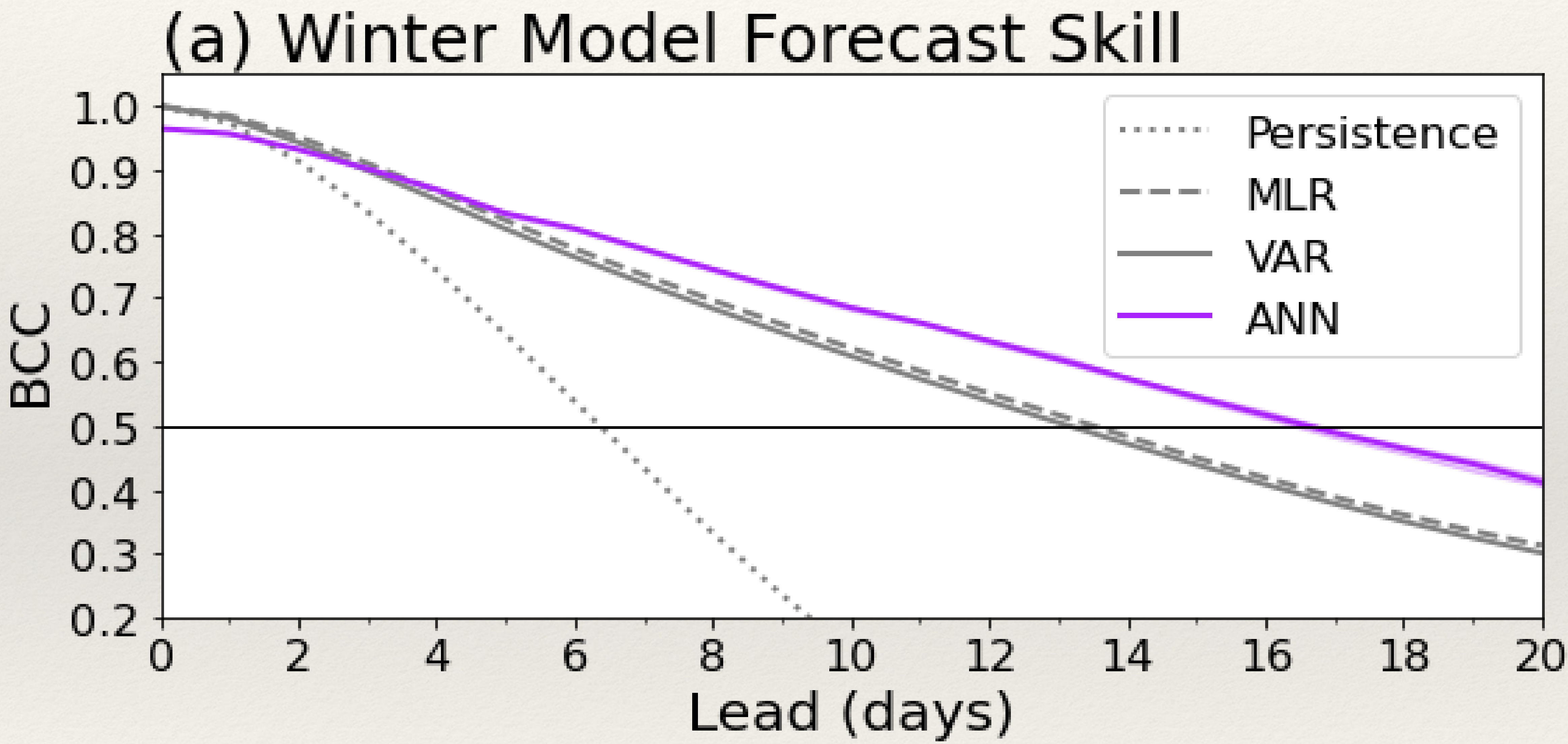
- Deterministic model which outputs *numerical values*



Training Data: May 1979– Dec. 2010
Validation Data: Jan. 2010 – Nov. 2019

Martin et al. (2021; submitted)

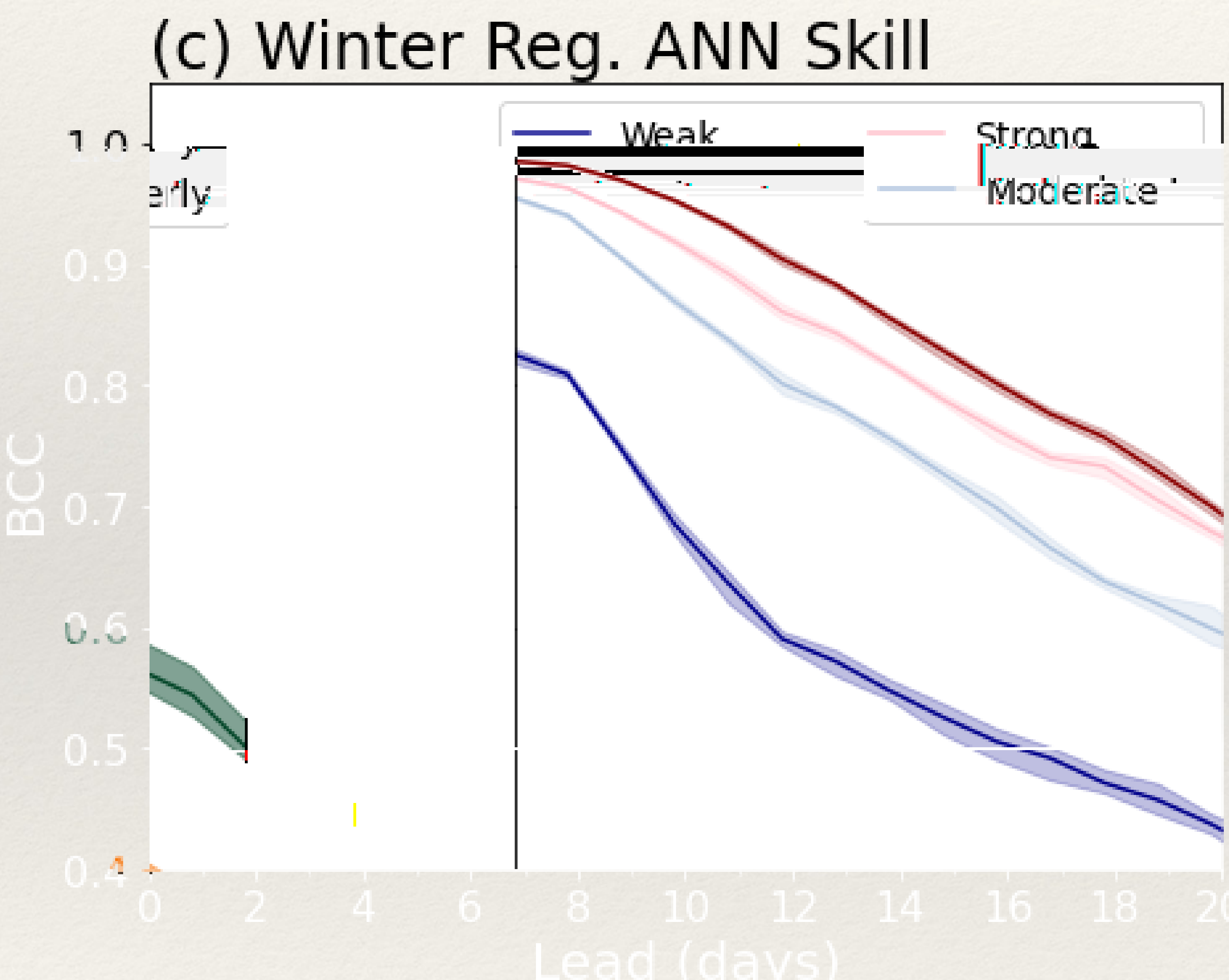
Regression ANN Model



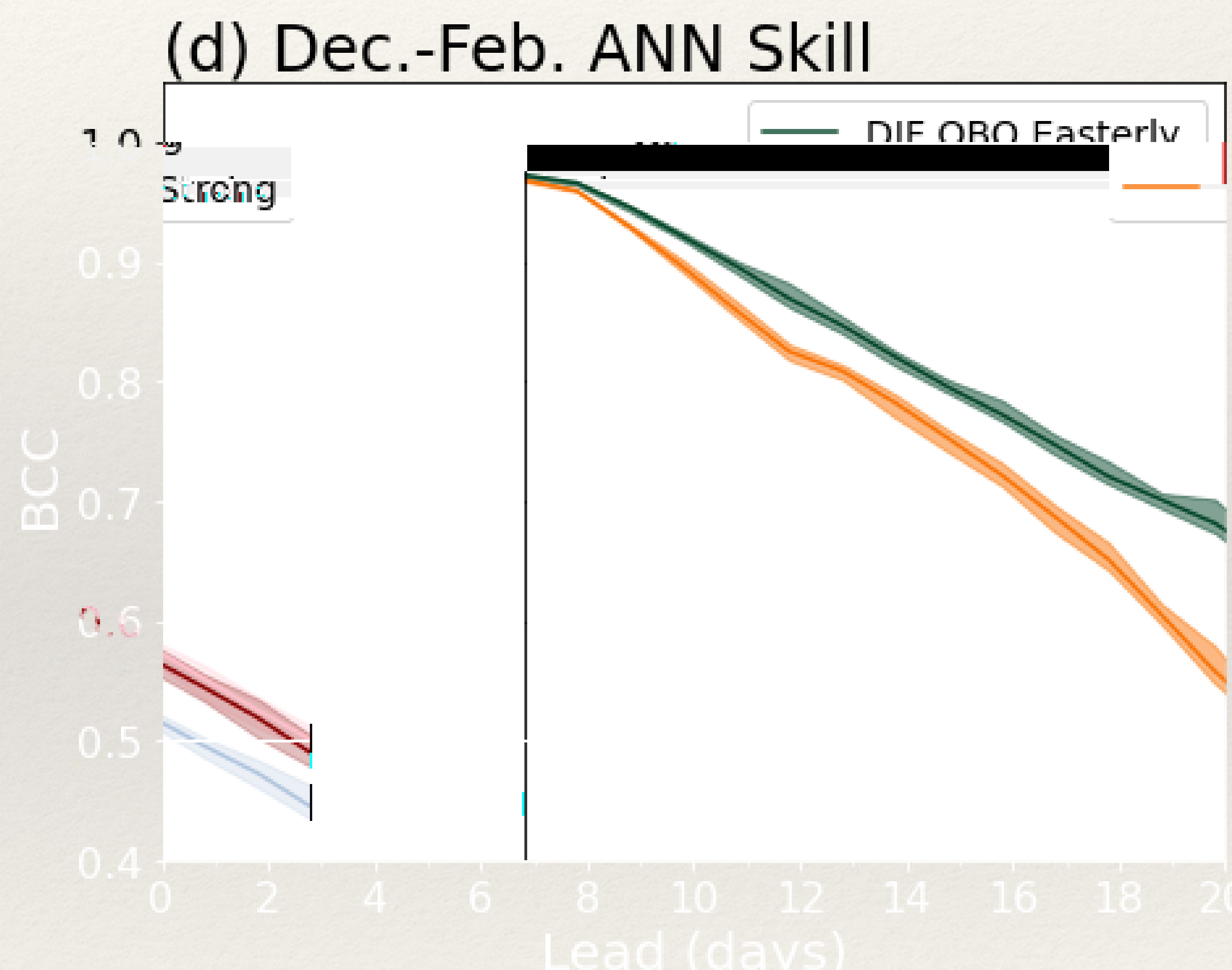
Martin et al. (2021; submitted)

Regression ANN Model

Skill vs. Initial MJO Amplitude



MJO Skill vs. Stratospheric State

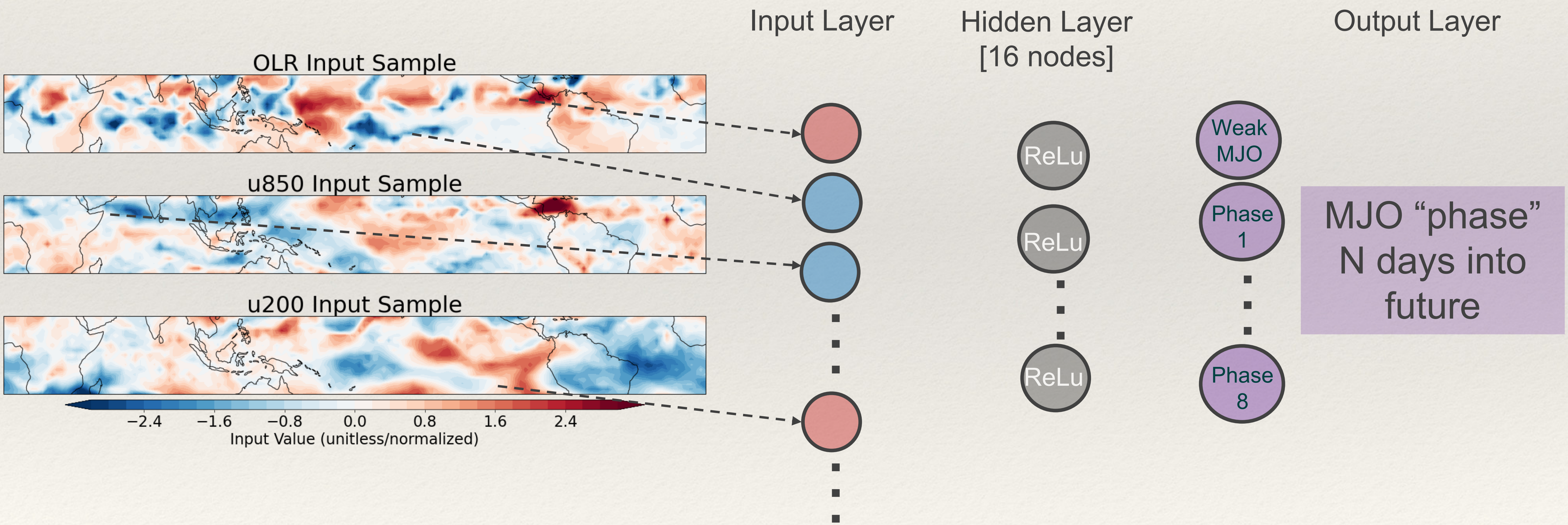


Martin et al. (2021; submitted)

