Relevant large-scale predictors for S2S precipitation forecast using XAI

Machine Learning Workshop | ECMWF | 30/03/2022

Niclas Rieger (@niclasrieger) | CRM, Barcelona, Spain

CAFE
Climate Advanced Forecasting
of sub-seasonal Extremes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N°813844

In collaboration with
Laura **Ferranti** & Linus **Magnusson** (ECMWF)
Álvaro **Corral** (CRM)
Estrella **Olmedo** & Antonio **Turiel** (ICM-CSIC)

Relevant predictors of S2S precipitation

Objective

Identify large-scale patterns that provide opportunities for subseasonal precipitation forecast

1. What climate variables are the most important drivers?
2. Which regions play an important role?
3. When are these large-scale patterns active?

Approach

Interpretable dimensionality reduction

(regularized PCA)

+

RNN

+

eXplainable AI

(LRP)

1

Create set of potential predictors

Wide range of different **climate variables** and climate **indices**

2

Improve S2S precipitation forecast

Purely statistical model based on small **LSTM-RNNs**

3

Uncover relevant predictors

Make RNNs **explainable** using layer-wise relevance propagation (**LRP**)

Create set of potential predictors

1

Wide range of different **climate variables** and climate **indices**

Potential predictors | 1

- **Predictors**
  - **Climate variables**
    - *ERA5 Reanalysis*
    - *(1979-2021)*
    - *Daily*
    - **Atmosphere**: T2m, z500, RH
    - **Land**: Soil moisture
    - **Ocean**: SST, SSS
    - **Cryosphere**: SIC, Snow
  - **Climate indices**
    - *NOAA PSL*
    - *(1979-2021)*
    - *Monthly interpolated to daily*
    - **Pacific**: ONI, SOI, EMI, PDO, PNA
    - **Atlantic**: NAO, AMM, TNA, TSA
    - **Tropics**: MJO, QBO, Solar
    - **Poles**: AAO, AO

