ENHANCING SEASONAL PREDICTIONS WITH MACHINE LEARNING: A GLOBAL PERSPECTIVE ON SST INFLUENCE IN EARLY WINTER

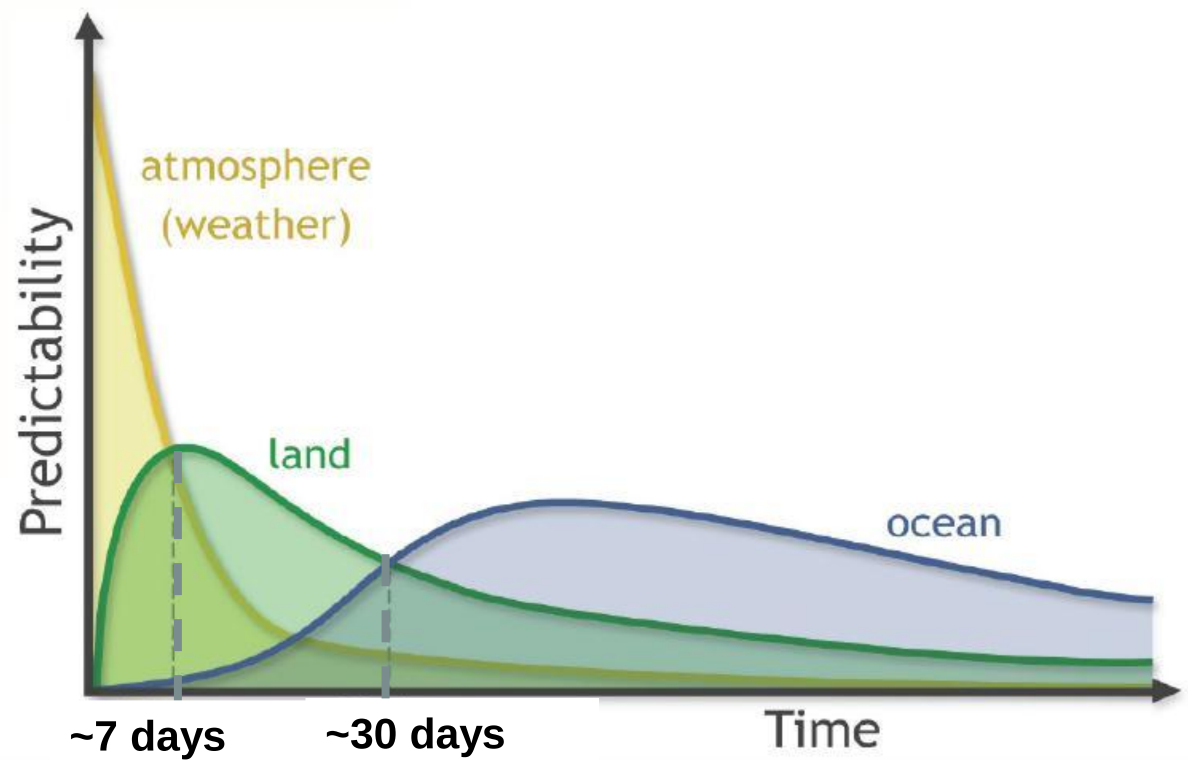
Víctor Galván Fraile, Marta Martín del Rey, Irene Polo, Belén Rodríguez Fonseca, Magdalena Alonso Balmaseda, Esteban Rodríguez Guisado, María Moreno García.

CONTENTS

1. Introduction: Seasonal Forecast

❖ Sources of predictability:

- ENSO (biggest single signal).
- Other tropical and extratropical SST.
- Local surface conditions.
- Sea-ice and Stratospheric processes.



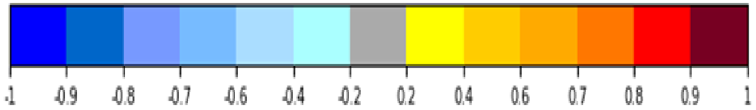
Source: Oskvig, K. et al (2020)

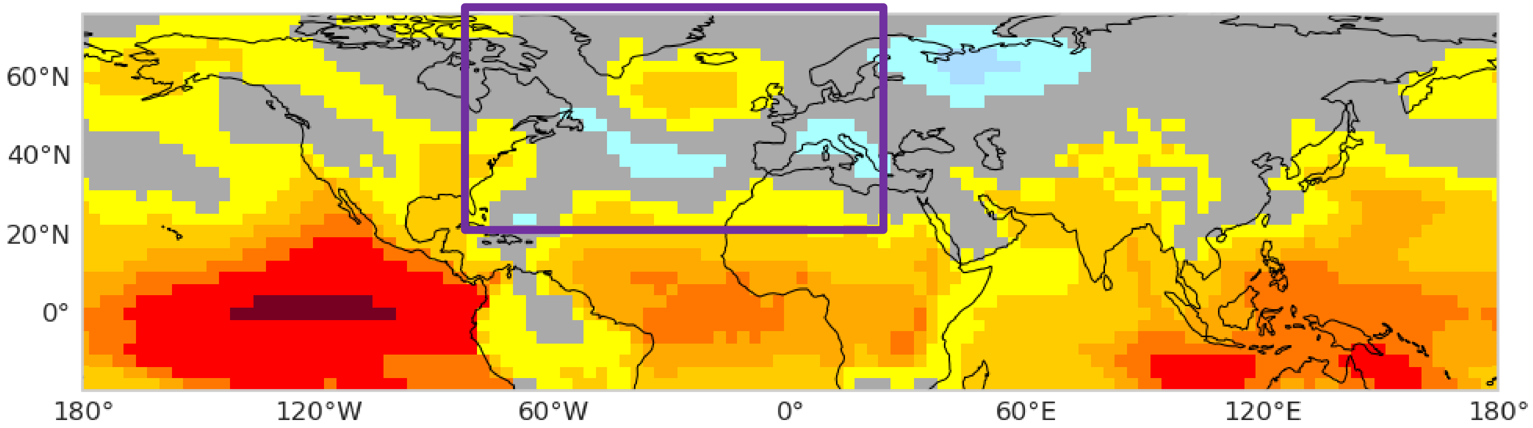
1. Introduction: Seasonal Forecast

- Dynamical models

  - Non-linear interactions from sources of predictability.
  - Ensemble forecast map uncertainty in initial state to uncertainty in outcome.
  - Model errors and biases.