Classification ANN Model

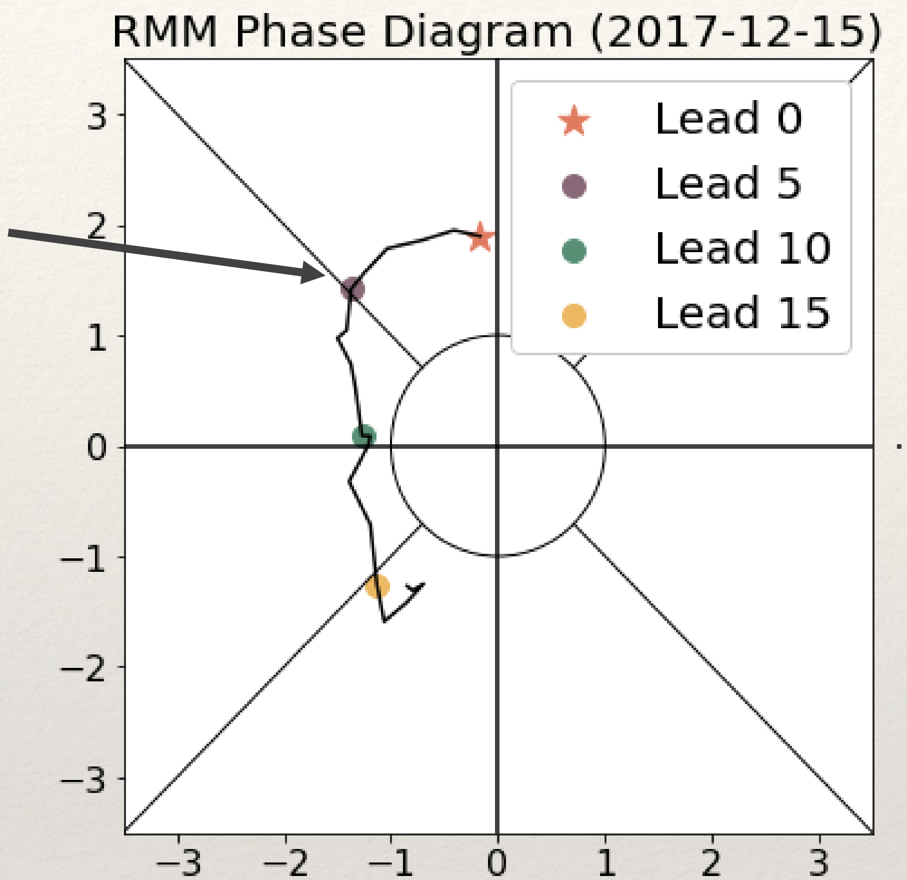
Classification Model

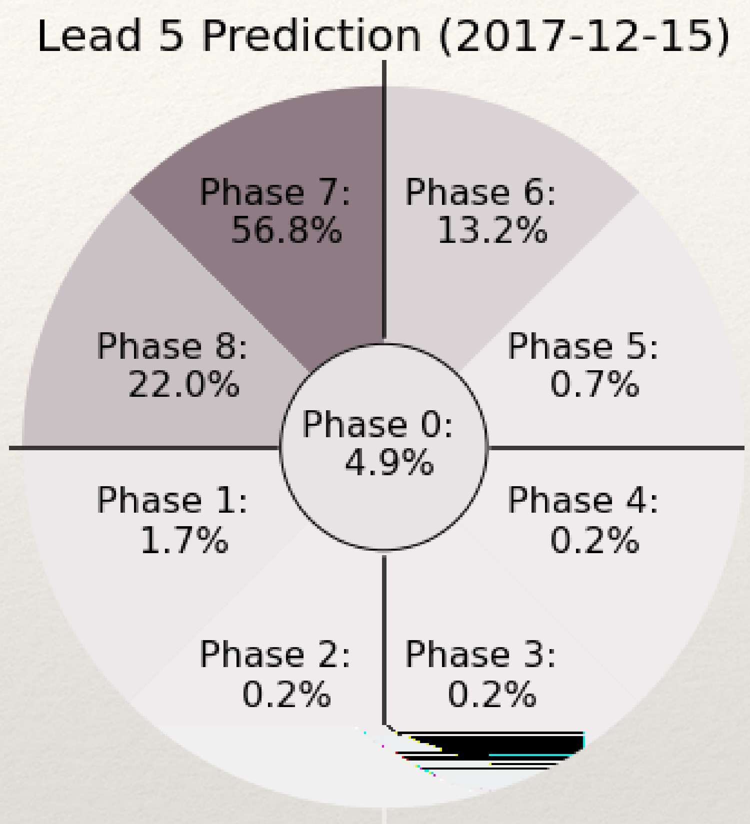
- Model which outputs the *probability* across various categories



Martin et al. (2021; submitted)

Classification ANN Model



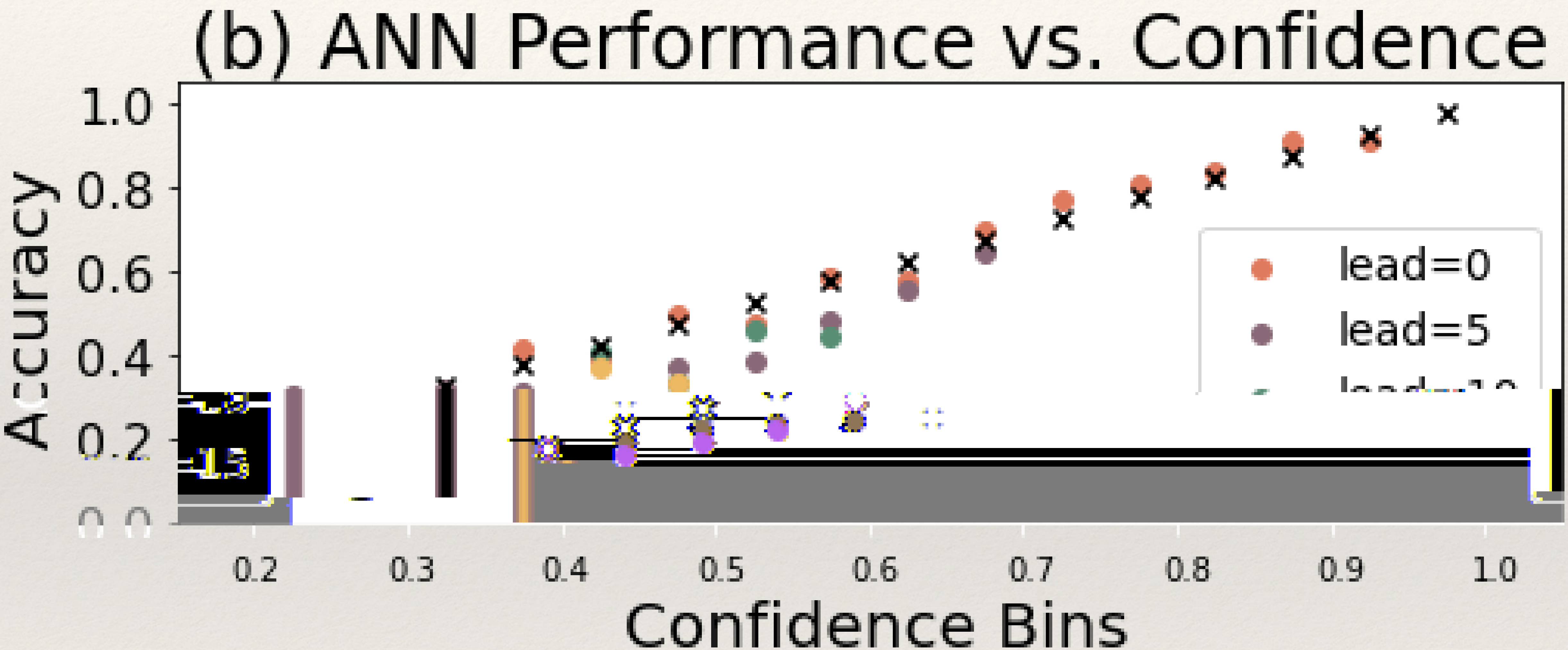


Classification Model

- A model which outputs the *probability* across various categories

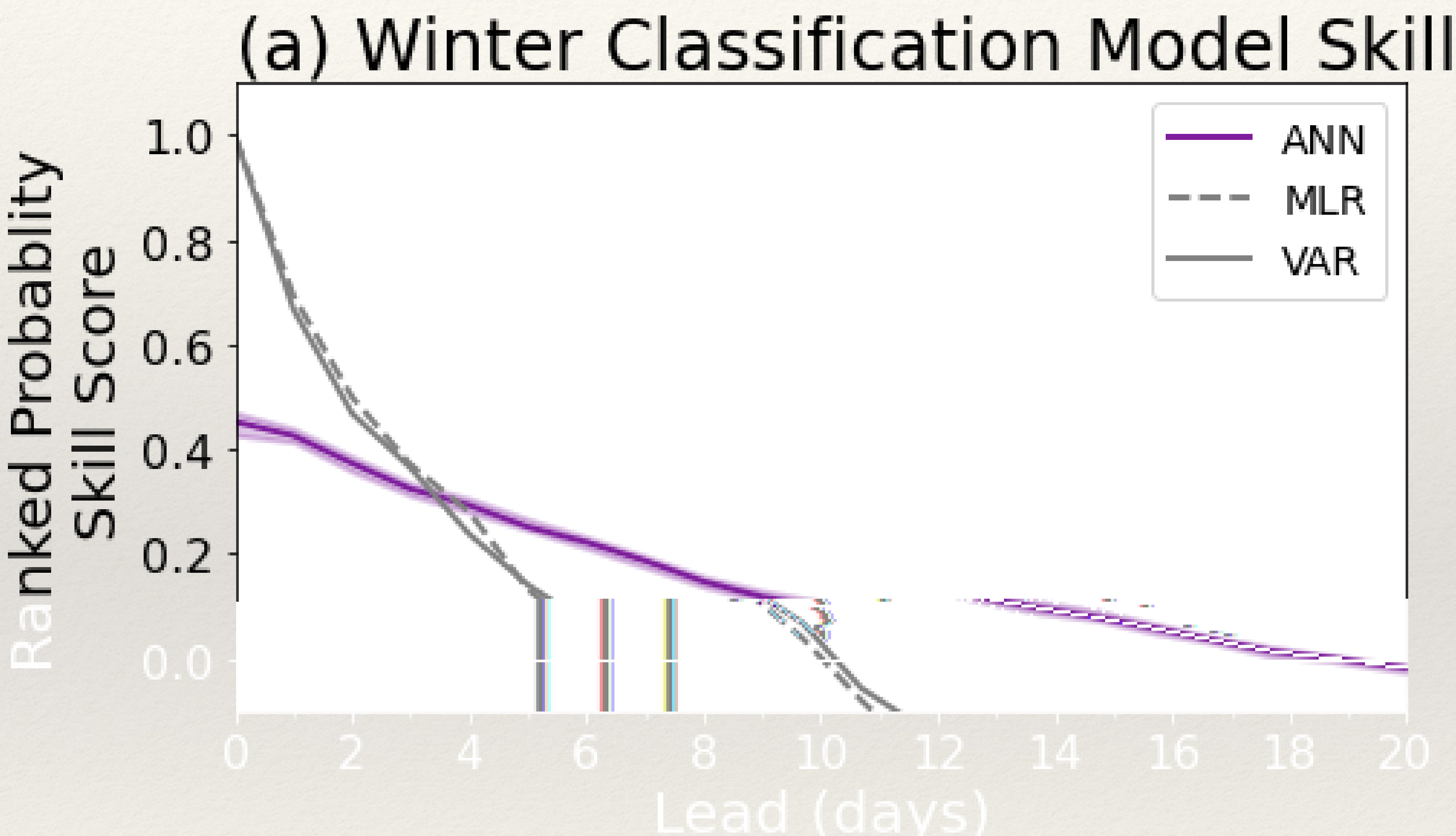
Martin et al. (2021; submitted)

Classification ANN Model



Martin et al. (2021; submitted)

Classification ANN Model



Martin et al. (2021; submitted)

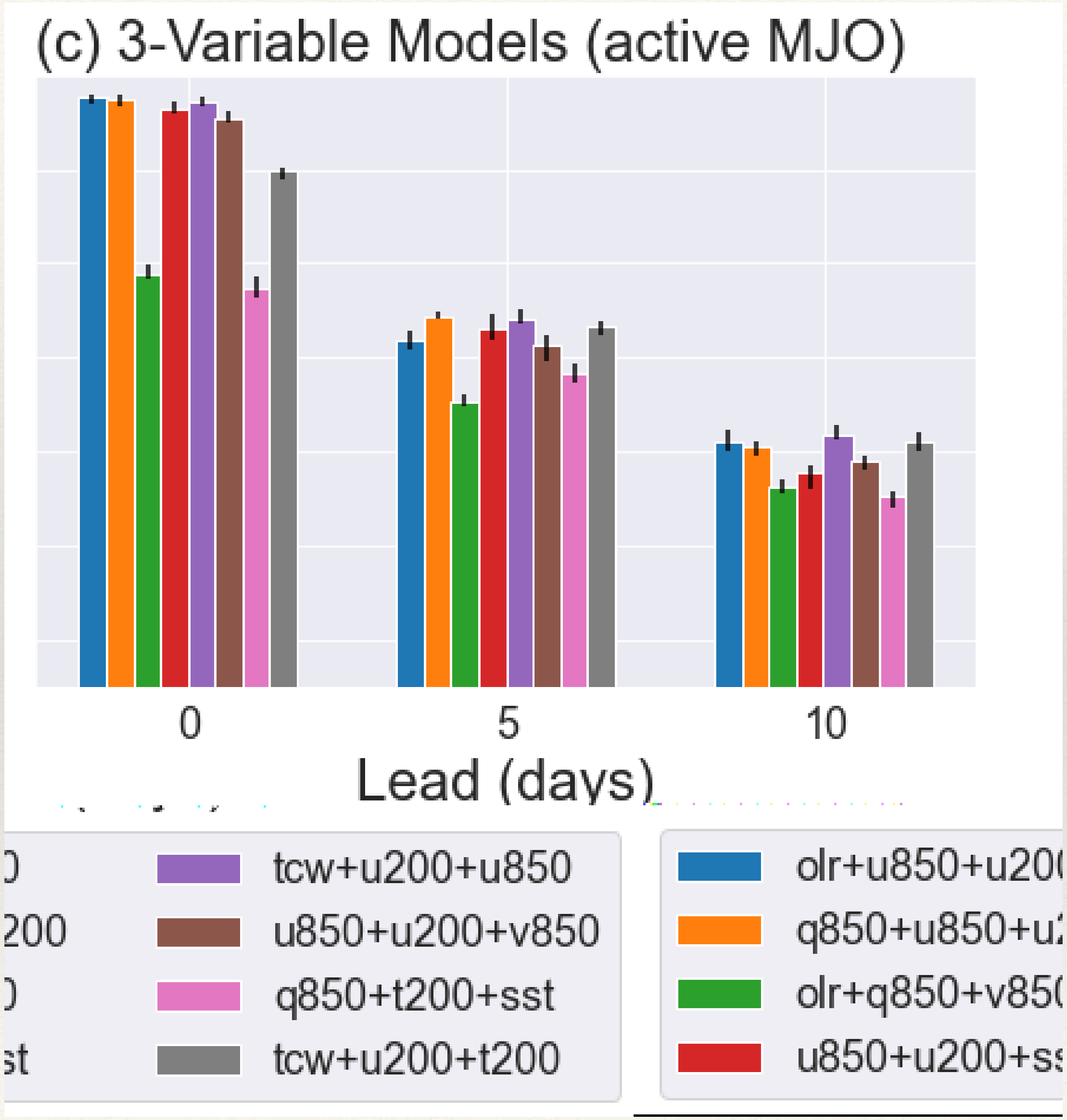
Machine Learning & the MJO

- **How skillful are ML models at predicting the MJO?**
  - ANN approaches can provide skillful MJO prediction out to past 2 weeks, better than traditional statistical models
- **How might ML be useful to study and understand the MJO, in addition to predict it?**

Thanks!