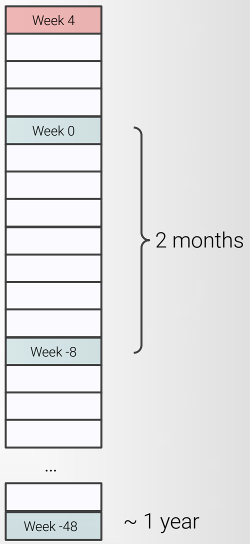
Multivariate PCA + Varimax rotation

40 PCs

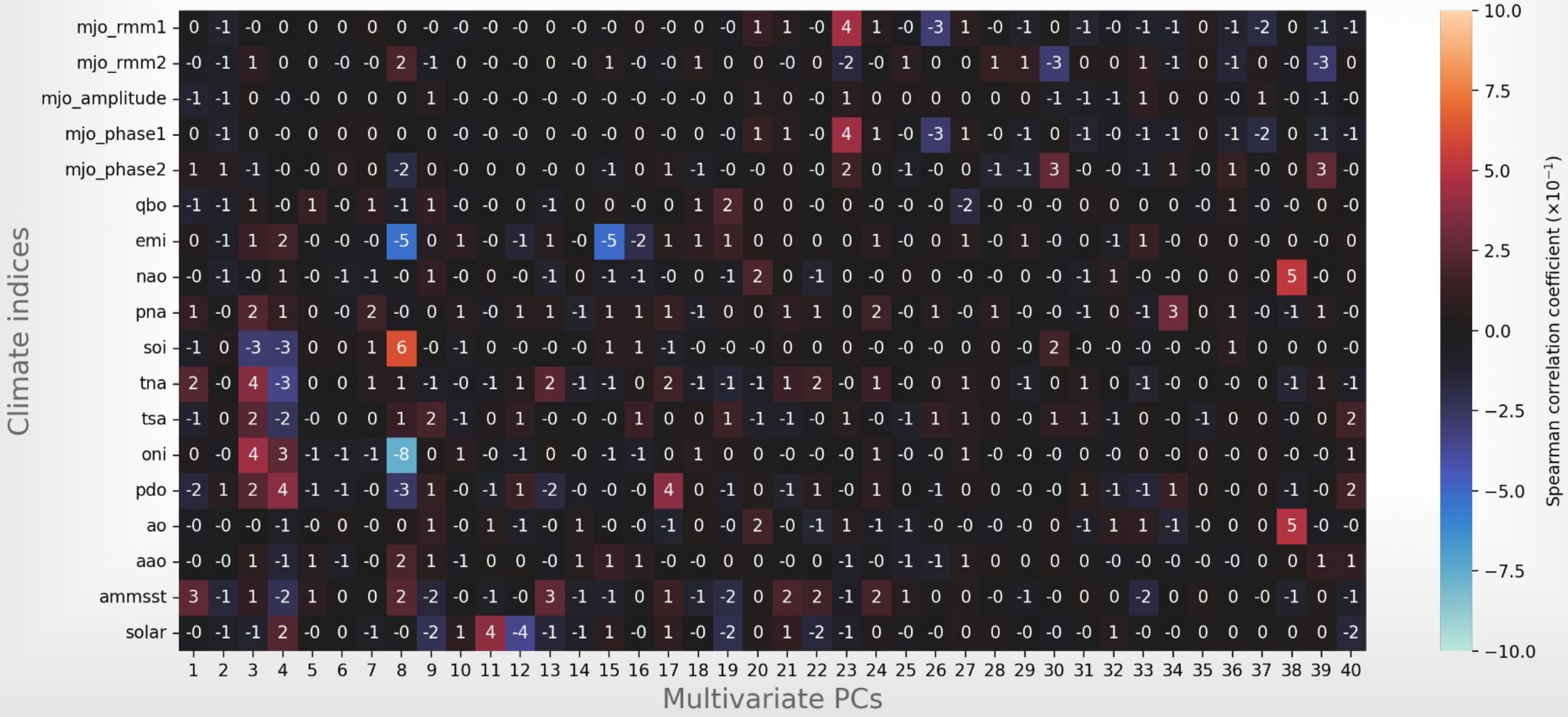
18 climate indices

Input

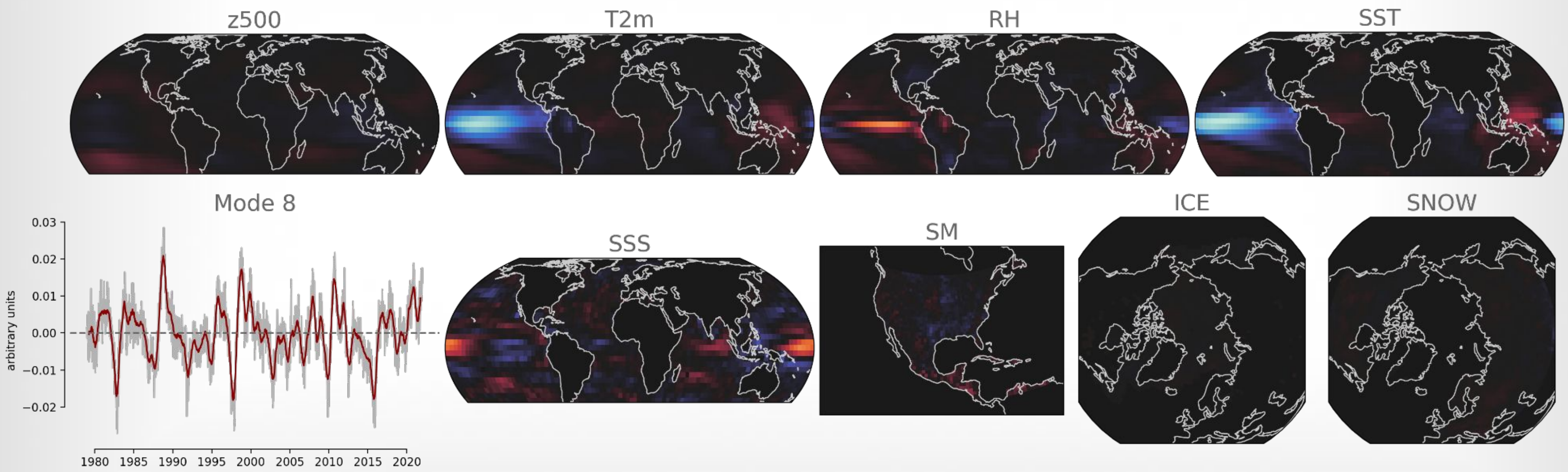
1 sequence for each of 58 predictors



Potential predictors | 1



Multivariate PCA + Varimax rotation

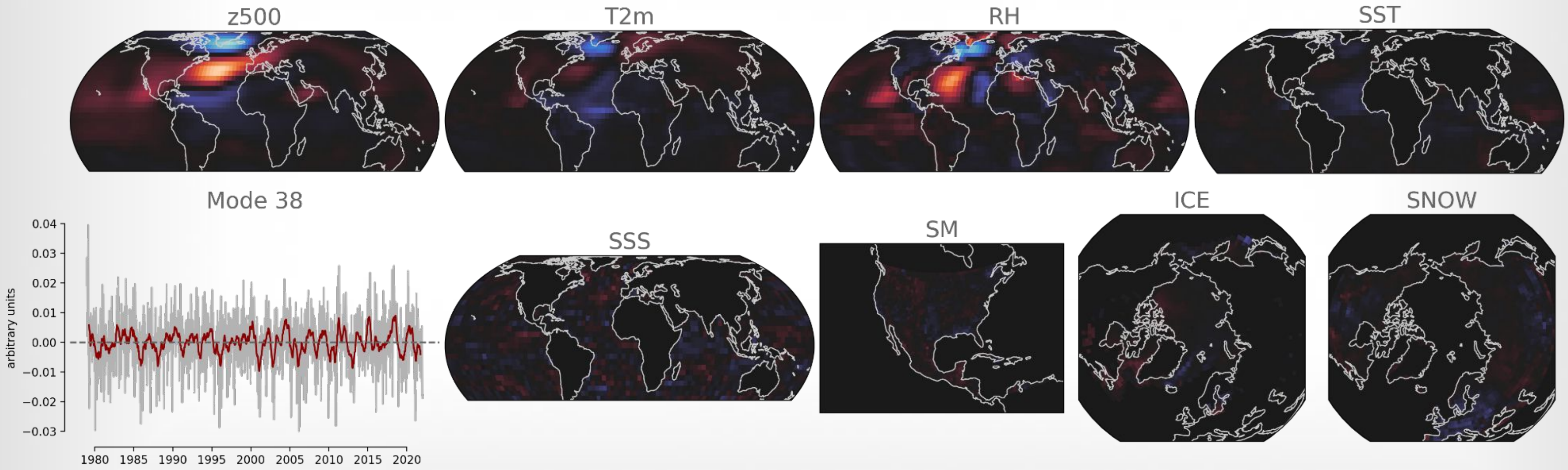


Example

ENSO (EP)

Potential predictors | 1

Multivariate PCA + Varimax rotation



Example

NAO - like

Create set of potential predictors

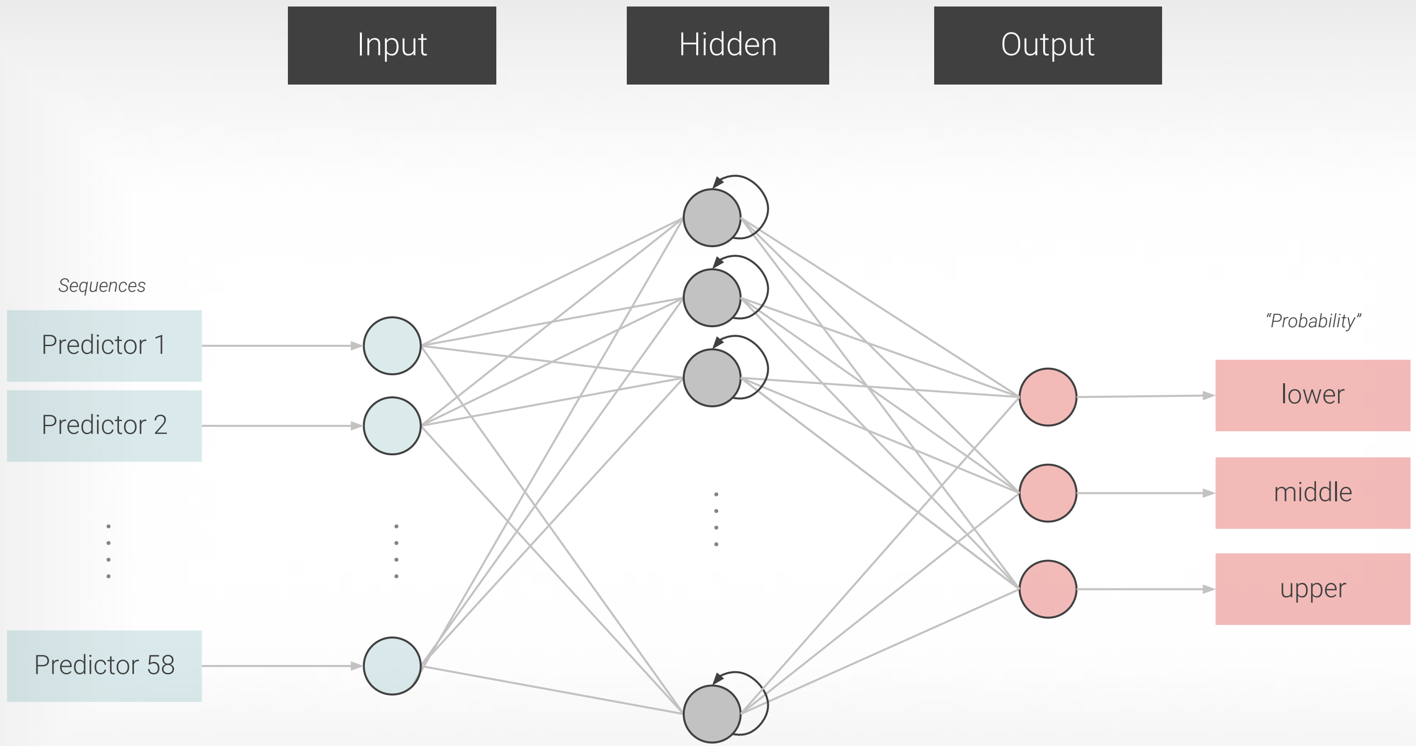
1

Wide range of different **climate variables** and climate **indices**

Improve S2S precipitation forecast

2

Purely statistical model based on small **LSTM-RNNs**



Architecture details

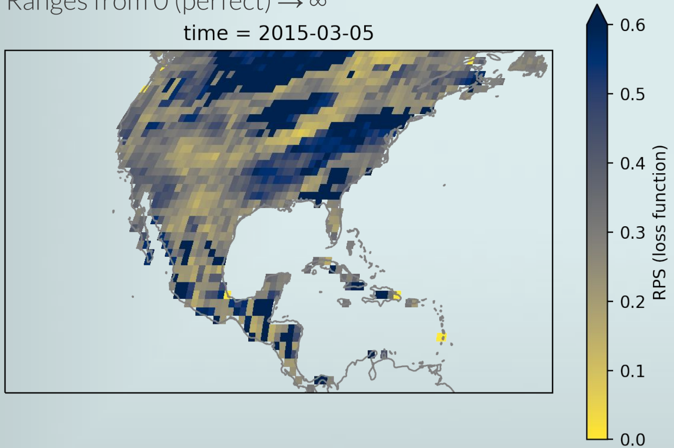
- 128 neurons
- LSTM cells
- Recurrent dropout: 0.5
- Activation: Sigmoid
- Learning rate: 1e-4
- Batch size: 16
- Training: 1 epoch

Ranked Probability Score (RPS)

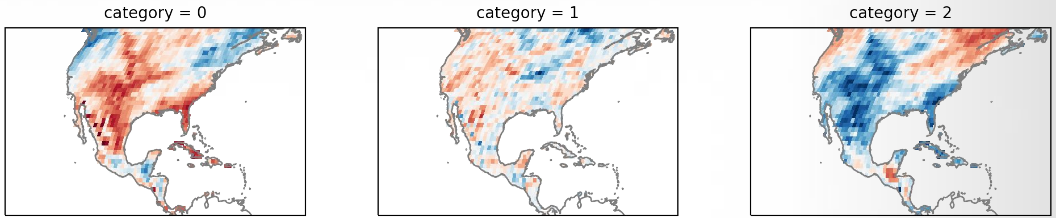
$$\text{RPS} = \sum_m^M \left[\left(\sum_i^m \hat{y}_i\right) - \left(\sum_i^m y_i\right)\right]^2$$