SEAS5 (ECMWF) skill in SLP for Nov-Dec initialized in October.





Source: Johnson, S. et al (2019)

1. Introduction: Seasonal Forecast

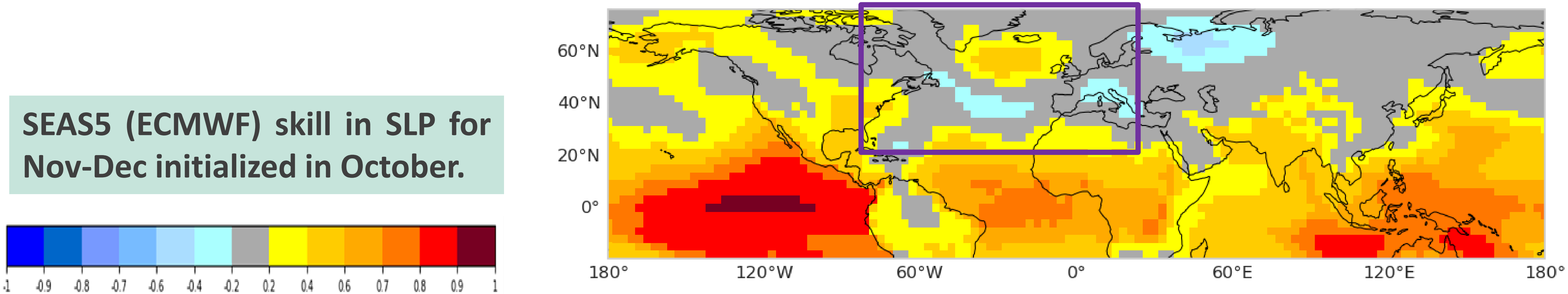
❖ Dynamical models

- Non-linear interactions from sources of predictability.
- Ensemble forecast map uncertainty in initial state to uncertainty in outcome.
- Model errors and biases.

❖ Statistical models

- Works with reality instead of error-prone numerical models.
- Aims to extract predictable signals.
- Limited number of past cases.
- A non-stationary climate is a challenge.

SEAS5 (ECMWF) skill in SLP for Nov-Dec initialized in October.

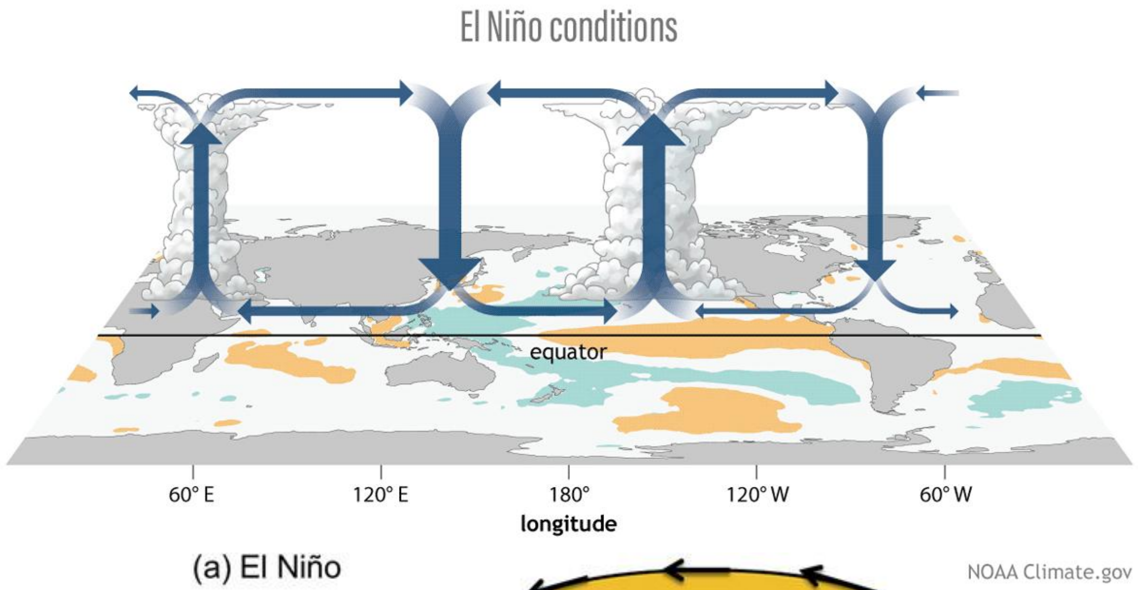


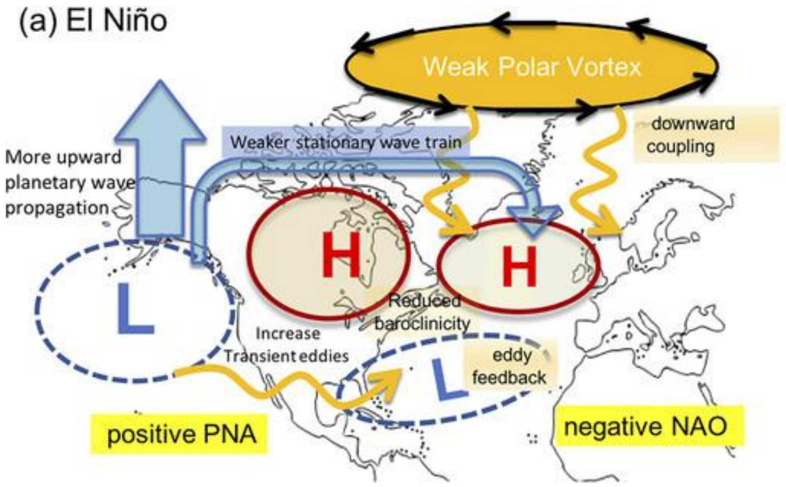
Source: Johnson, S. et al (2019)

1. Introduction: El Niño-Southern Oscillation

- What is ENSO?
  - Air-sea coupled mode (Bjerknes feedback).
  - Changes in Walker circulation.
  - Atmospheric Gill response.
  - Rossby waves.

- ENSO teleconnections:
  - Different pathways.
  - Non-linearities.
  - Non-stationarities.
  - Late Winter more robust.