Martin et al. (2021; submitted)

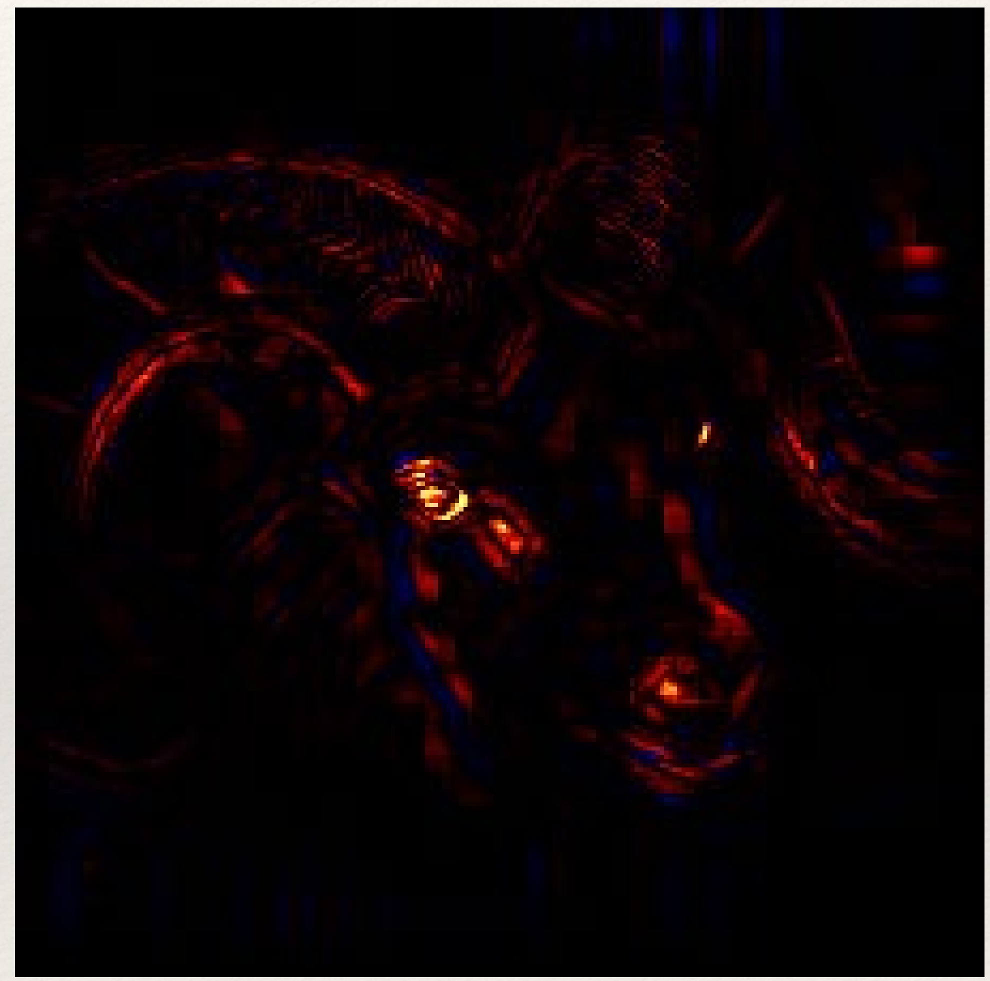
A Machine-Learning Framework for the MJO

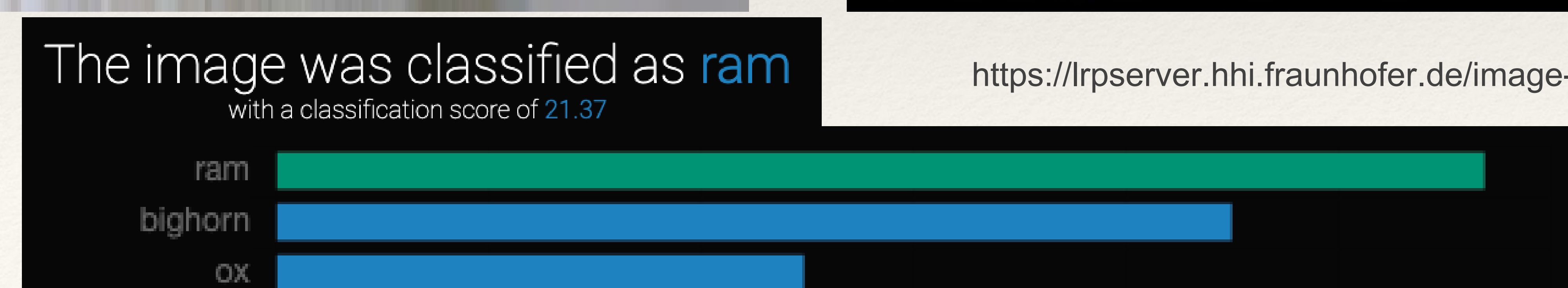


Martin et al. (2021; submitted)

A Machine-Learning Framework for the MJO

"Layerwise-relevance propagation" & other tools can help understand how the models work

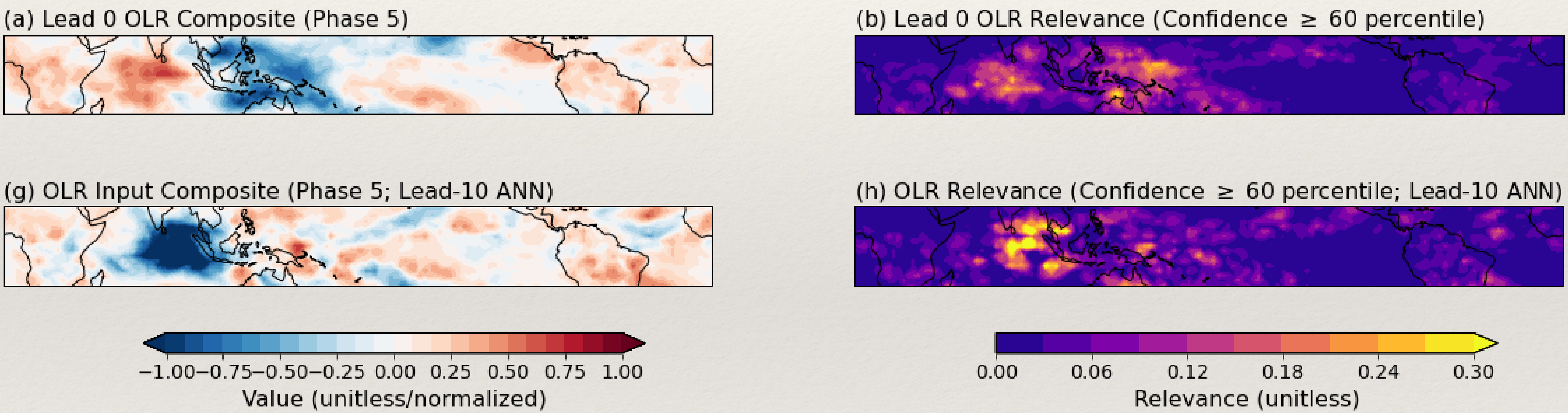




https://lrpserver.hhi.fraunhofer.de/image-classification

A Machine-Learning Framework for the MJO

"Layerwise-relevance propagation" & other tools can help understand how the models work

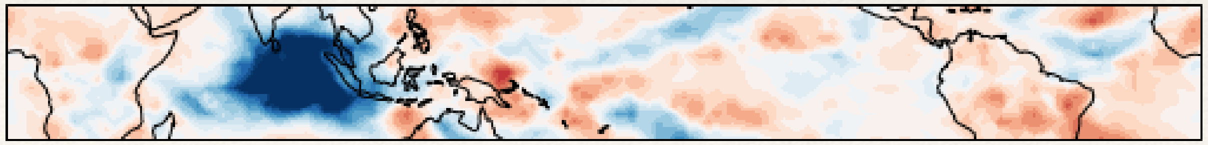


Martin et al. (2021; submitted)

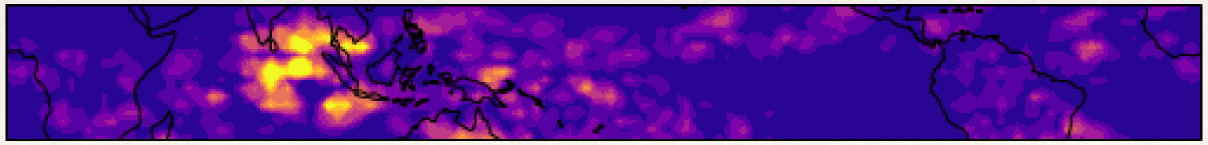
Machine Learning & the MJO

- **<u>How skillful are ML models at predicting the MJO? How might one frame MJO prediction in an ML context?</u>**
  - ANN approaches can provide skillful MJO prediction out to past 2 weeks, better than traditional statistical models
- **<u>How might ML be useful to study and understand the MJO, in addition to predict it?</u>**
  - ML models computationally efficient, flexible, and explainable
  - XAI methods & model experimentation might be useful tools to better understand sources & regions of model skill

(g) OLR Input Composite (Phase 5; Lead-10 ANN)



(h) OLR Relevance (Confidence ≥ 60 percentile; Lead-10 ANN)



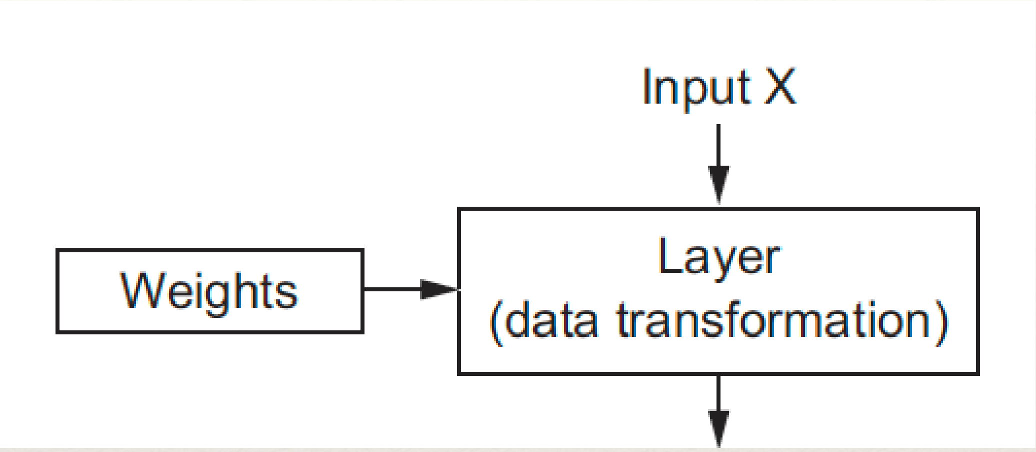
Thanks!

Martin et al. (2021; submitted)

Additional Slides

Artificial Neural Networks

Fully-connected artificial neural network



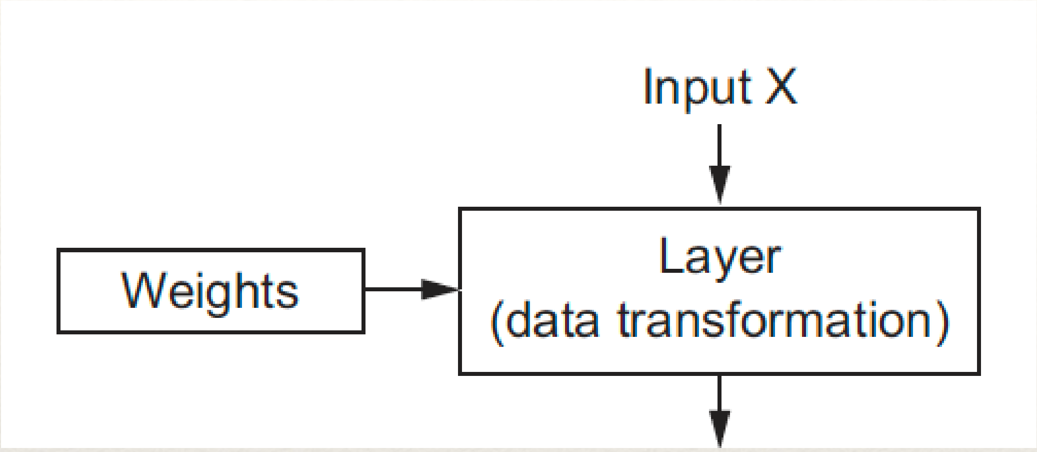
A *linear* data transformation might take the form:

$$W \cdot X + b$$

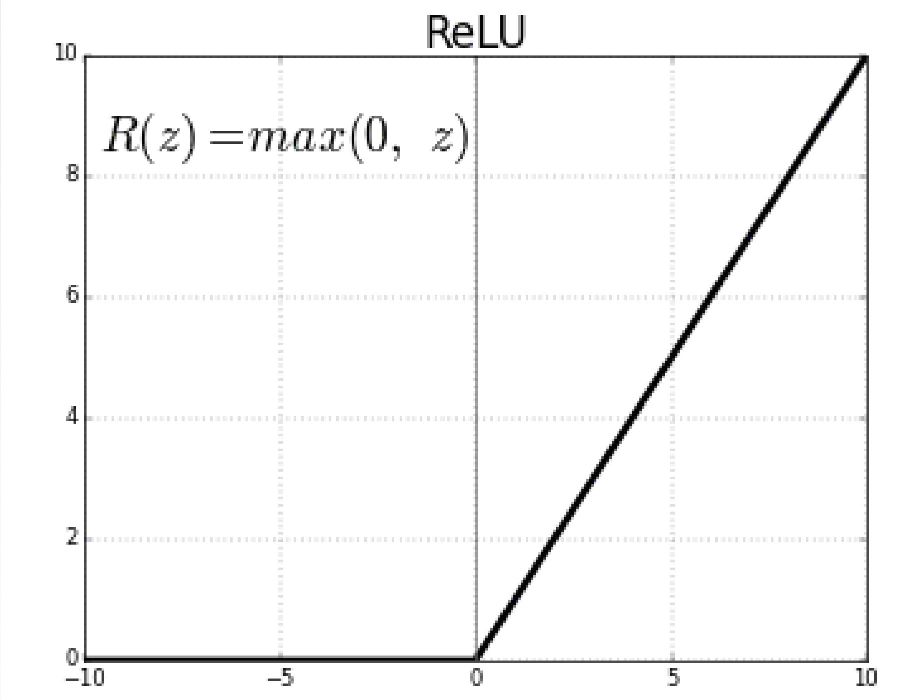
Chollet 2018

Artificial Neural Networks

Fully-connected artificial neural network

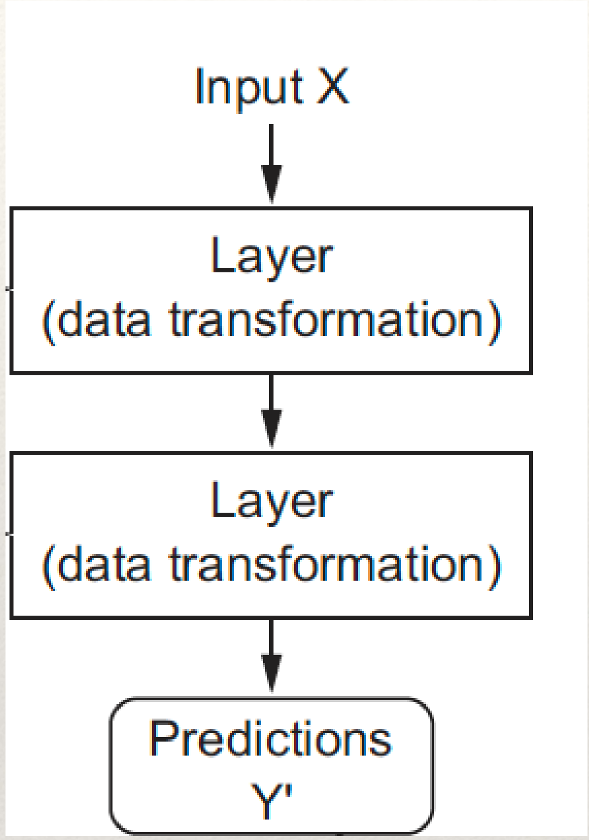


Neural network models introduce *non-linearity* into their transformations



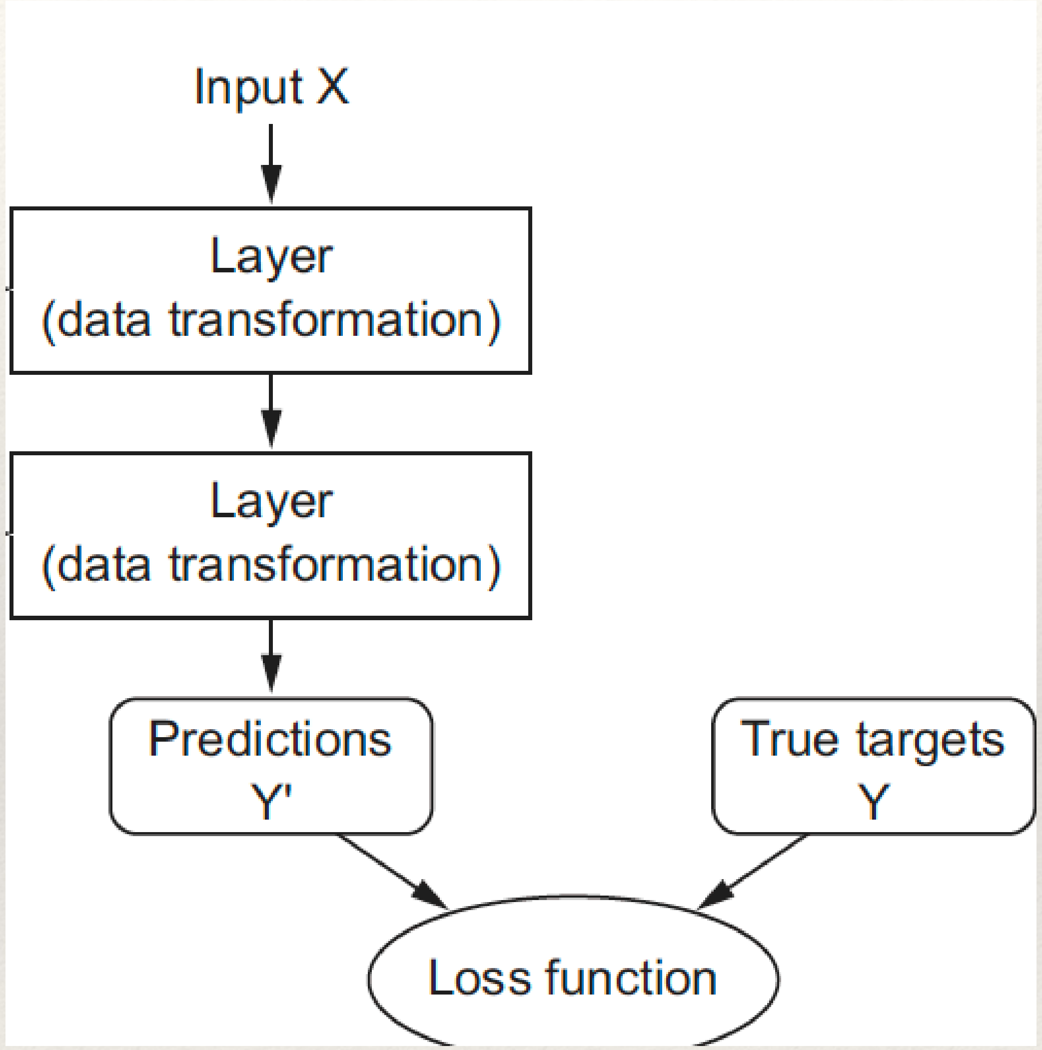
Chollet 2018

Artificial Neural Networks



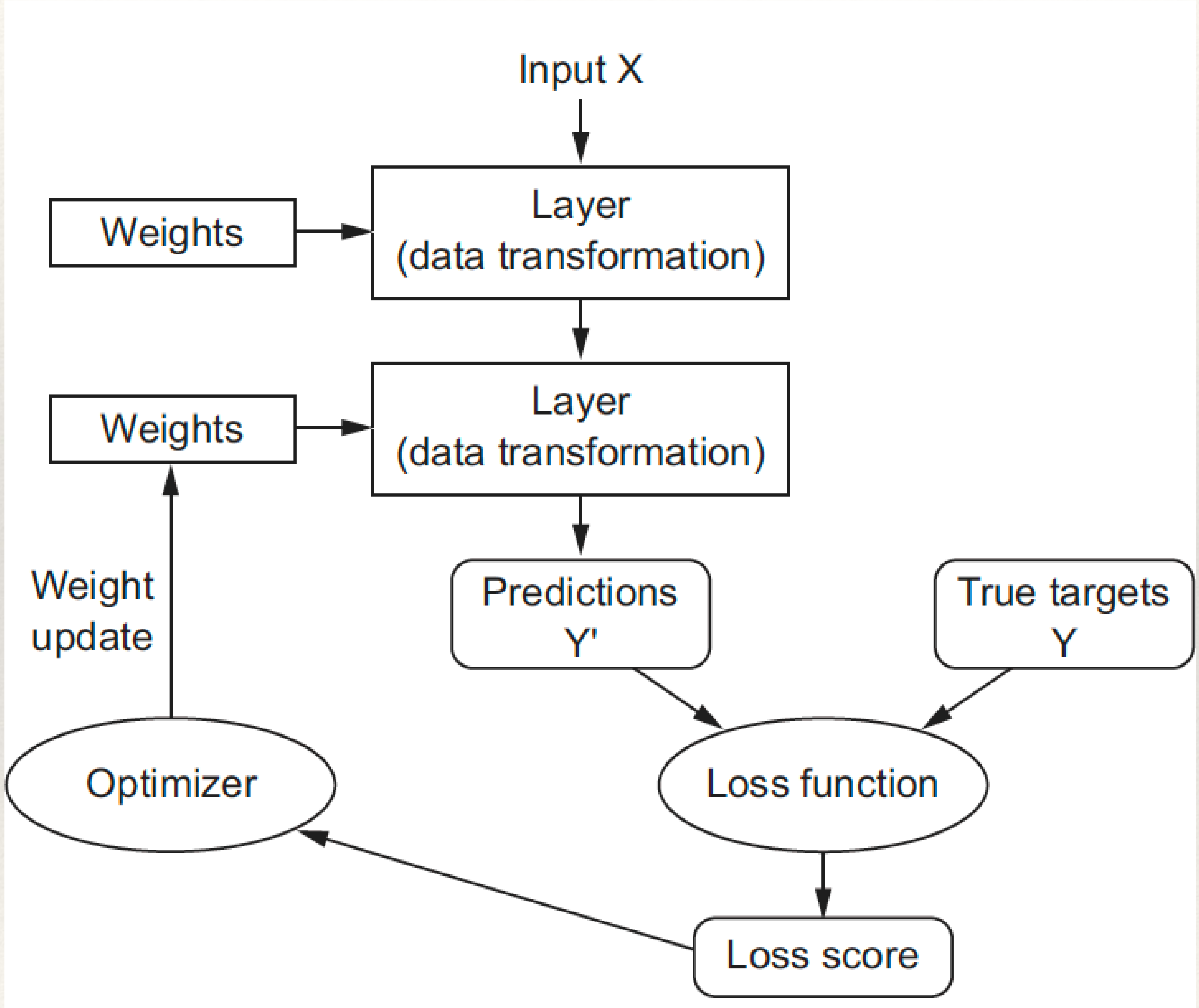
Chollet 2018

Artificial Neural Networks



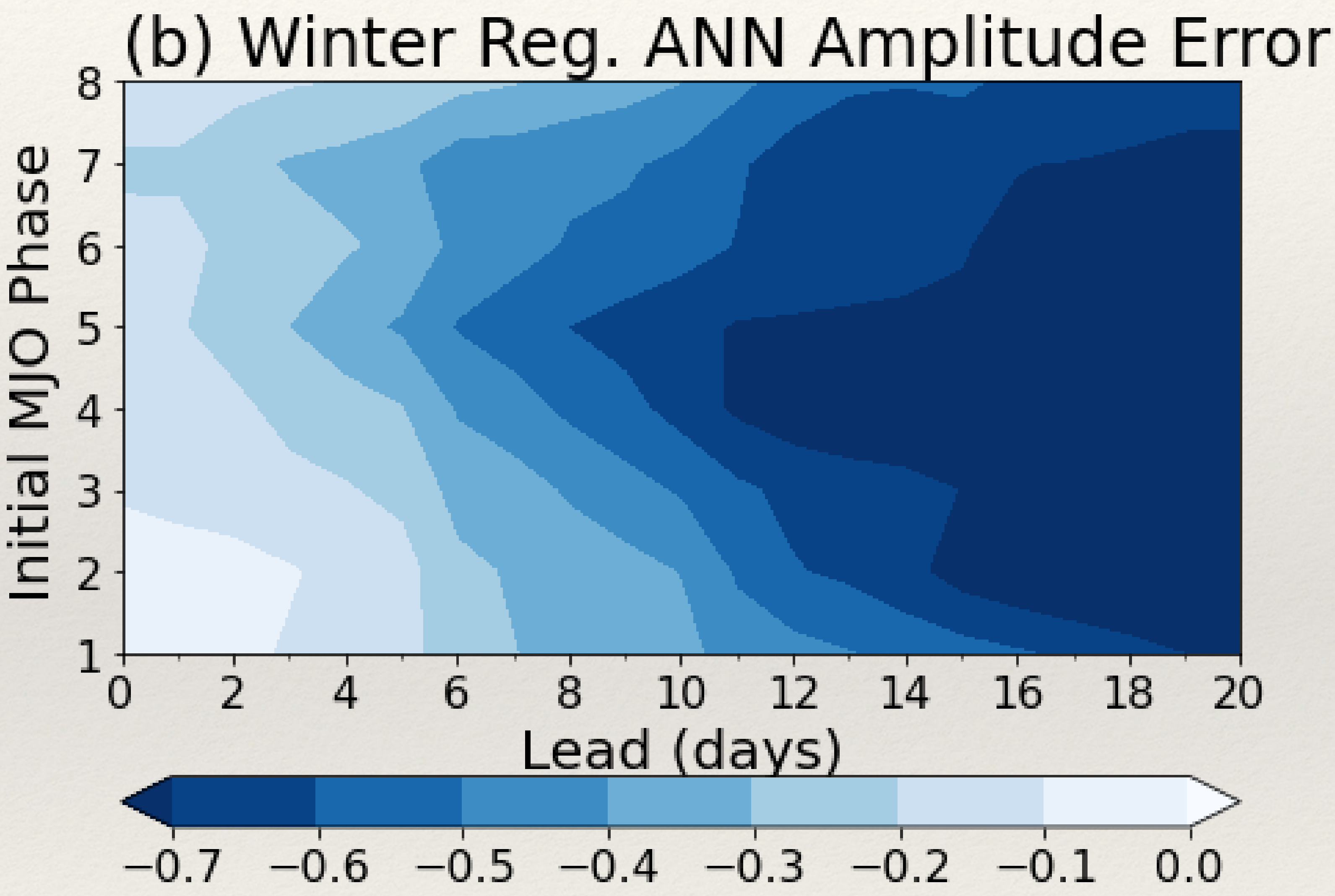
Chollet 2018

Artificial Neural Networks



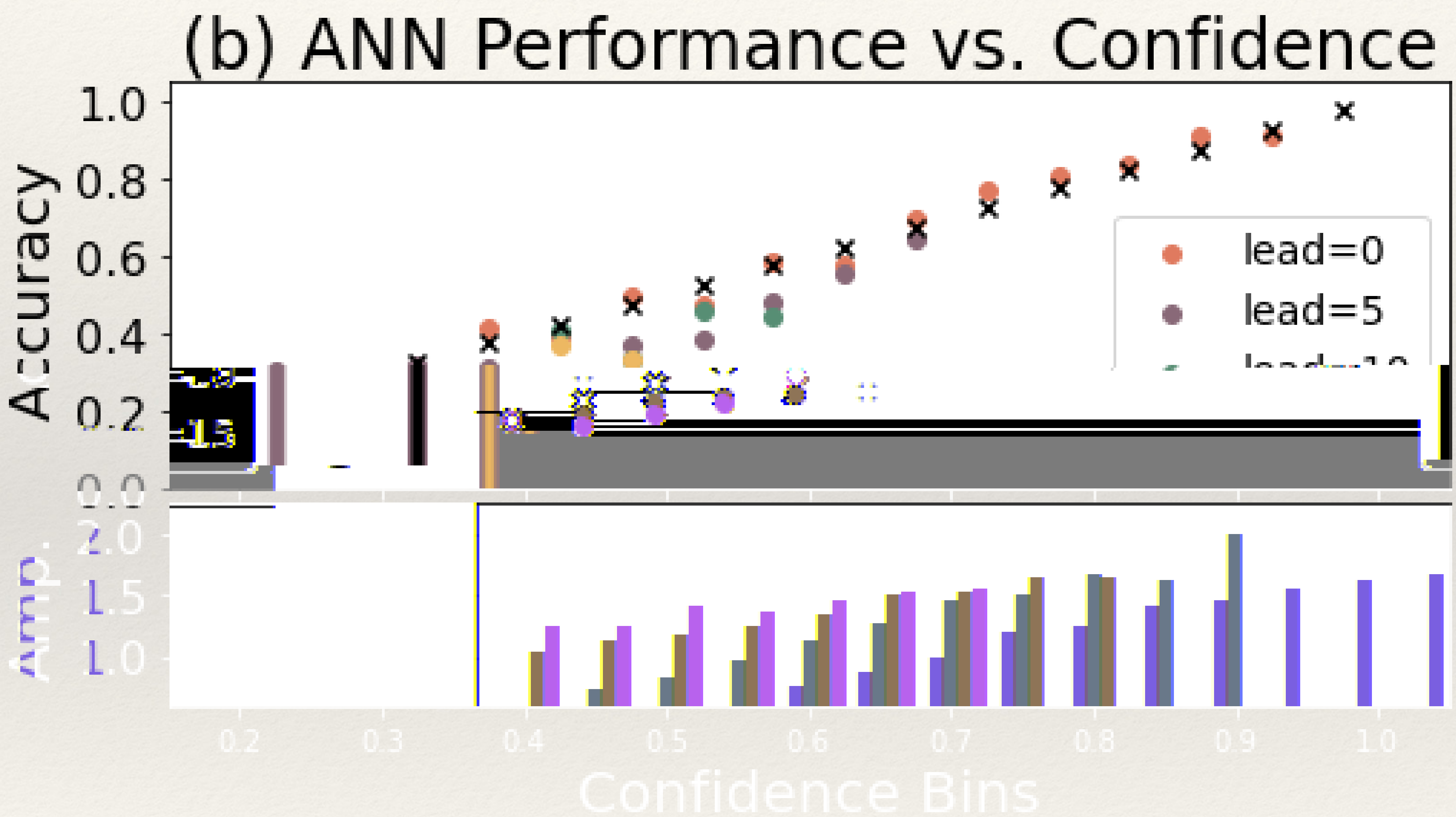
Chollet 2018

Regression ANN Model



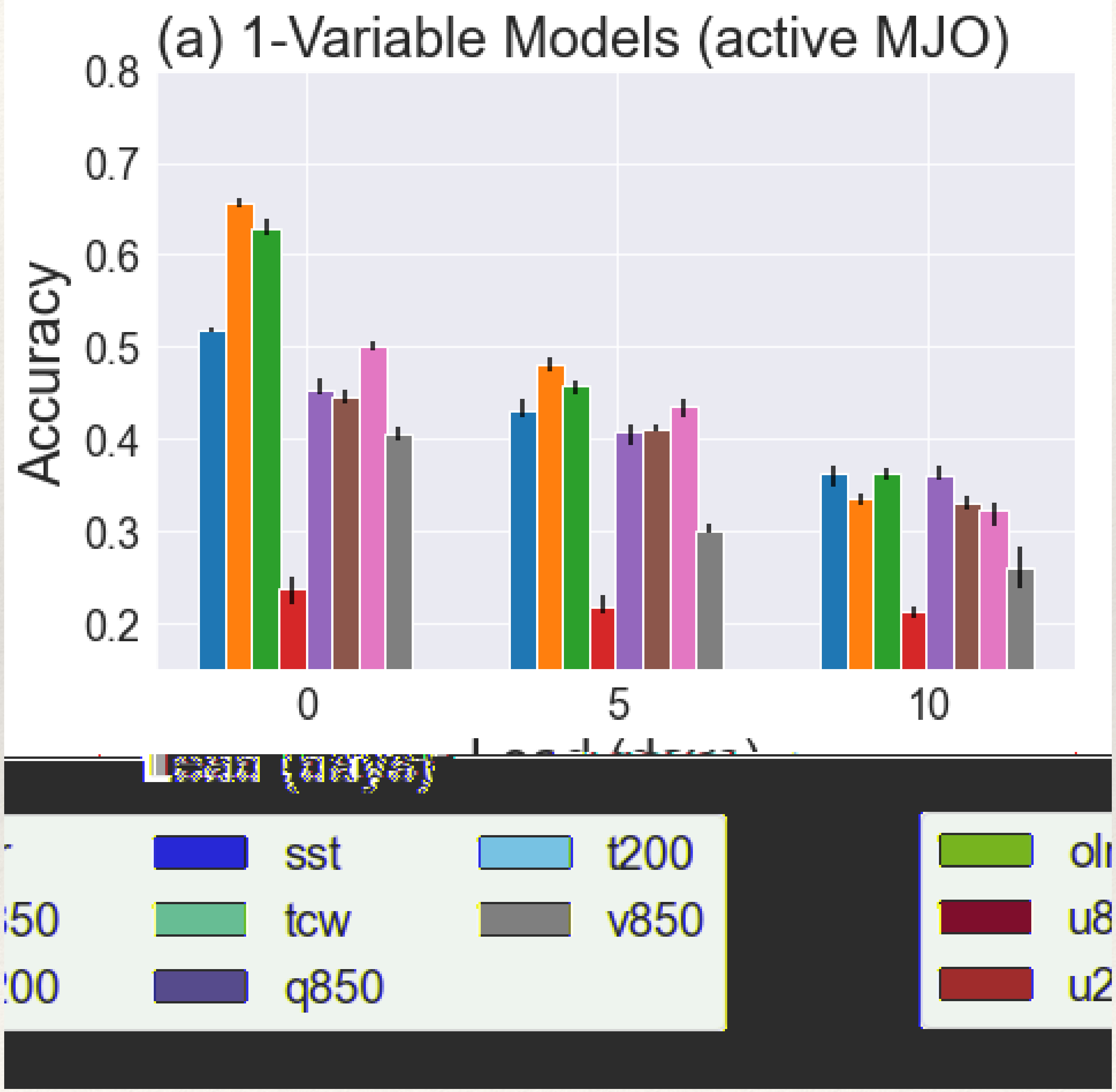