M : number of categories
ŷ : forecast
y : observation

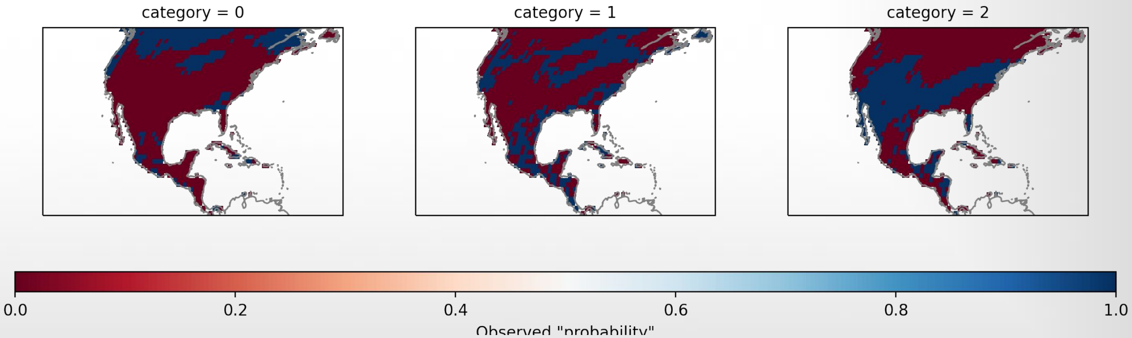
Ranges from 0 (perfect) $\longrightarrow \infty$



Forecast ŷ



Observation y



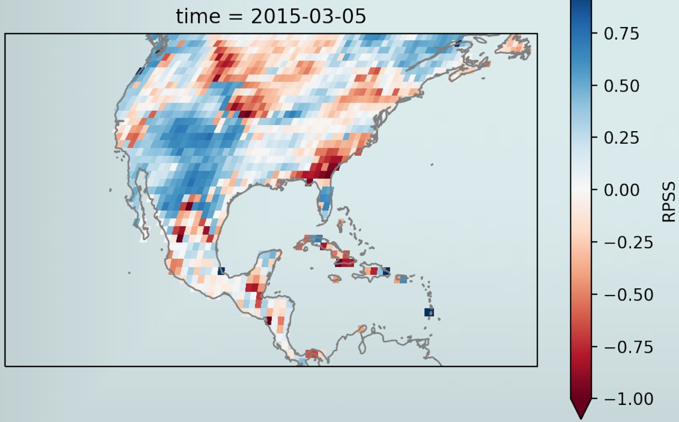
Evaluation metric | 2

Ranked Probability Skill Score (RPSS)

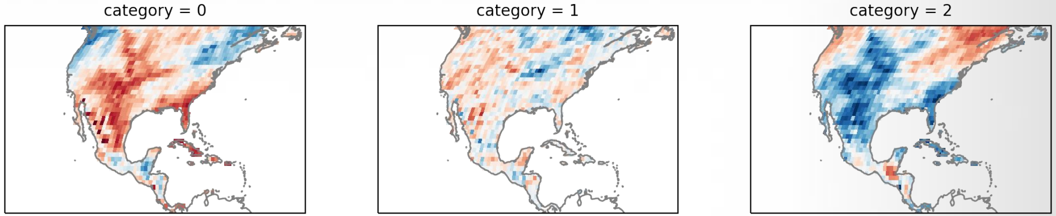
$$\text{RPSS} = 1 - \frac{\text{RPS}}{\text{RPS}_{ref}}$$

Compare the RPS of the forecast against another (reference) forecast.

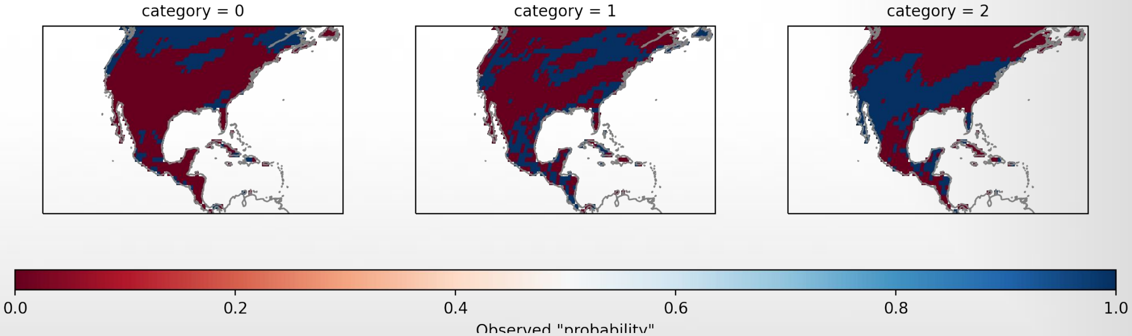
Climatology** is often the reference.
Ranges from -∞ → 0 (climatology) → +1 (perfect)