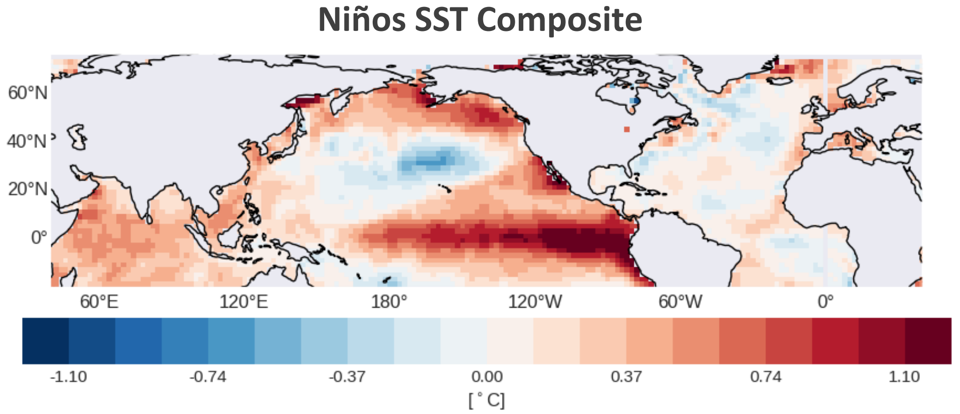


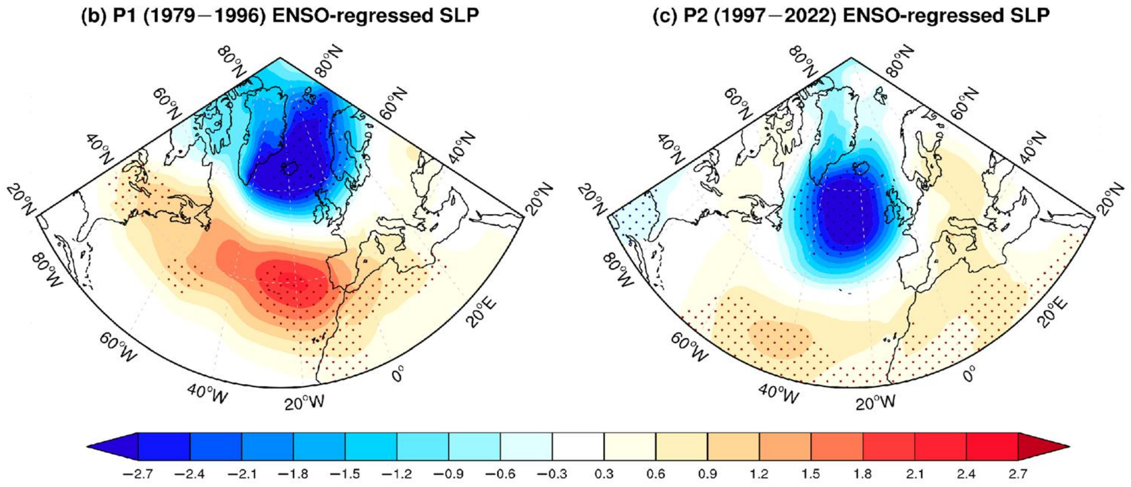
Source: Jiménez, B. et al (2018)

1. Motivation: ENSO Europe-North Atlantic Early Winter (ND) Teleconnection

- Analyze the ND ENSO teleconnection with machine learning.
- Analyze models skill and error.
- Look for predictable signals through model attributions.

Source: Hou et al. (2023)





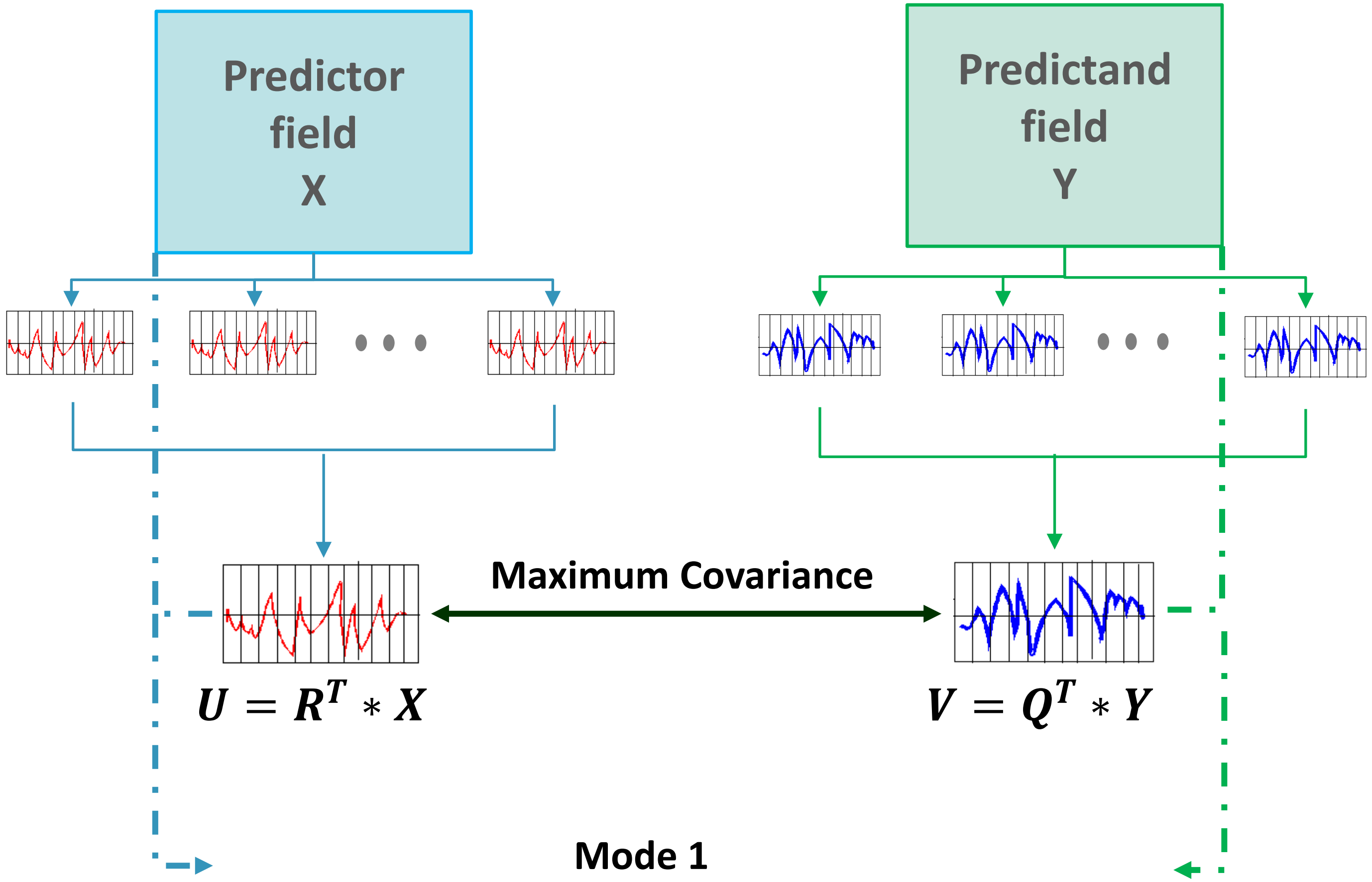
2. Methodology: Maximum Covariance Analysis