Martin et al. (2021; submitted)

Classification ANN Model



Martin et al. (2021; submitted)

A Machine-Learning Framework for the MJO

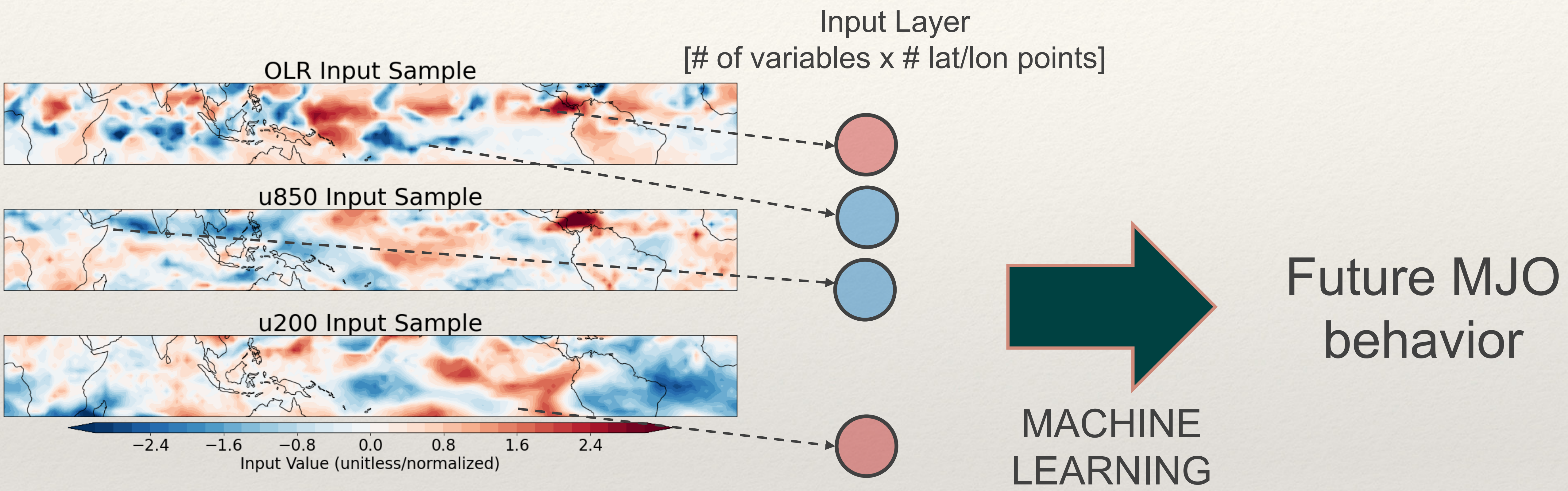


Martin et al. (2021; submitted)

Machine Learning & the MJO

- How skillful are machine learning models at predicting the MJO?
- How might ML be useful to study and understand the MJO?

Machine Learning & the MJO



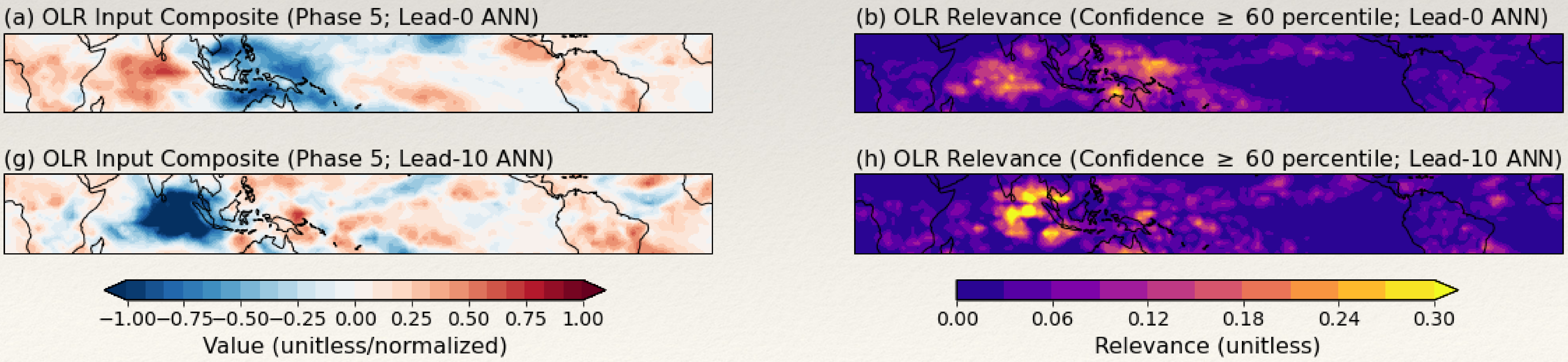
Training Data: May 1979– Dec. 2010
Validation Data: Jan. 2010 – Nov. 2019

Machine Learning & the MJO

- How skillful are machine learning models at predicting the MJO?
- How might ML be useful to study and understand the MJO?