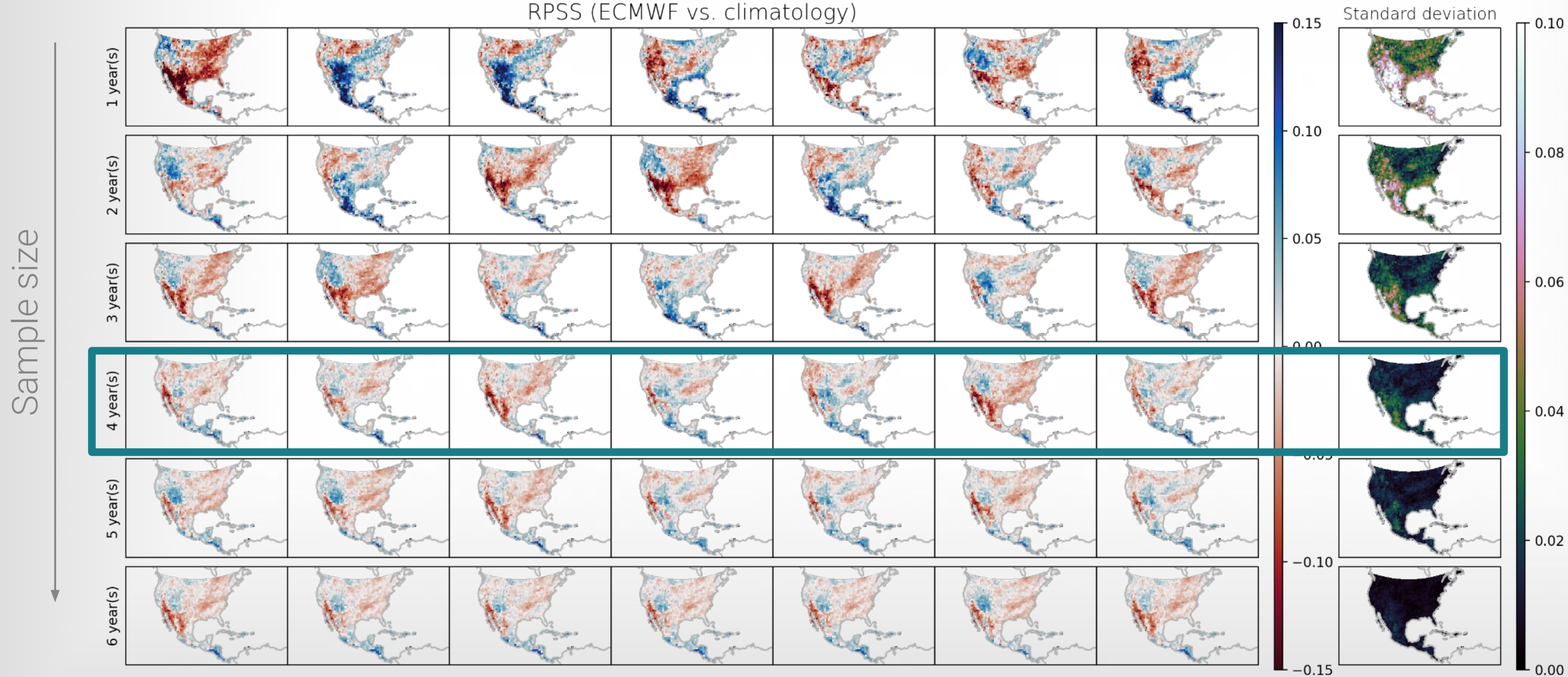
Forecast ŷ



Observation y



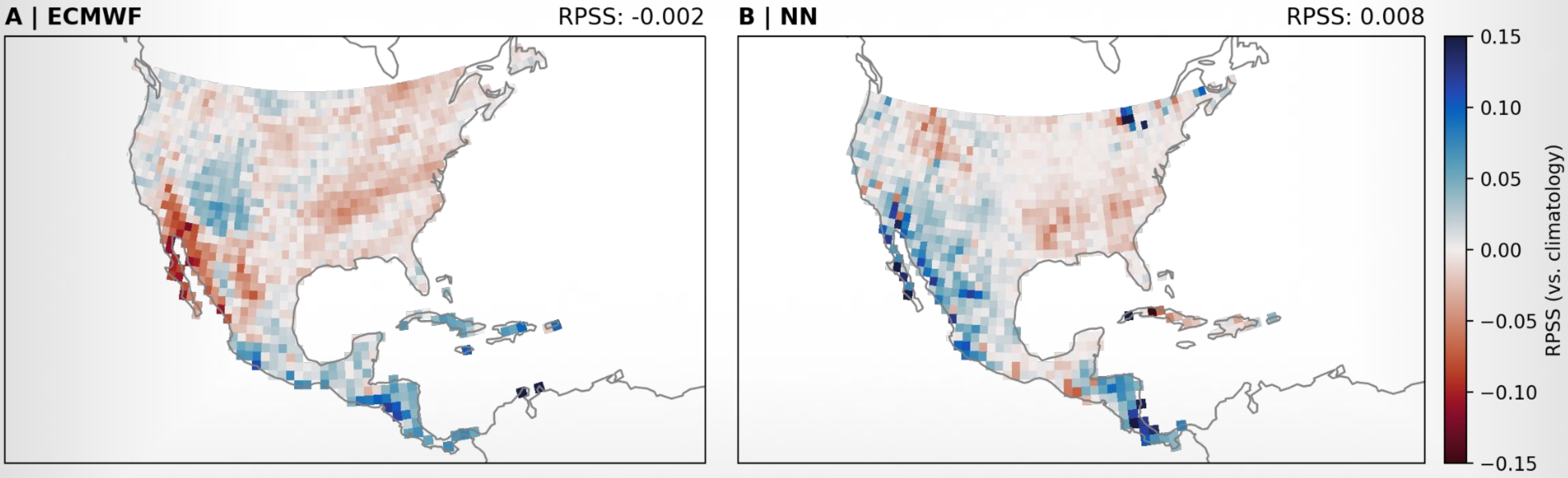
Effect of sample size on RPSS



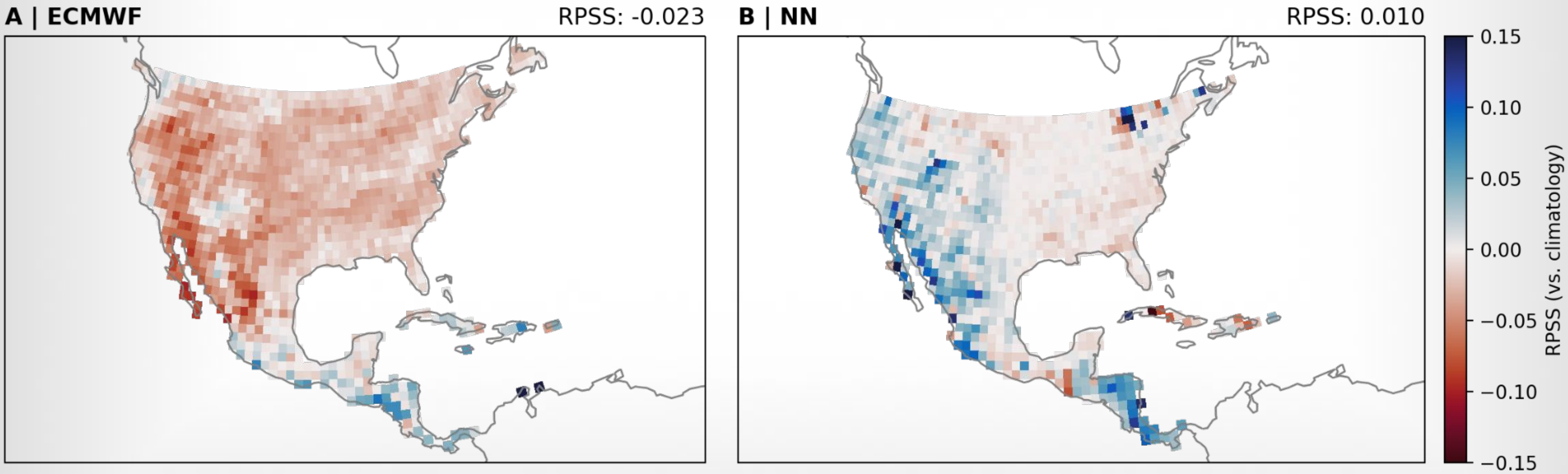
High year-to-year variability

4 years seem to be a good compromise

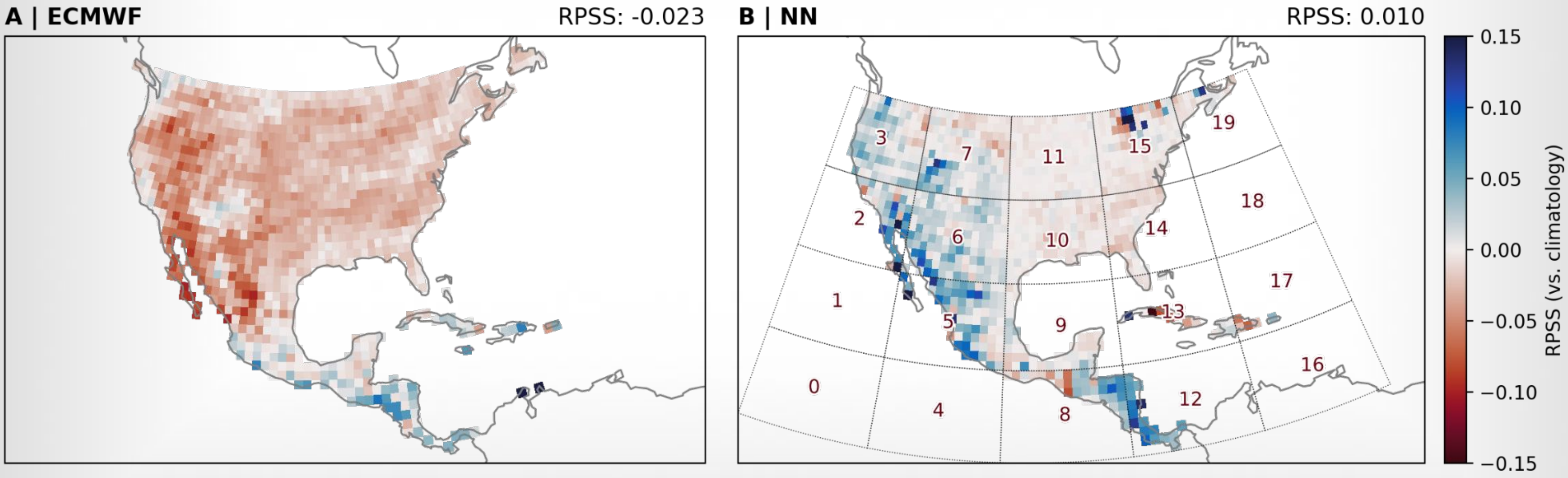
Forecast evaluation 2018-2021 (Week 3)



Forecast evaluation 2018-2021 (Week 4)



Forecast evaluation 2018-2021 (Week 4)



1

Create set of potential predictors

Wide range of different **climate variables** and climate **indices**

2

Improve S2S precipitation forecast

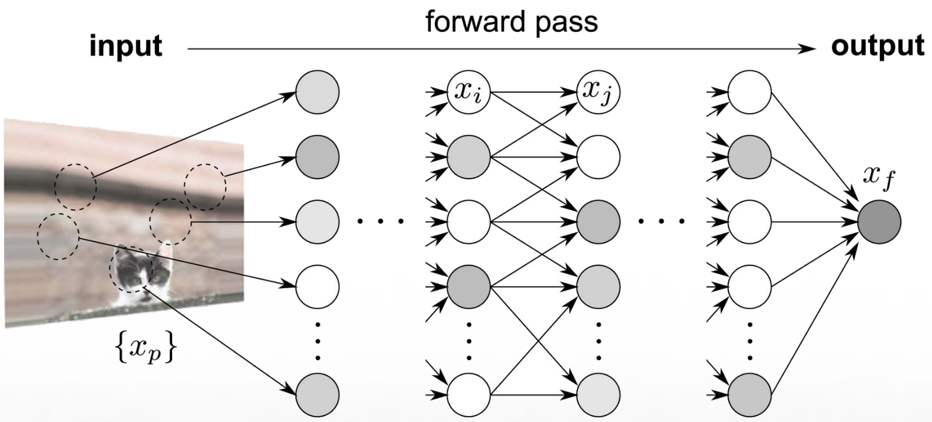
Purely statistical model based on small **LSTM-RNNs**