- Finds patterns such that time expansion coefficients (which are the projections onto the patterns) have **maximum covariance** and the patterns are **orthogonal** to each other.

- Finds a mapping function $\mathbf{\Phi}$ from the predictor field onto the predictand field, which is adjusted in the training phase.

$$\widehat{Y} = \mathbf{\Phi} * X$$

- **Objective**: **Maximize** the covariance during training.

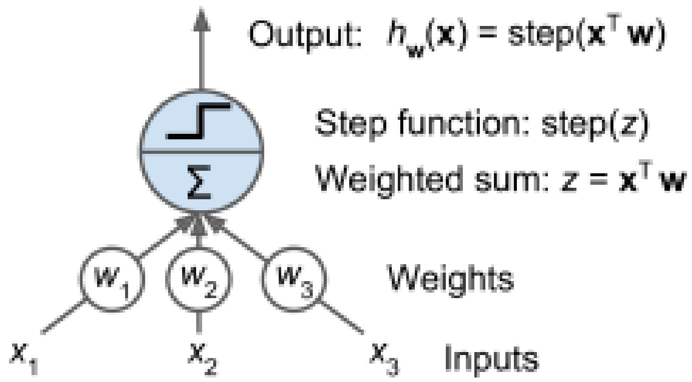


2. Methodology: Deep Learning (NN4CAST)

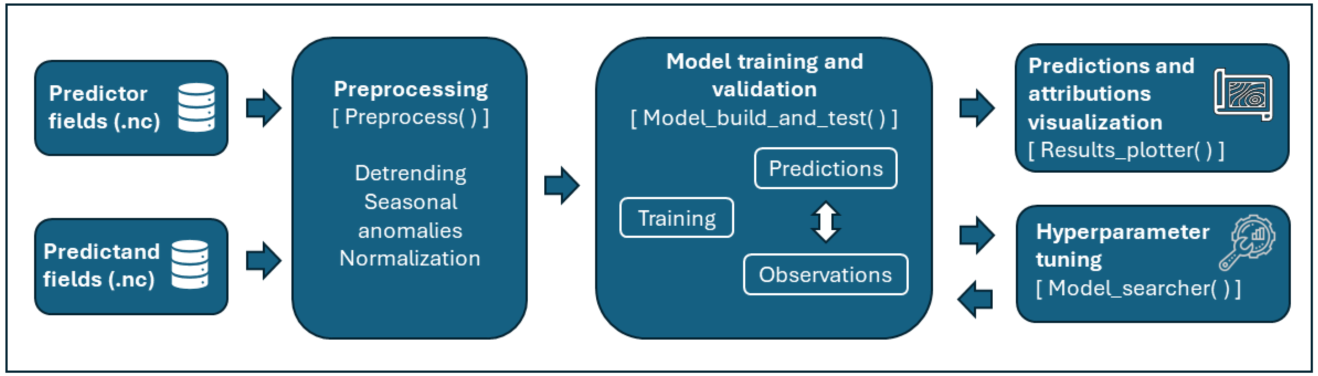
- Types of Machine Learning:
  - **Supervised learning.**
  - Unsupervised learning.
  - Reinforcement learning.

- **Artificial Neural Networks** (NNs):
  - Linear combination of inputs.
  - Non-linear activation function.
  - Input, output and hidden layers.
  - Universal function approximator.

- **Objective**: **Minimize** the mean squared error during training.



Source: Géron, A. (2022)



Source: Galván-Fraile, V. et al. (2024)

2. Methodology:

- **NN**: Finds a non-linear combination of the predictors (inputs) to make better predictions by minimizing a loss function.
  - **Objective**: **Minimize** the mean squared error during training.

- **MCA**: Finds patterns such that time expansion coefficients (which are the projections onto the patterns) have **maximum covariance** and the patterns are **orthogonal** to each other.
  - **Objective**: **Maximize** the covariance during training.

NN4CAST: An end-to-end deep learning application for seasonal climate forecasts

Víctor Galván Fraile[1], María Belén Rodríguez-Fonseca[1,2], Irene Polo[1], Marta Martín-Rey[1], and María N. Moreno-García[3]

[1]Group on Tropical Climate Variability and Atmospheric Teleconnections (TROPA), University Complutense of Madrid (UCM), Madrid, Spain
[2]Institute of Geosciences (CSIC), Madrid, Spain
[3]Data Mining Research Group, University of Salamanca (USAL), Salamanca, Spain

Correspondence: Víctor Galván Fraile (vgalva01@ucm.es).

Spy4Cast v1.0: a Python Tool for statistical seasonal forecast based on Maximum Covariance Analysis

Pablo Duran-Fonseca[1] and Belén Rodríguez-Fonseca[1,2]

[1]Universidad Complutense de Madrid, Madrid, Spain
[2]Instituto de Geociencias IGEO, CSIC-UCM, Madrid, Spain

Correspondence: Pablo Duran-Fonseca (pdrm56@gmail.com). Belén Rodríguez-Fonseca (brfonsec@ucm.es).