Machine Learning & the MJO

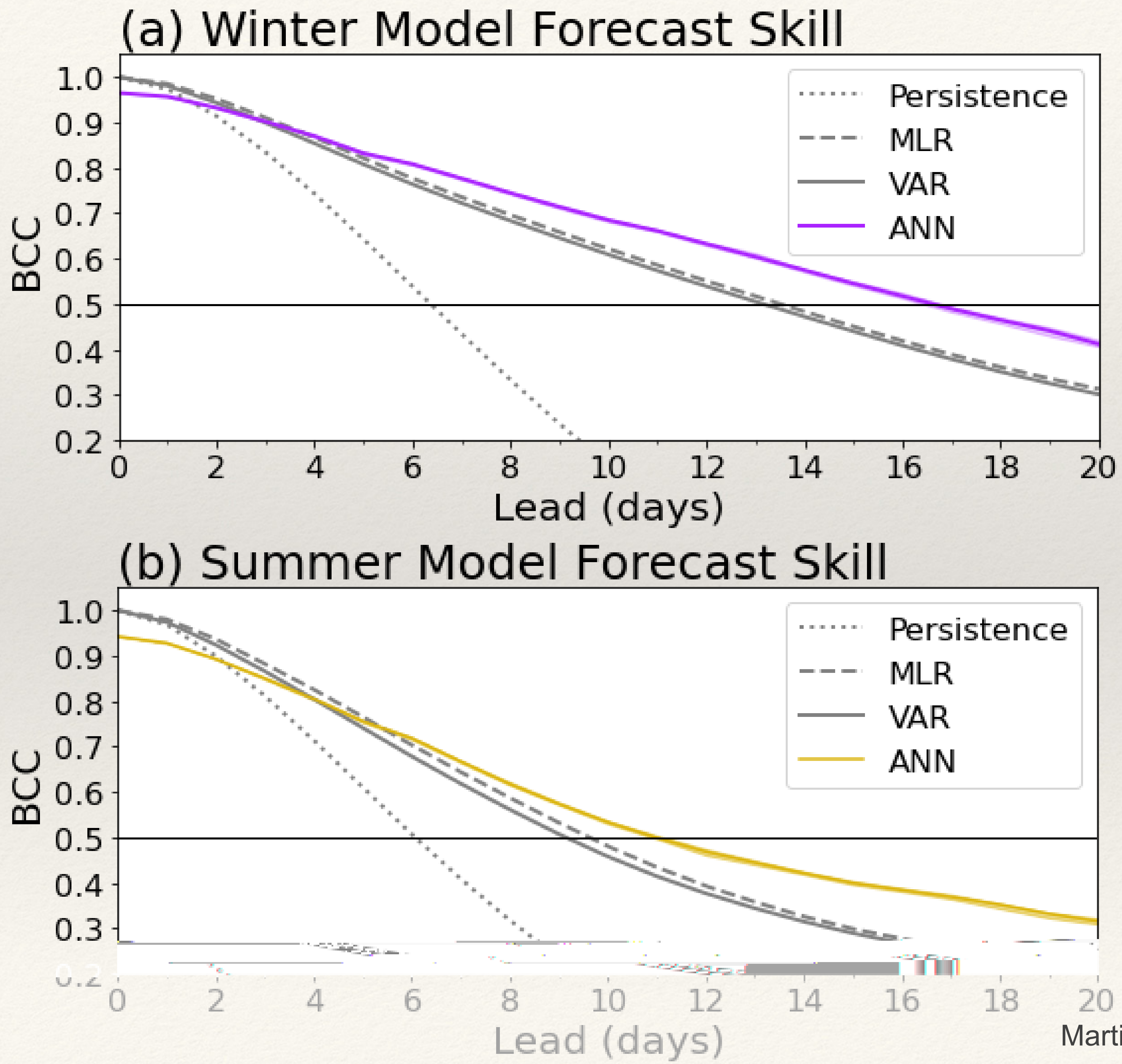
- **<u>How might ML be useful to study and understand the MJO, in addition to predict it?</u>**
  - ML models computationally efficient, flexible, and explainable
  - XAI methods & model experimentation might be useful tools to better understand sources & regions of model skill



Machine Learning & the MJO

Thanks!

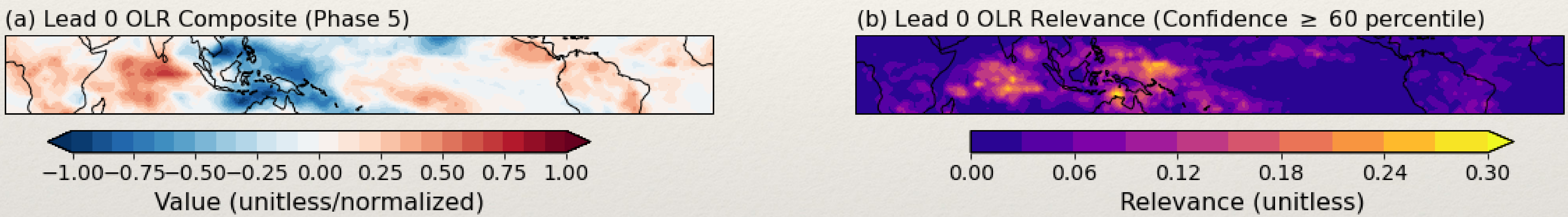
A Machine-Learning Framework for the MJO



Martin et al. (2021; submitted)

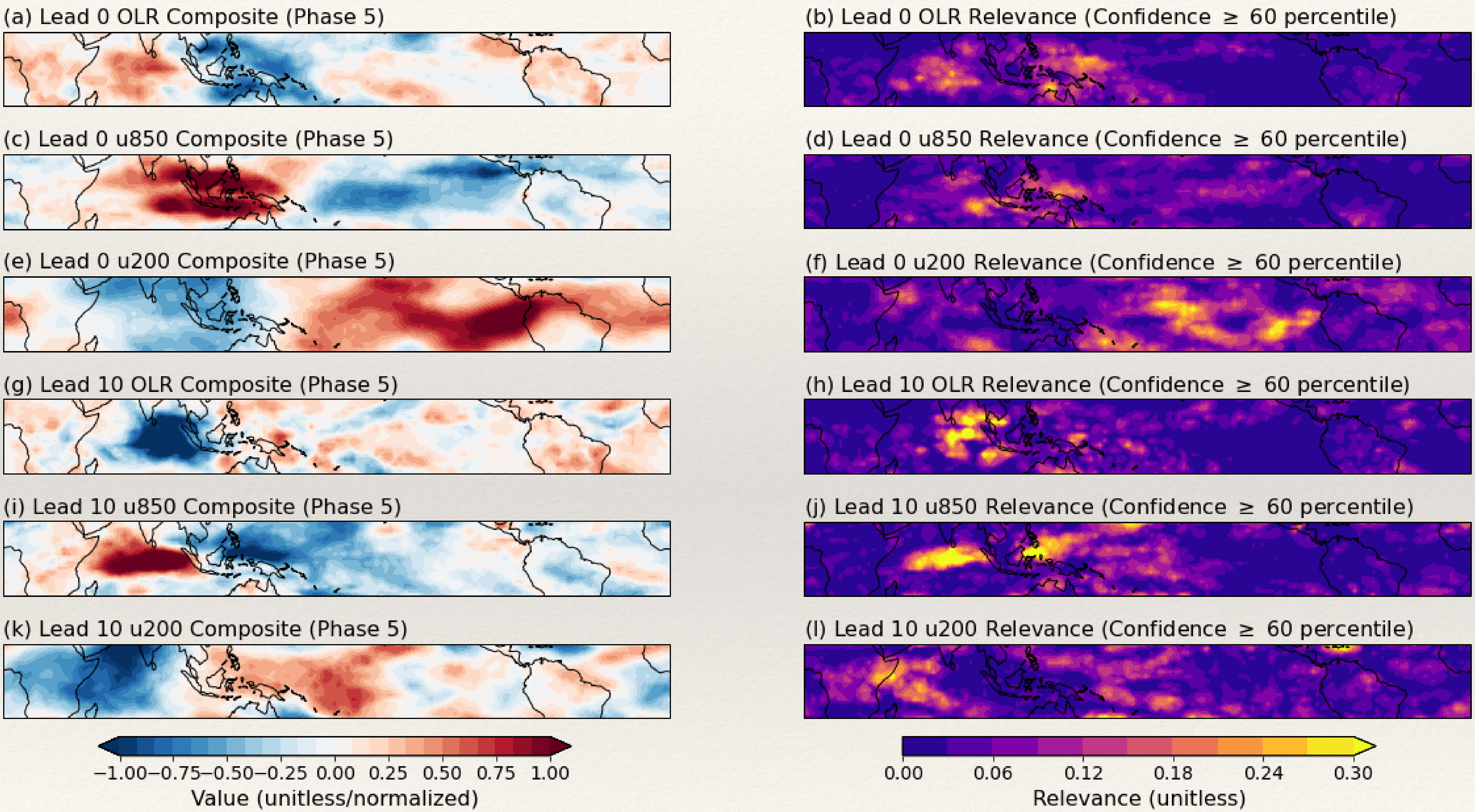
A Machine-Learning Framework for the MJO

“Layerwise-relevance propagation” & other tools can help understand how the models work



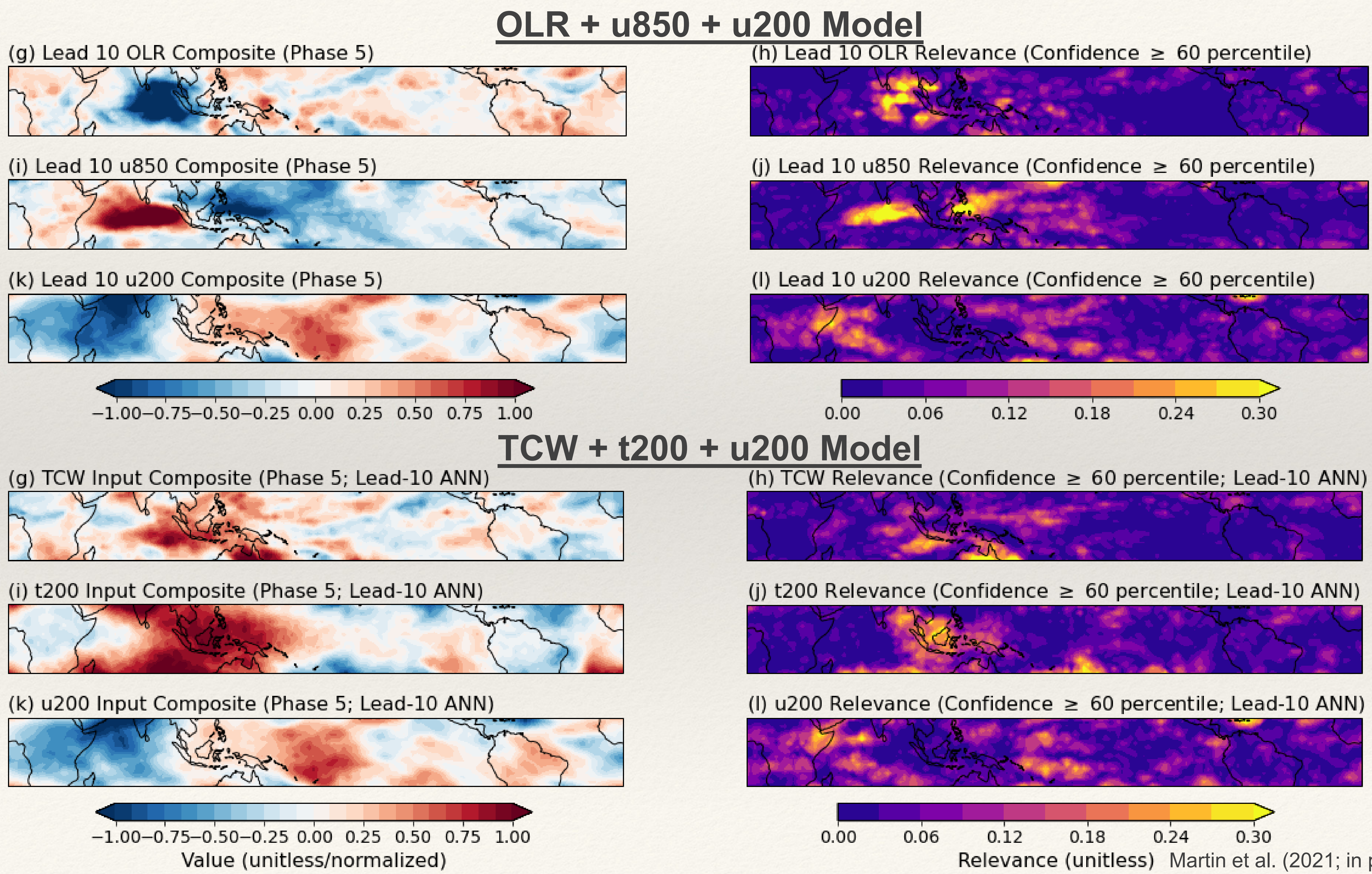
Martin et al. (2021; submitted)

A Machine-Learning Framework for the MJO



Martin et al. (2021; in prep)

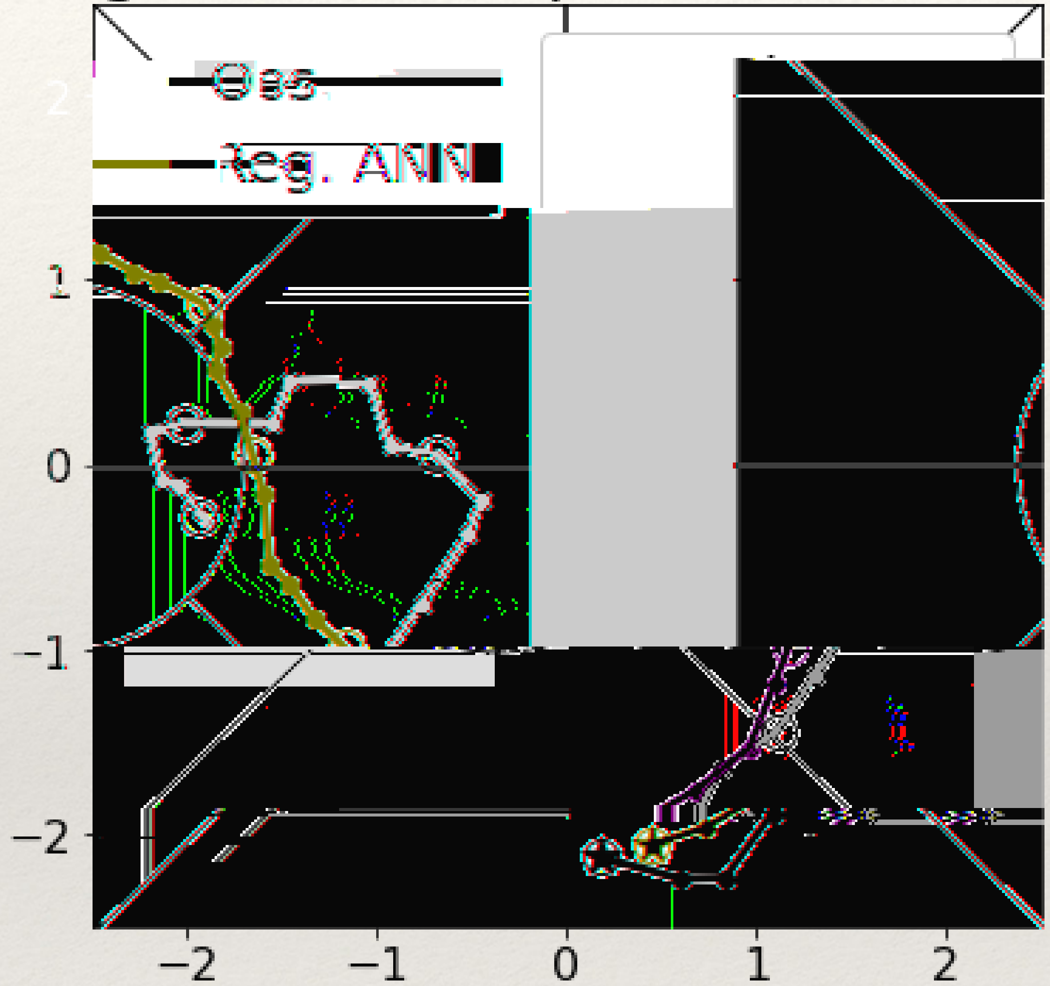
A Machine-Learning Framework for the MJO