3

Uncover relevant predictors

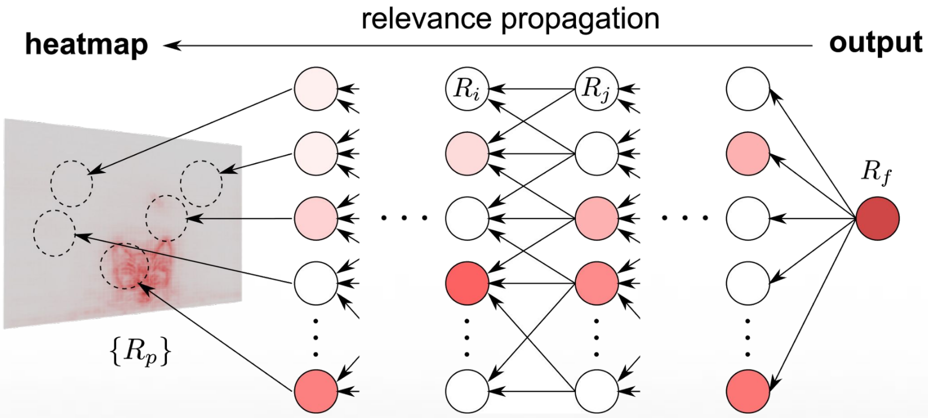
Make RNNs **explainable** using layer-wise relevance propagation (**LRP**)

Layer-wise relevance propagation



Reference: Bach et al (2015), Montavon et al. (2017), Arras et al. (2017)

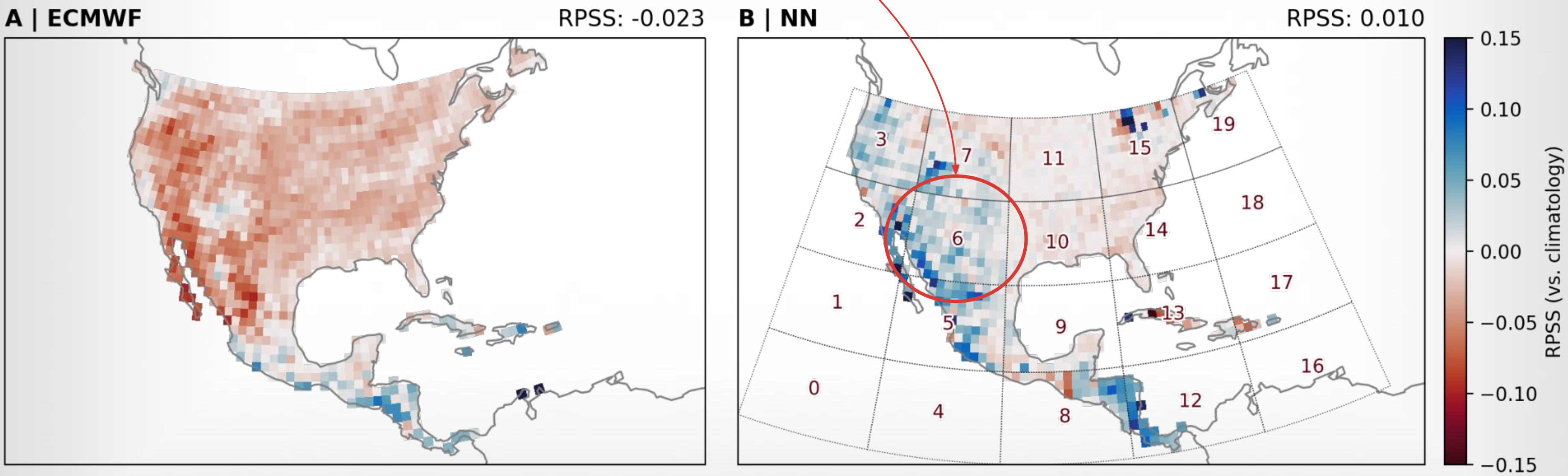
Layer-wise relevance propagation



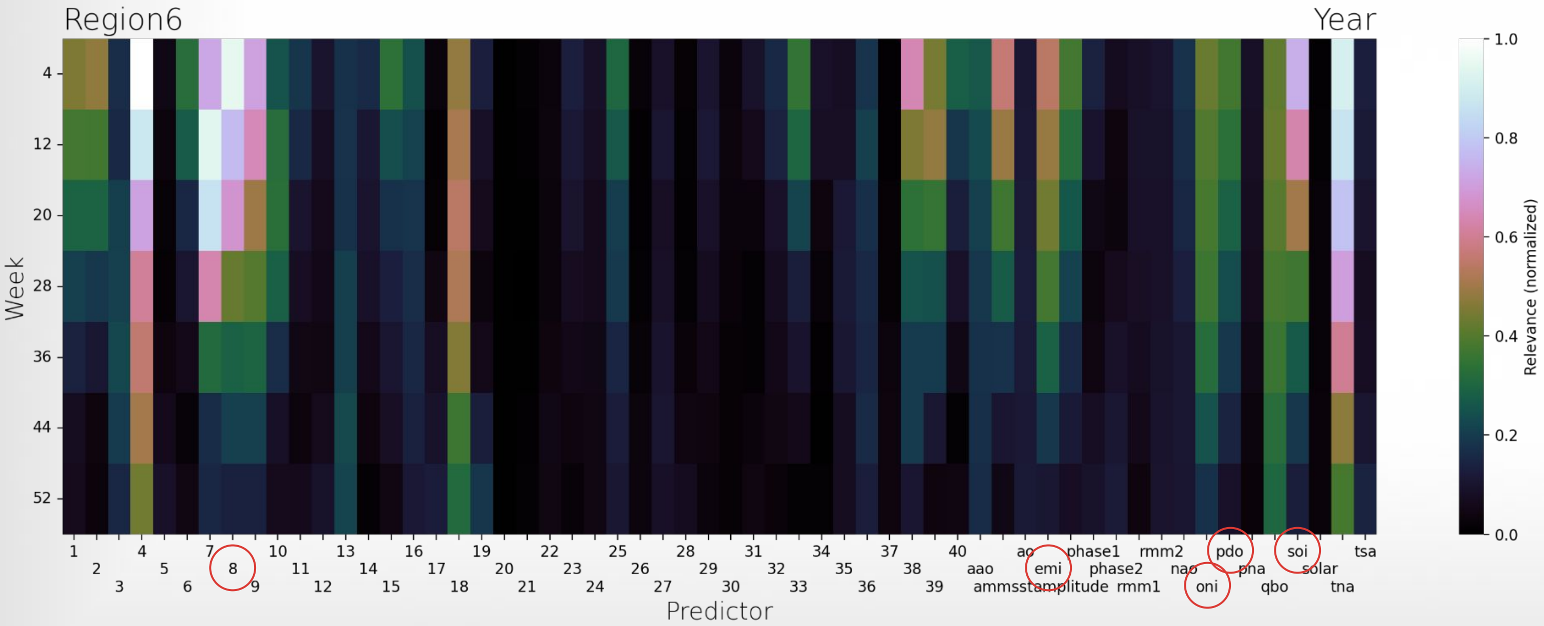
Reference: Bach et al (2015), Montavon et al. (2017), Arras et al. (2017)

Explainability | 3

What makes the forecast skillful in **region 6**?

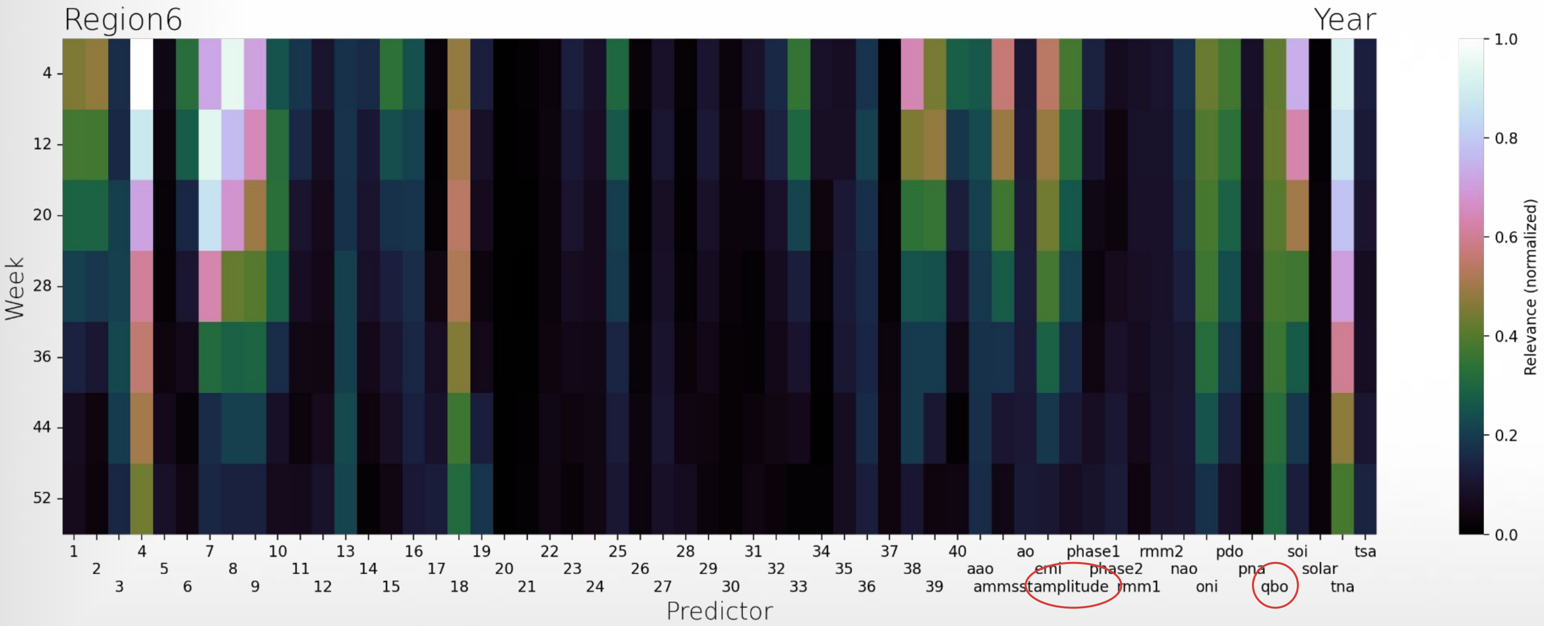


What makes the forecast skillful in **region 6**?