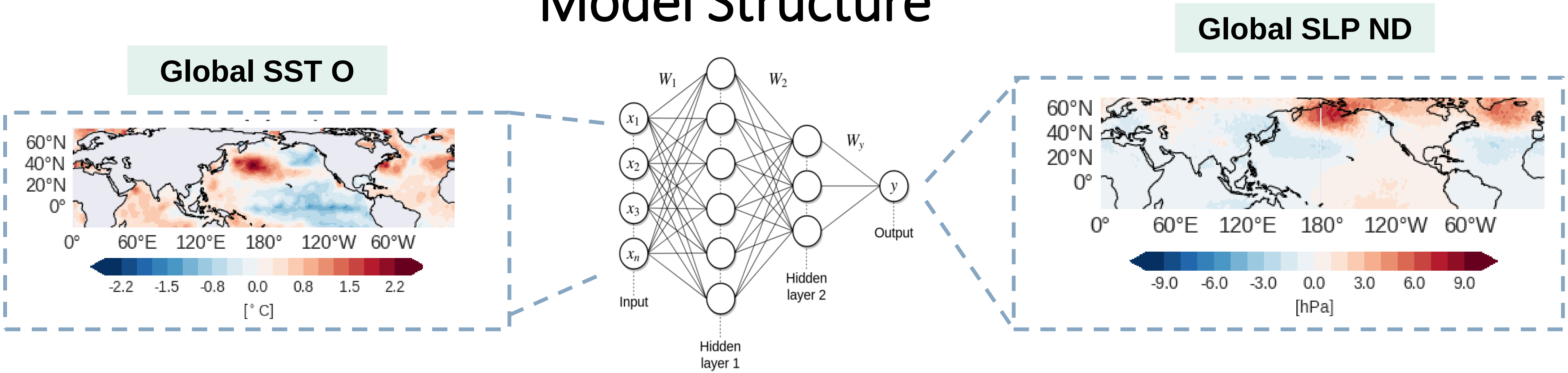
2. Data:

- **Predictor**: Global **SST** anomalies in **OCT**. (also analyzed from previous months) [HadISST].
- **Predictand**: Global **SLP** anomalies in **Nov-Dec** [ERA5].
- **Time period**: 1940-2019.

- **Leave-one-out Cross Validation** strategy.
- **Models:**

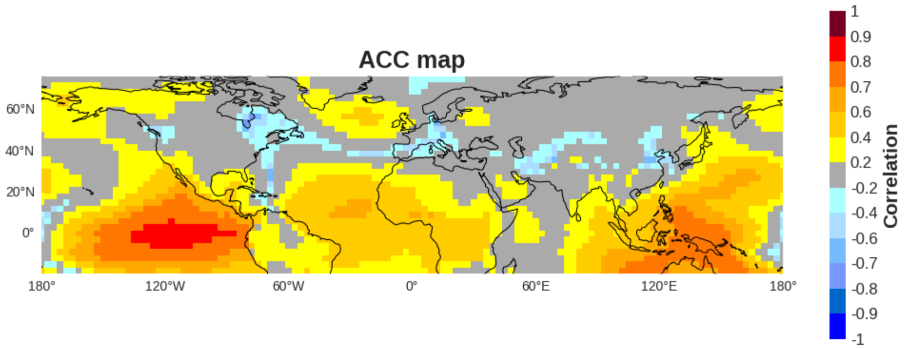
  1. MCA with 3 modes.
  2. NN with 6 hidden layers following an encoder-decoder architecture.