Martin et al. (2021; in prep)

Machine Learning & the MJO

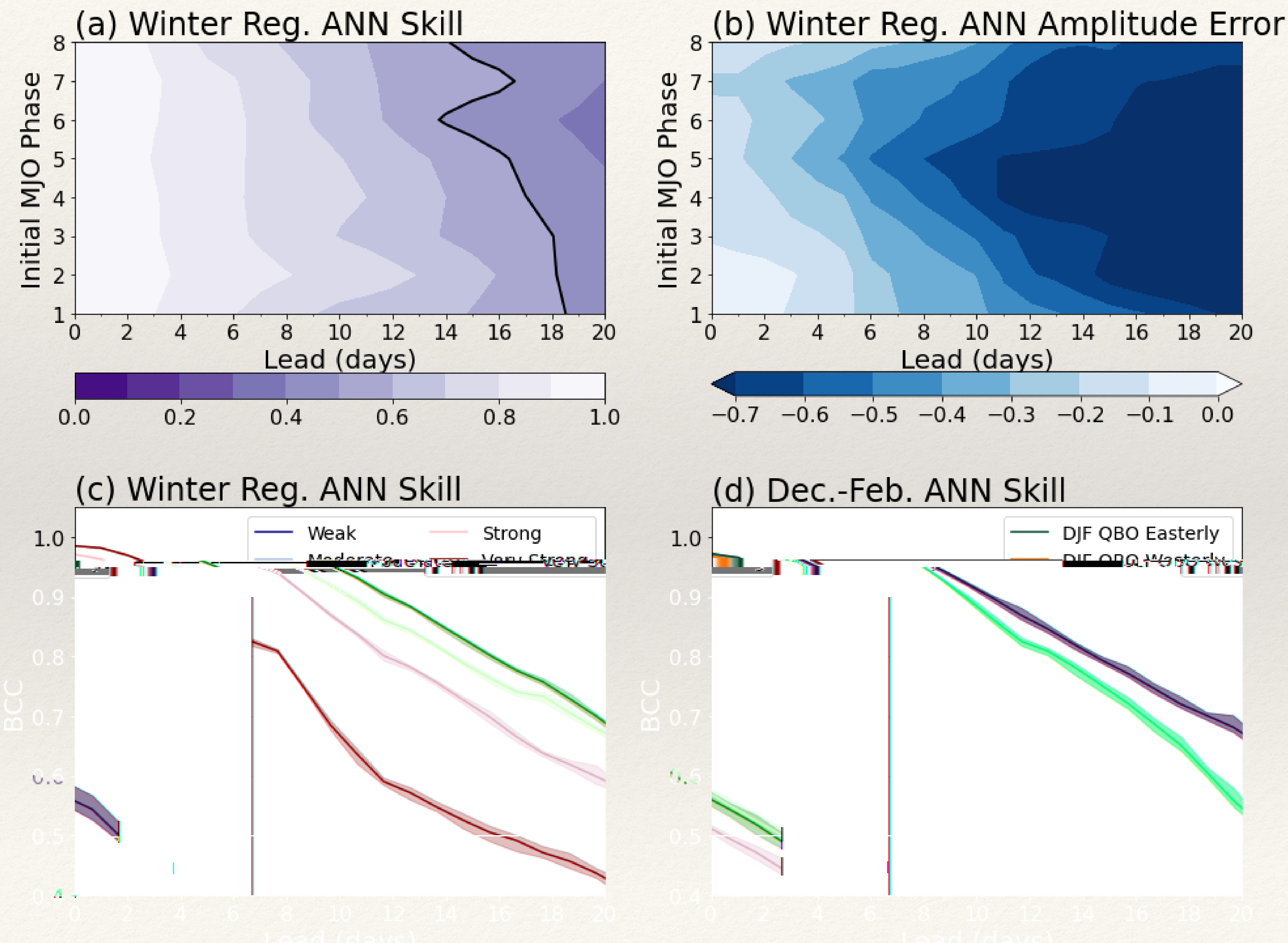
Reg. ANN Example (2011-11-26)

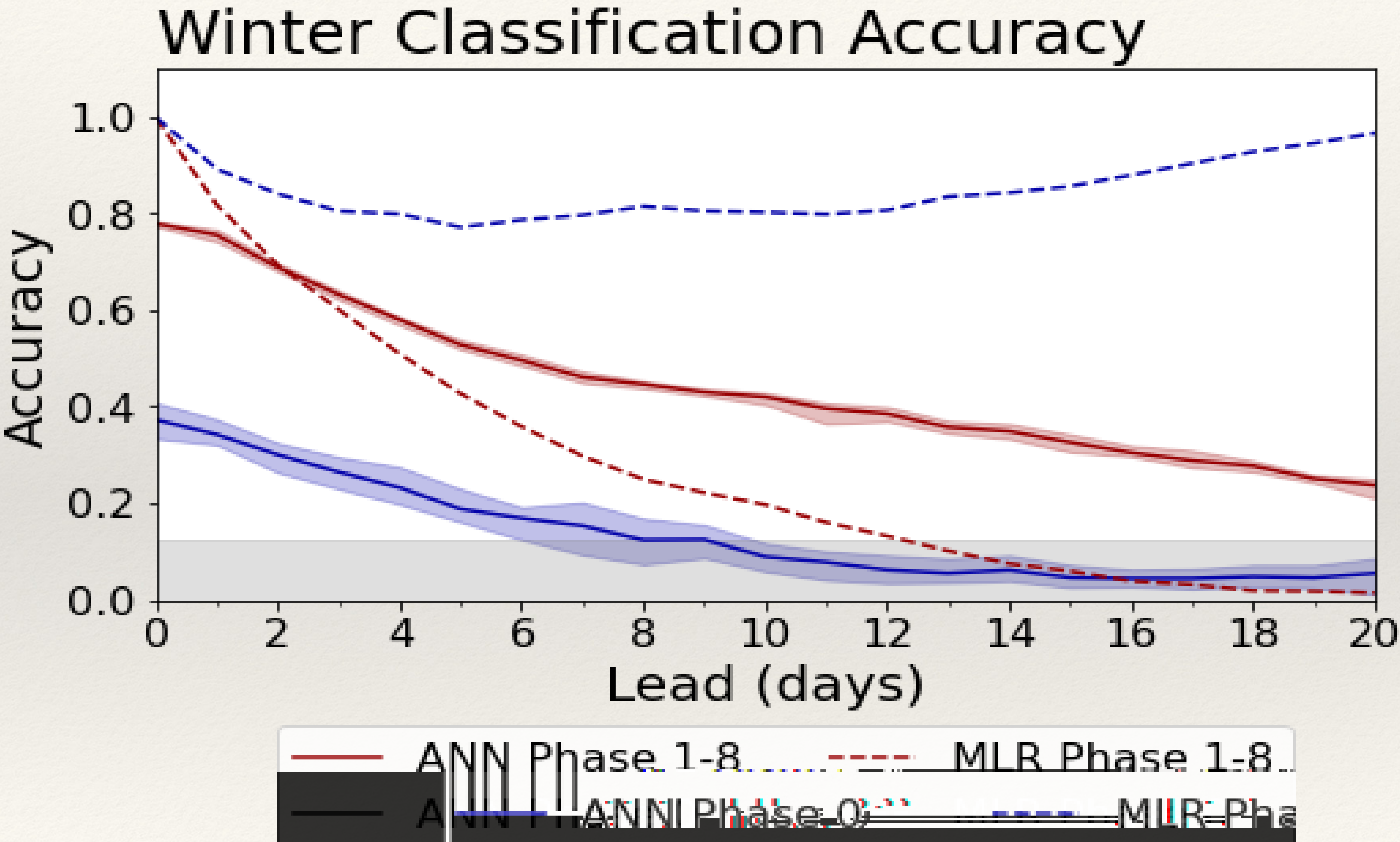


Regression Model

- A model which outputs *numerical values*

Martin et al. (2021; in prep)


(a) Winter Reg. ANN Skill
Initial MJO Phase
Lead (days)
(b) Winter Reg. ANN Amplitude Error
Initial MJO Phase
Lead (days)
(c) Winter Reg. ANN Skill
Weak
Strong
BCC
Lead (days)
(d) Dec.-Feb. ANN Skill
DJF QBO Easterly
BCC
Lead (days)


Winter Classification Accuracy
Accuracy
Lead (days)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
14
16
18
20
ANN Phase 1-8
MLR Phase 1-8
ANN Phase 0

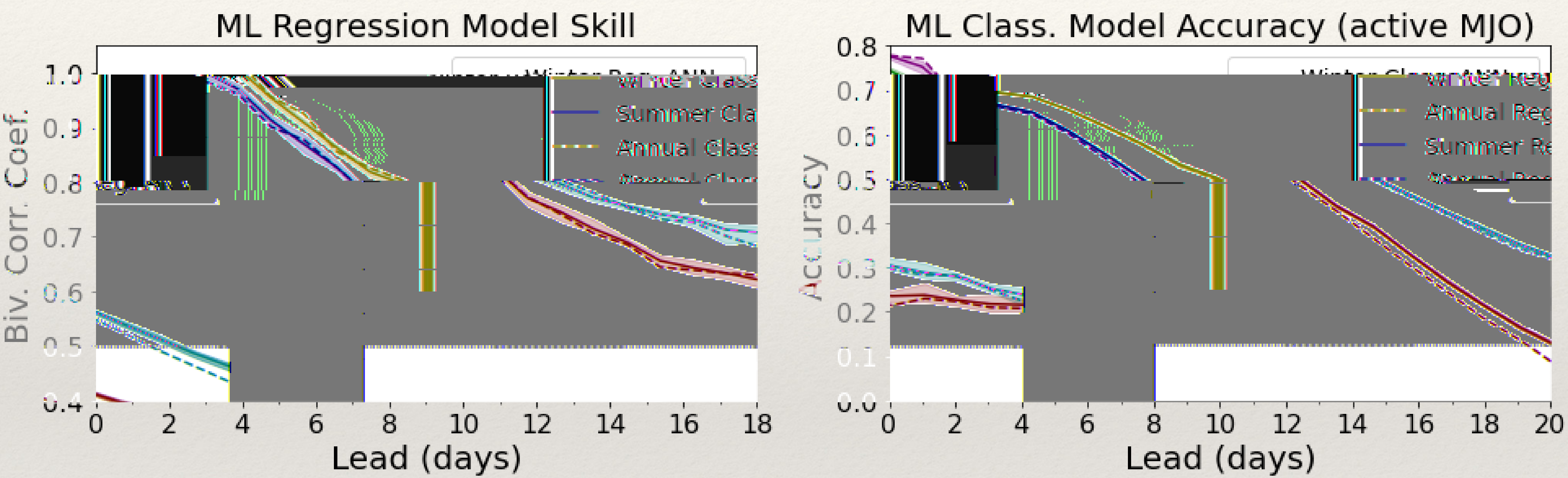

ML Regression Model Skill
Biv. Corr. Coef.
Lead (days)
ML Class. Model Accuracy (active MJO)
Accuracy
Lead (days)

Figure S3

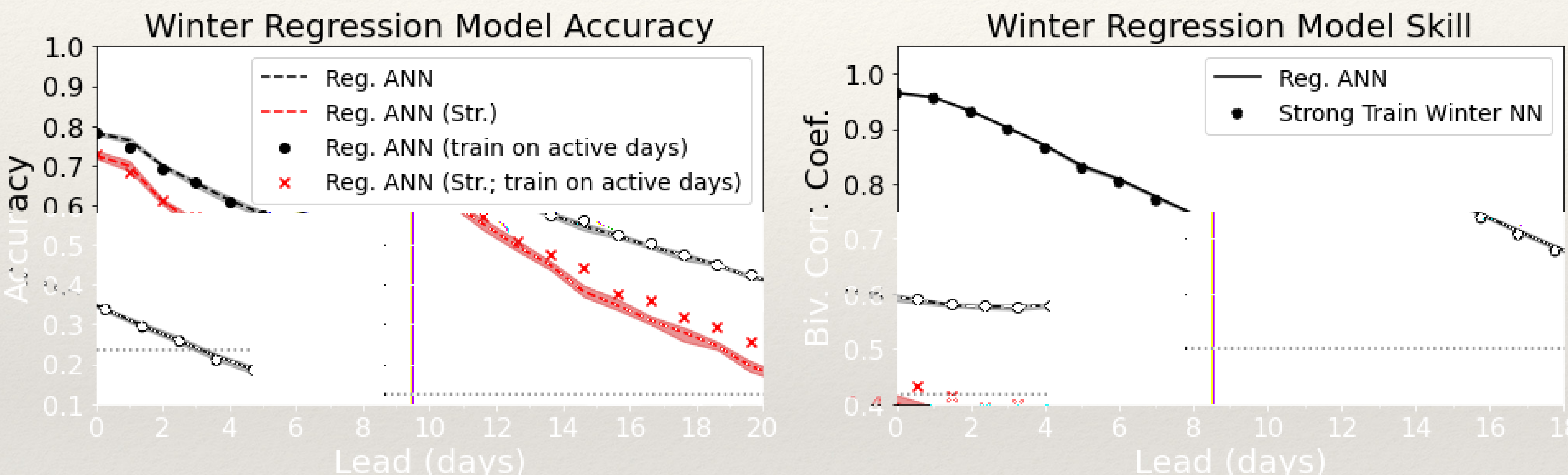
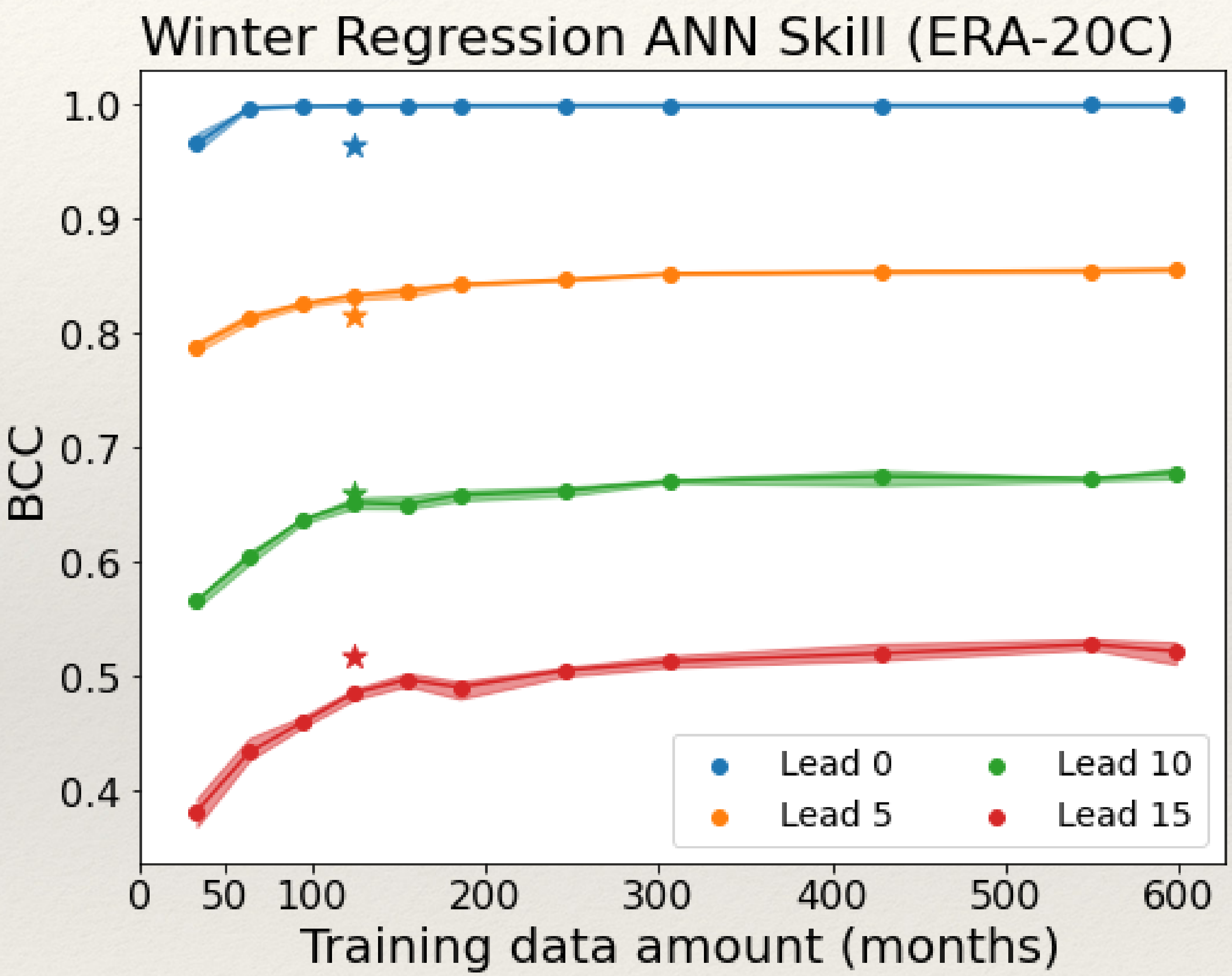