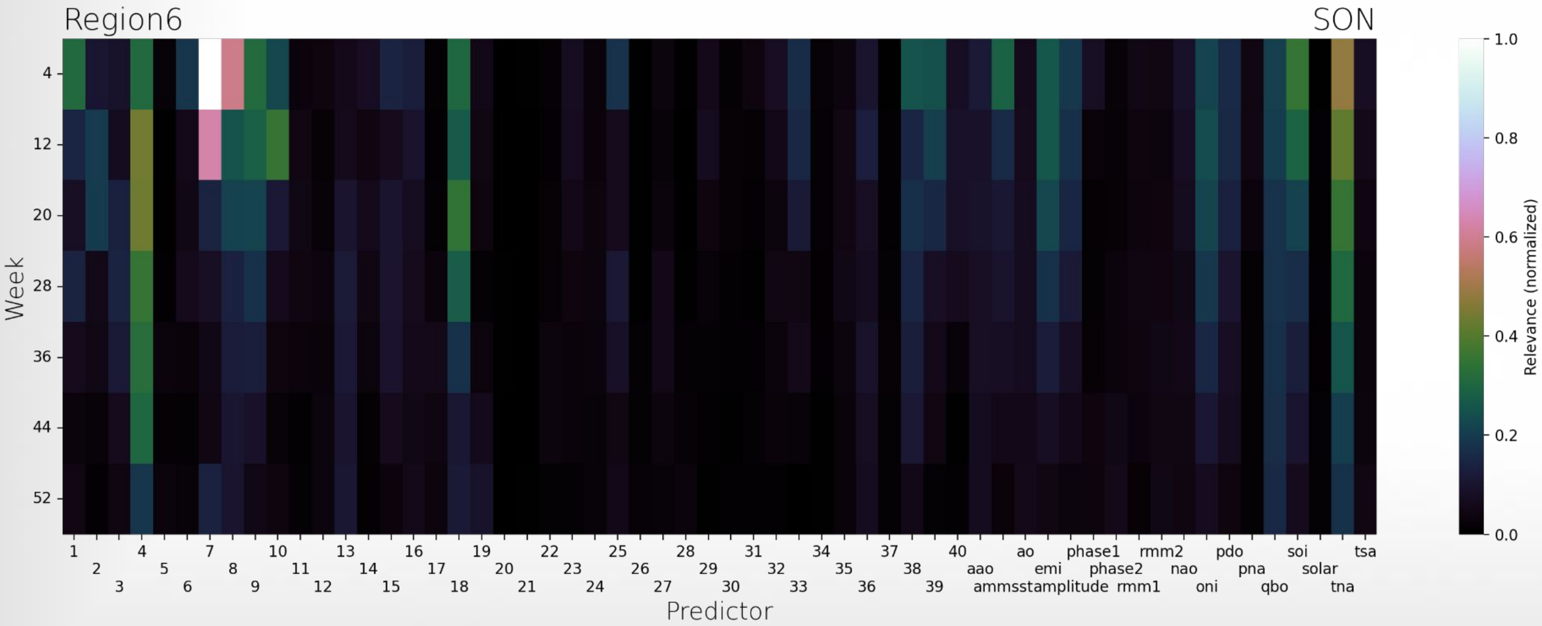
ENSO (CP, EP) + PDO are well known to influence North American climate

What makes the forecast skillful in **region 6**?