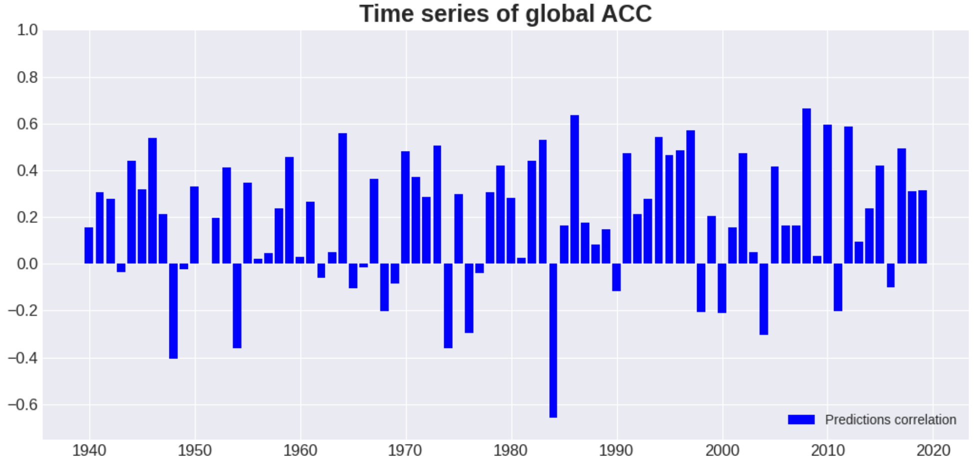
Model Structure

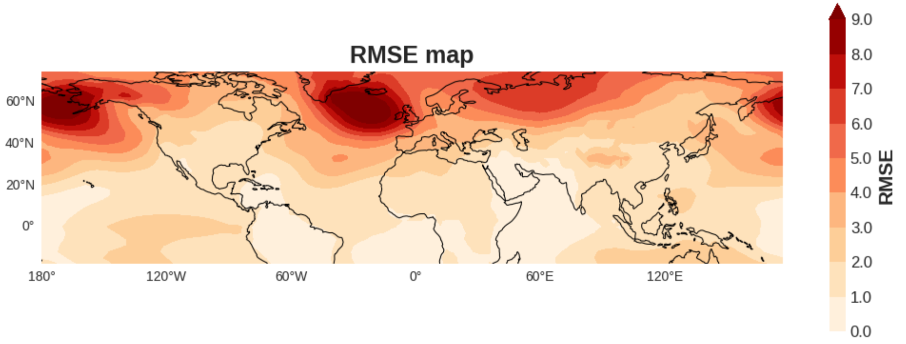


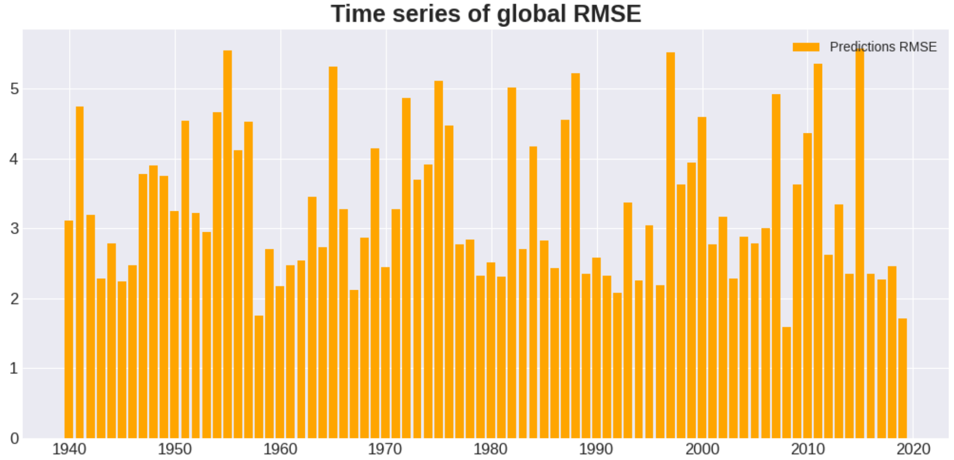
3. Application: **MCA** prediction of ND SLP

- Highest skill in the tropics.
- Skill in the North-Atlantic.
- Higher errors due to model definition.



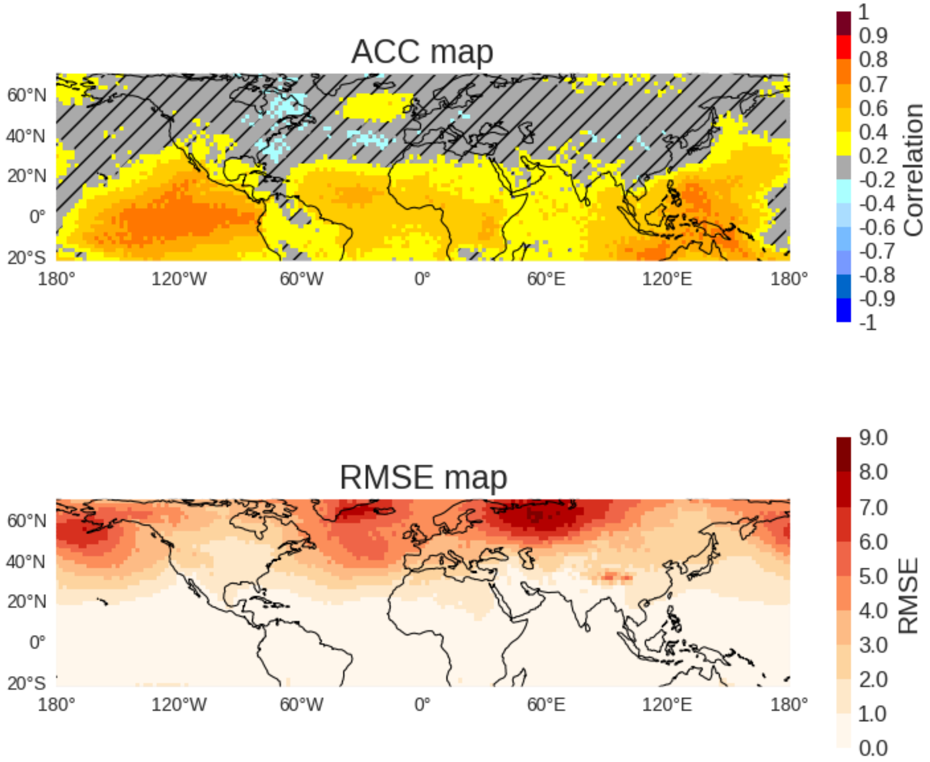


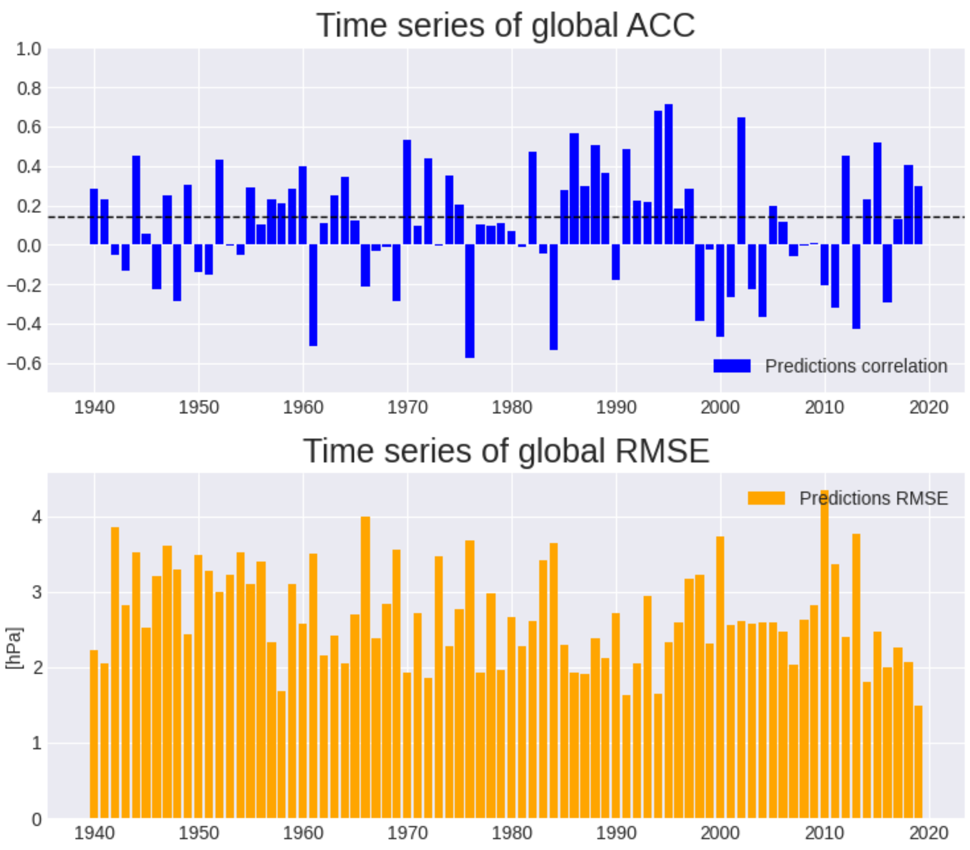




3. Application: **DNN** prediction of ND SLP

- Highest skill in the tropics.
- No skill in North-Atlantic.
- Interannual and decadal variability.
- Reduced errors due to model definition.