Figure S6: Strong training


Winter Regression ANN Skill (ERA-20C)
BCC
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
50
100
200
300
400
500
600
Training data amount (months)
Lead 0
Lead 5
Lead 10
Lead 15

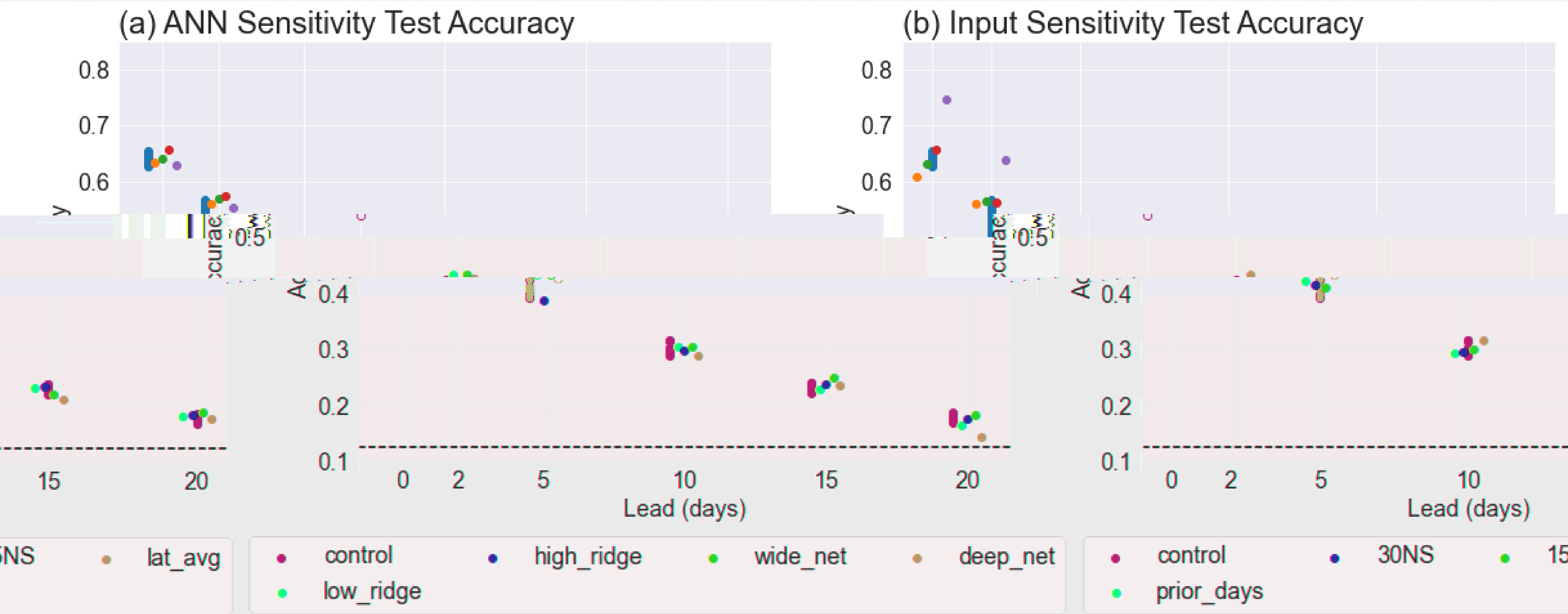


Figure S7: sensitivity tests

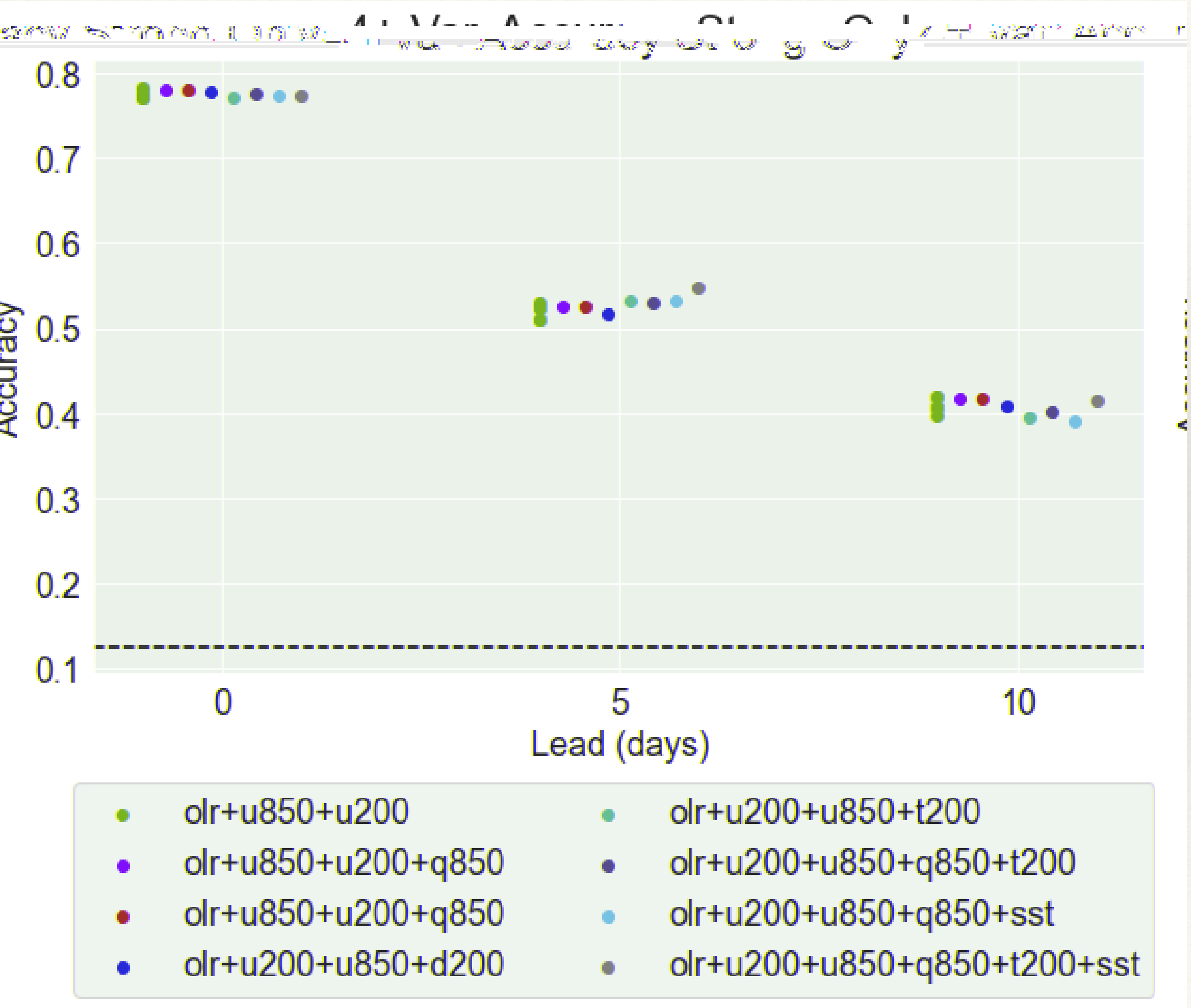


Fig. S11: Additional variable tests of combinations of 4, 5, or 6 inputs

1. “Persistence” model which simple persists the initial condition

$$RMM1(t_0 + \tau) = RMM1(t_0)$$
$$RMM2(t_0 + \tau) = RMM2(t_0)$$

2. **Vector autoregressive (VAR) scheme** (Maharaj & Wheeler 2005; Marshall et al. 2016)

Statistical bivariate forecast which captures 1-day typical change in RMM and steps forward (essentially akin to our prior "persistence" model)

$$\begin{bmatrix} RMM1(t) \\ RMM2(t) \end{bmatrix} = \boldsymbol{L} \begin{bmatrix} RMM1(t-1) \\ RMM2(t-1) \end{bmatrix}$$

MLR used to calculate **L** in each season

Marshall et al. 2016:

$$\mathrm{RMM1}_t = 0.9616\ (\mathrm{RMM1}_{t-1}) - 0.1135\ (\mathrm{RMM2}_{t-1})$$

$$\mathrm{RMM2}_t = 0.1257\ (\mathrm{RMM1}_{t-1}) + 0.9875\ (\mathrm{RMM2}_{t-1})$$

M21 "**L**":

```
LR.coef_
```

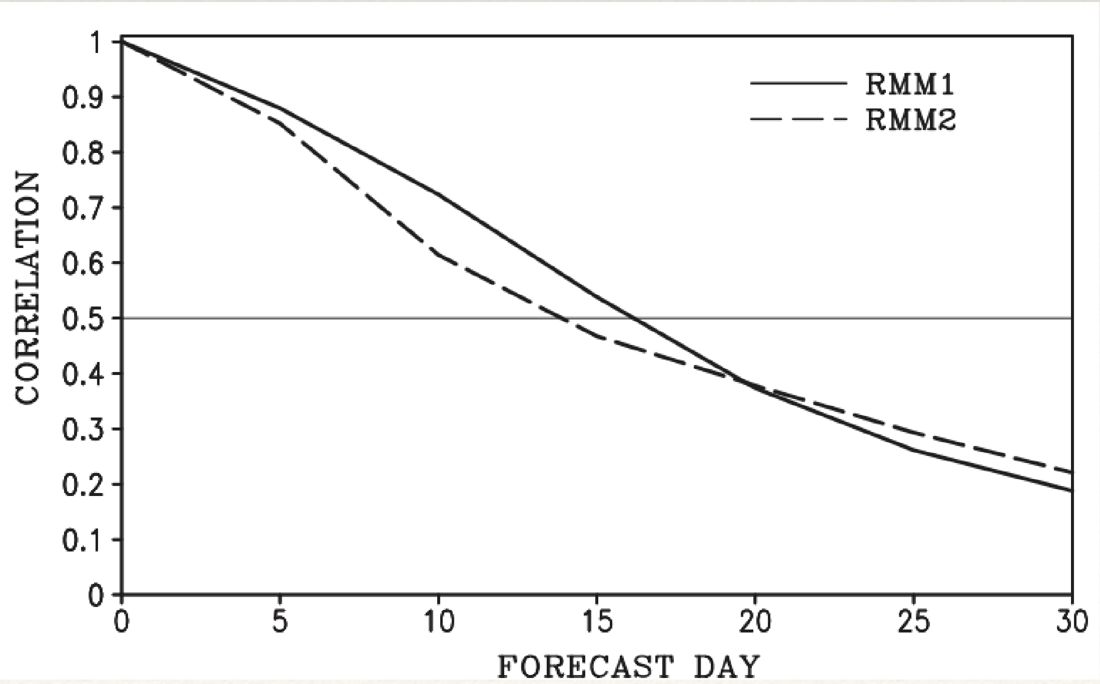
```
array([[ 0.96745844, -0.11490354],
       [ 0.12136748,  0.98466266]])
```

3. **Multiple linear regression (MLR) scheme** (Kim 2008; Jiang et al. 2008; Kang & Kim 2010; Seo et al. 2009, Wang et al. 2019)

Predicts RMM at lead τ given RMM at initial time and on prior days. Follow Kim & Kang (2010) who found j=2 (e.g. day 0 and day -1) is ok (Seo et al. 2009 used pentad data and retained more days, but change seemed relatively small).

$$\begin{bmatrix} RMM1(t_0+\tau) \\ RMM2(t_0+\tau) \end{bmatrix} = \boldsymbol{L}_\tau \quad \Sigma_{j=1} \begin{bmatrix} RMM1(t_0-j+1) \\ RMM2(t_0-j+1) \end{bmatrix}$$

Kang & Kim 2010



Present MLR model
(different train/validation period