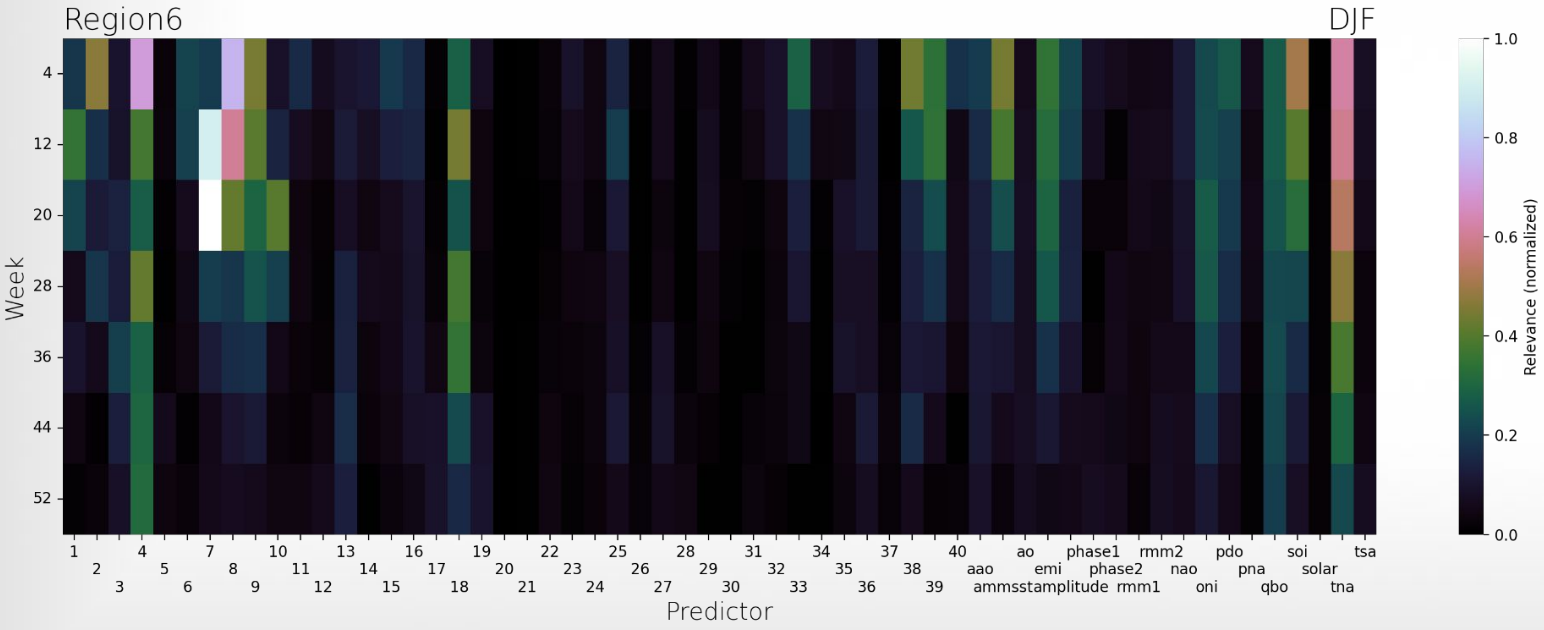


Influence of MJO & QBO has also been established recently

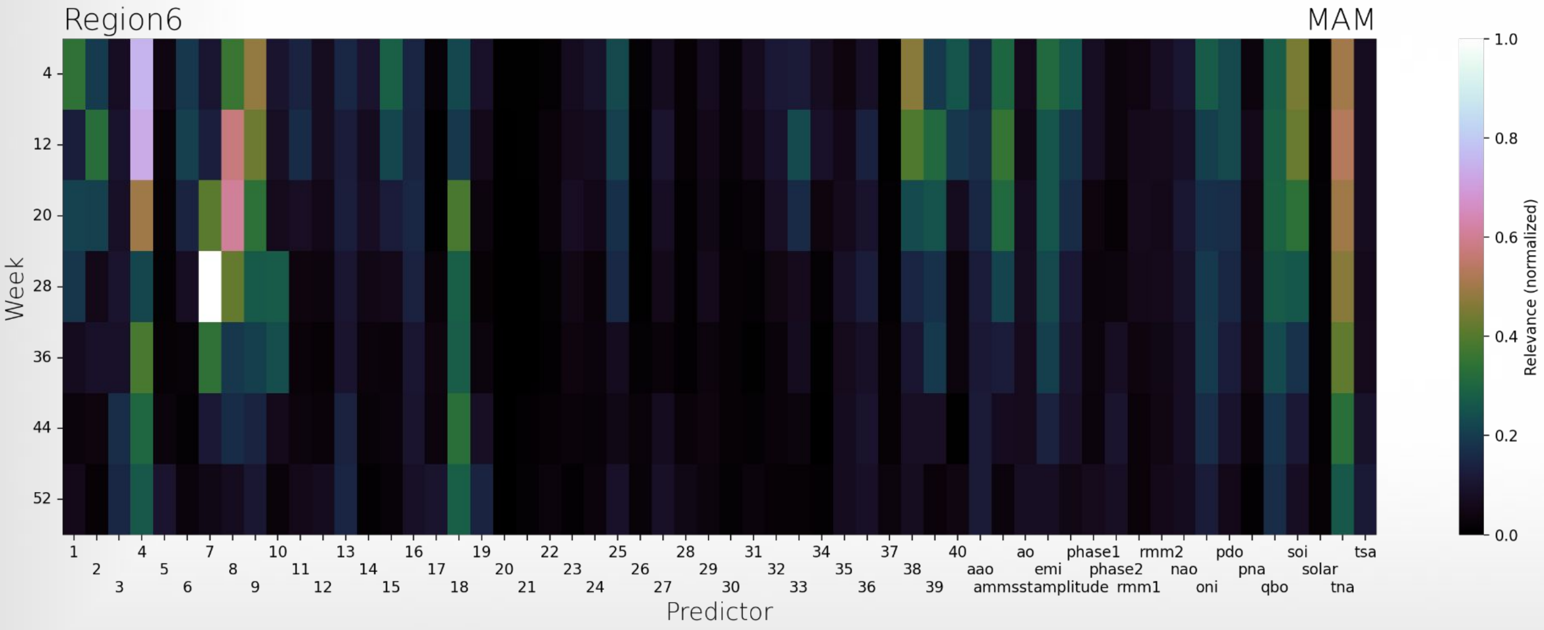
What makes the forecast skillful in **region 6**?



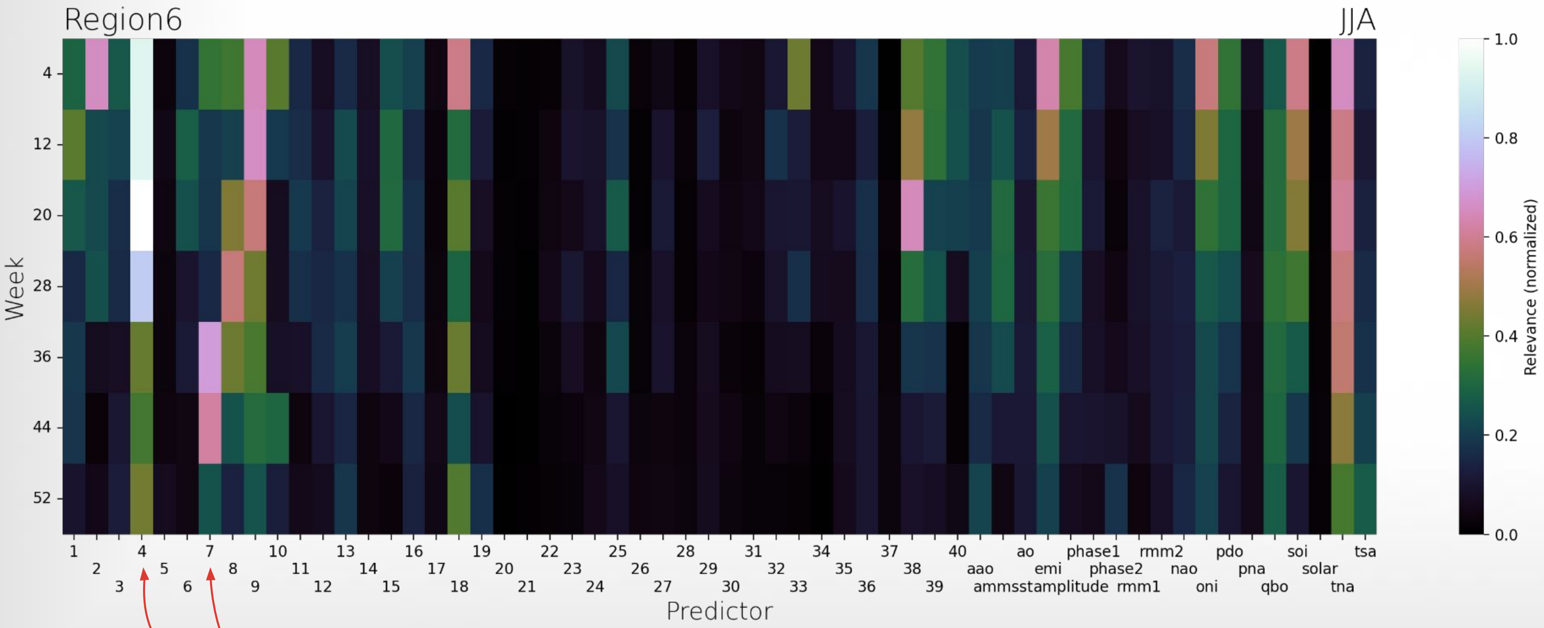
What makes the forecast skillful in **region 6**?



What makes the forecast skillful in **region 6**?



What makes the forecast skillful in **region 6**?