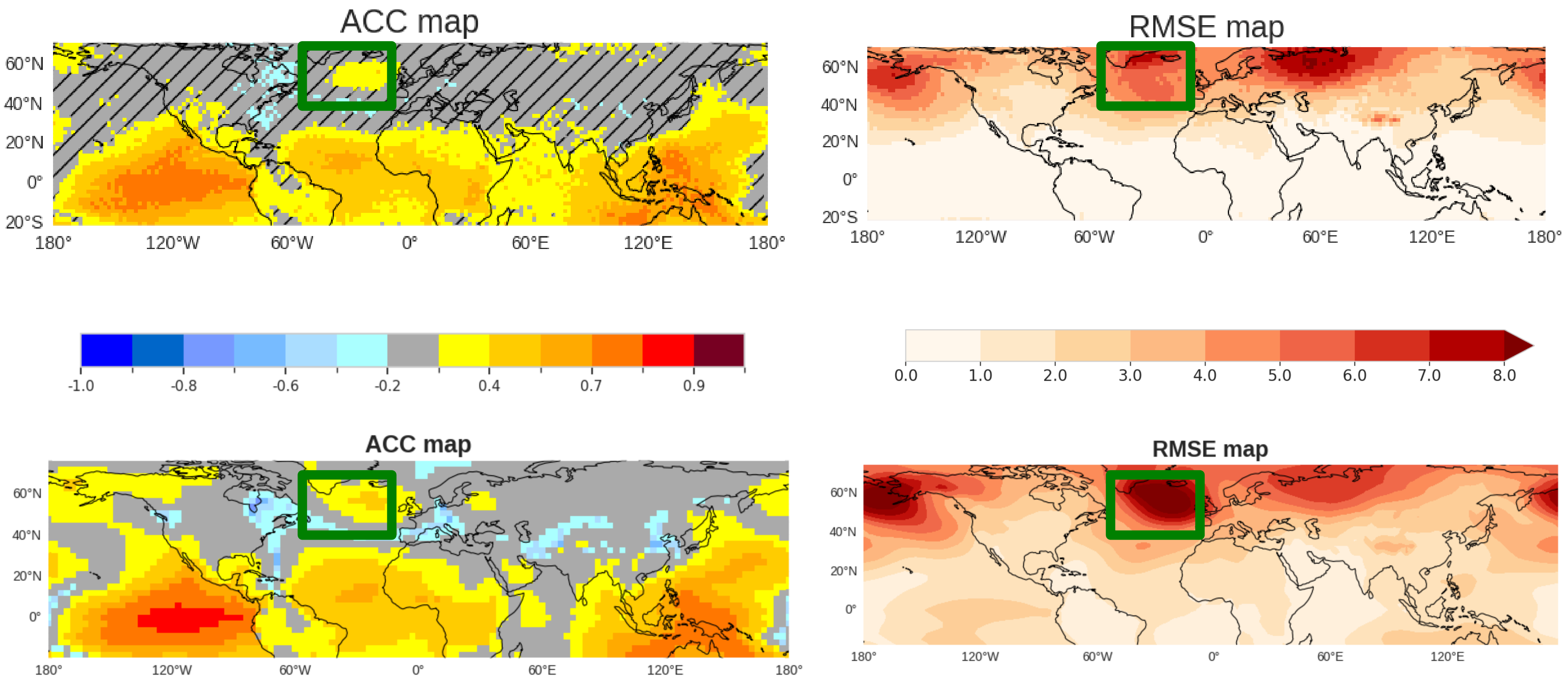


3. Application: Prediction of **ND SLP**

NN

Hindcast period: 1940-2019

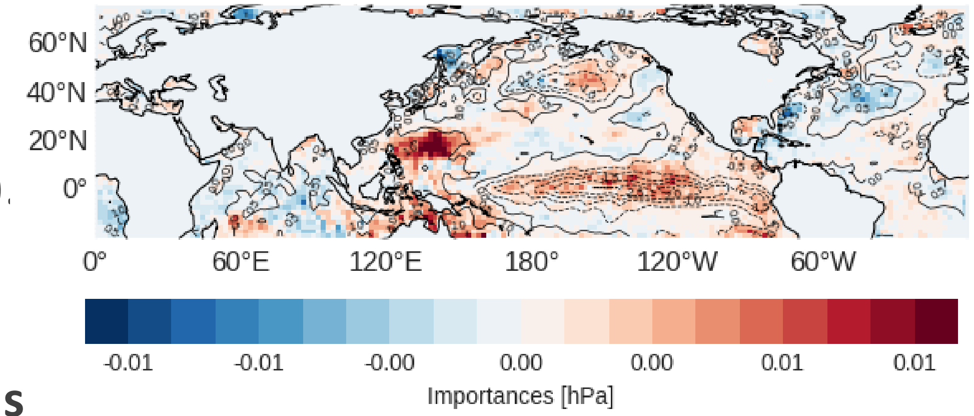
MCA

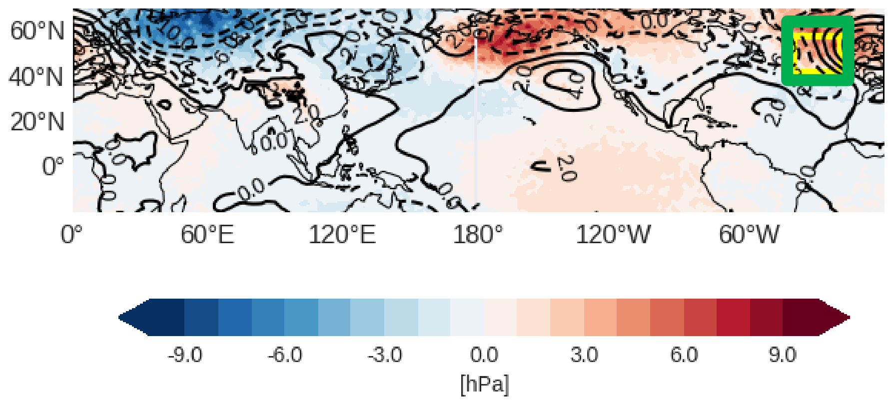


3. Application: **DNN** attributions of ND SLP

- Analyze concrete predictions (1988, 1987).
- Study **model attributions** (Integrated Gradients).
- Discover predictable signals?