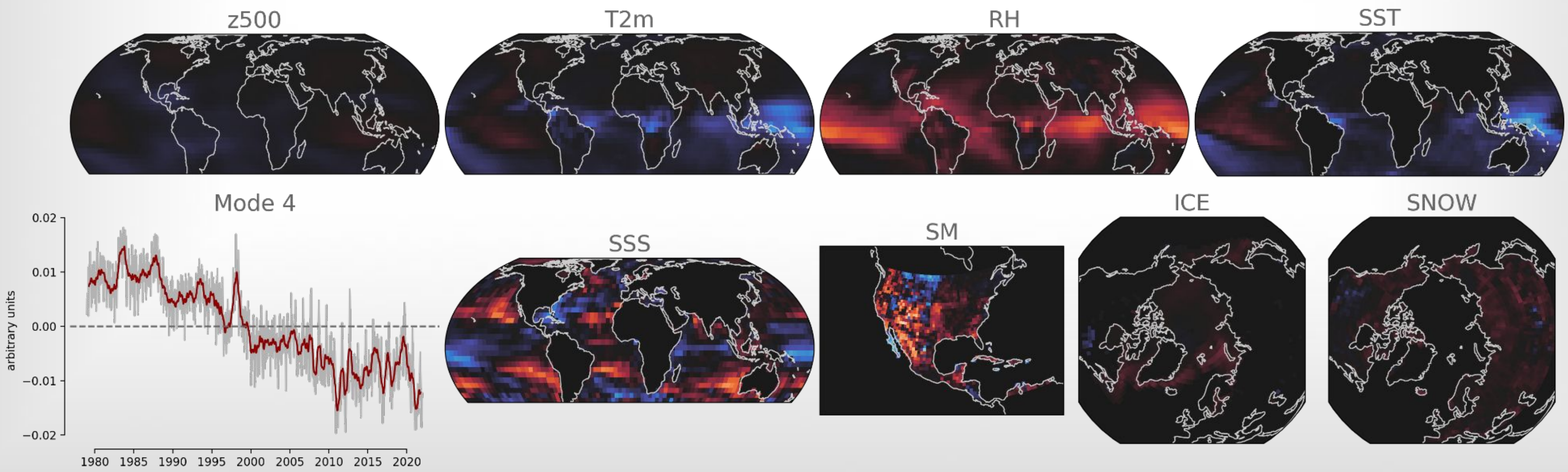


Let’s take a look at these two predictors

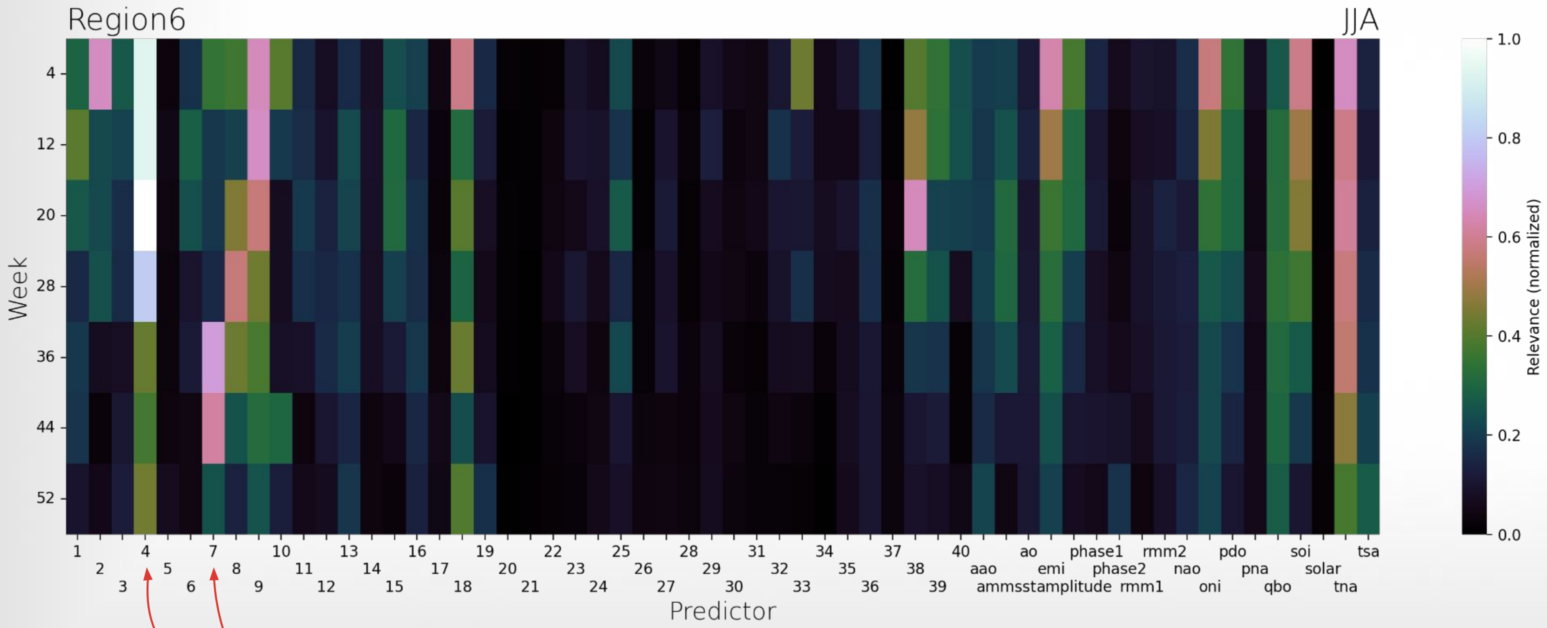
What makes the forecast skillful in **region 6**?

Mode 4

- important for JJA forecasts
- describes a reduction of RH, SM and SSS off the west coast of North America
- associated to current drought conditions in western North America?



What makes the forecast skillful in **region 6**?

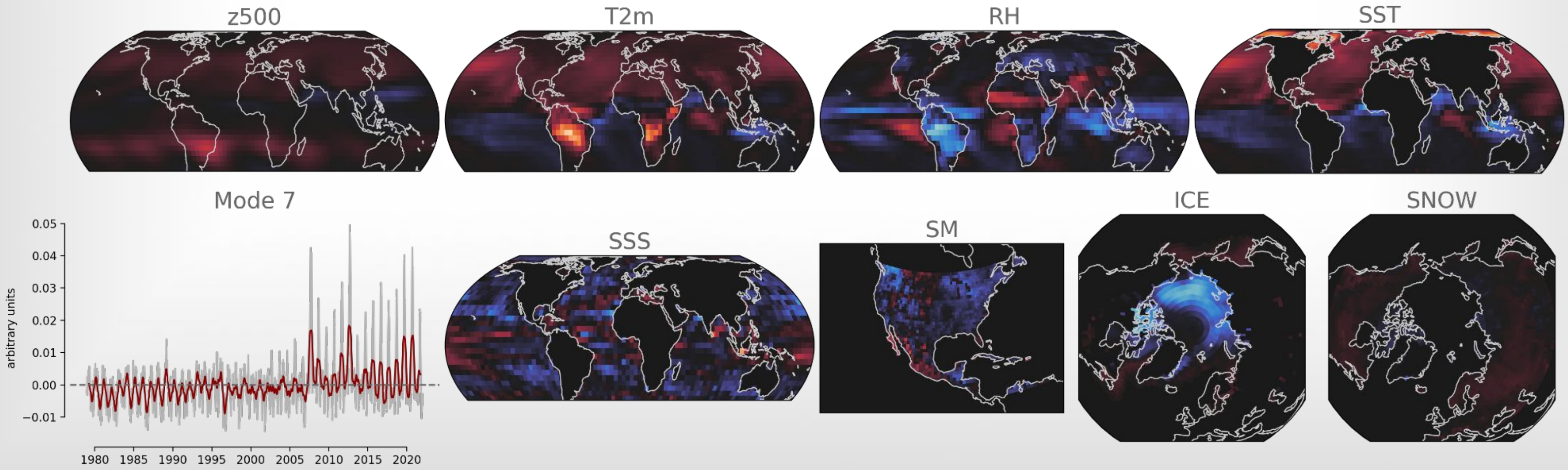


Let's take a look at these two predictors

What makes the forecast skillful in **region 6**?

Mode 7

- characterized by seasonality
- state of late summer provides predictive information
- related to sea ice extent, T2m, RH and SST



Final thoughts

1 | Selecting predictors

- Multivariate EOF analysis can be used to efficiently reduce dimensionality of multiple climate variables
- Interpretability ⇒ use regularization technique e.g. Varimax rotation

2 | Predicting precipitation

- Good skill for western North America indicates that some relevant predictors are not sufficiently well assimilated by ECMWF

3 | Explainability

- Importance of hydrological variables (RH, SM, SSS)
- Relevant predictors may have a time lag (e.g. sea ice concentration in the Arctic)
- With increased applications of ML methods in S2S forecasts, XAI techniques offer the potential to discover new predictors

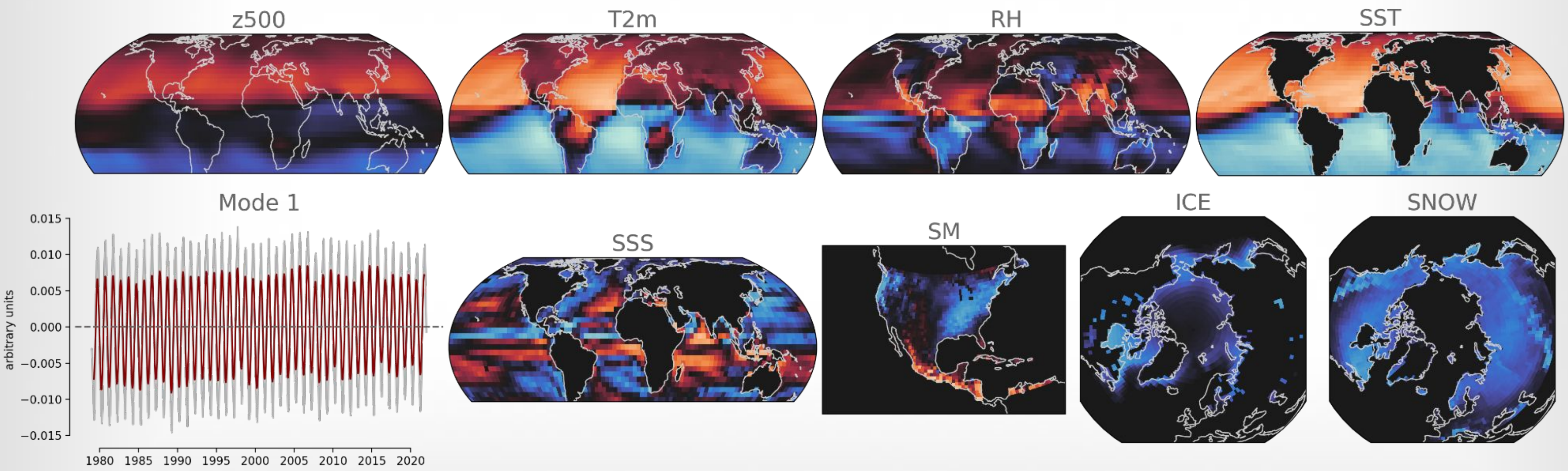
You may want to check out:

⇒ Python package xeofs
https://github.com/nicrie/xeofs

⇒ Short course on XAI by Ryan Lagerquist:
https://www.ai2es.org/products/education/

⇒ Follow activities of some XAI research groups e.g. Barnes (CSU), McGovern (UO)

Multivariate PCA + Varimax rotation



Example

Seasonal cycle