La Niña 1988

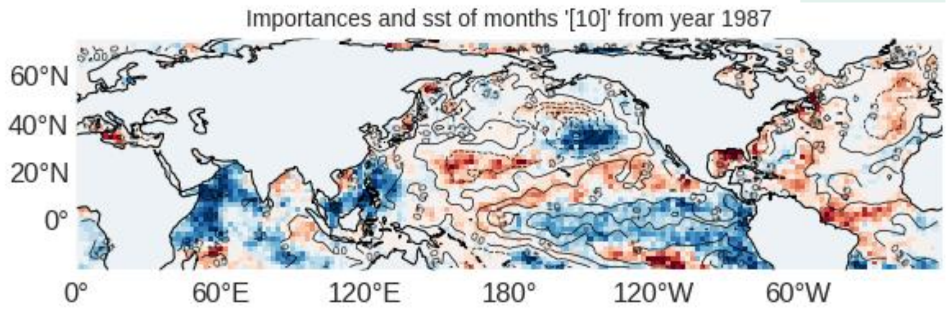


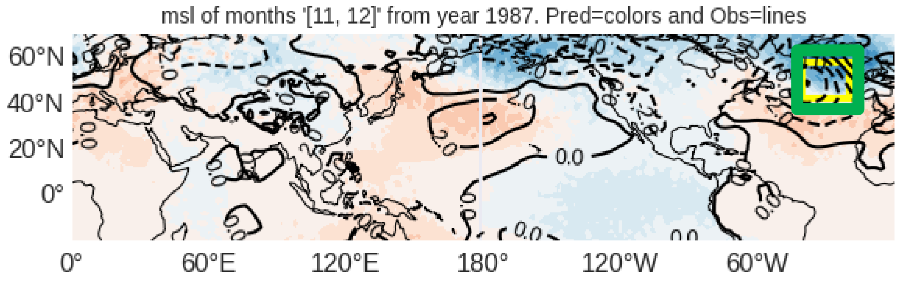


Contour: SST
Shaded: Importances

Contour: SLP Observed
Shaded: SLP Predicted

El Niño 1987



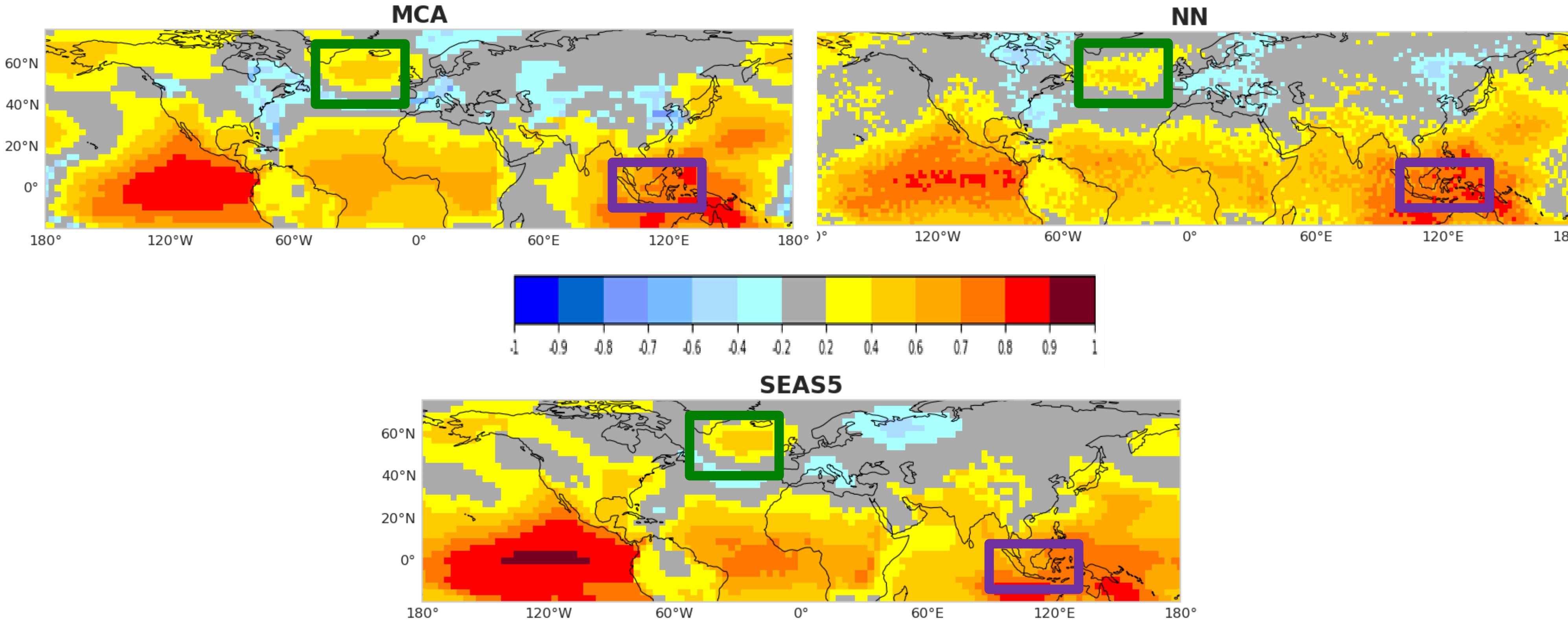


3. Application: **Models Comparison**

Initialization: Oct
Period: 1981-2016

- Warm Pool Index.

- North Atlantic Index



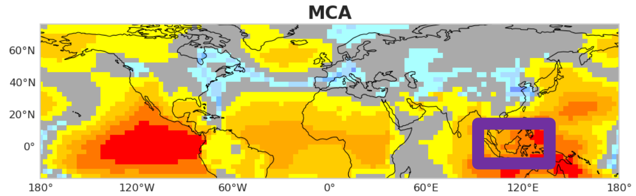
3. Application: **Models Comparison**

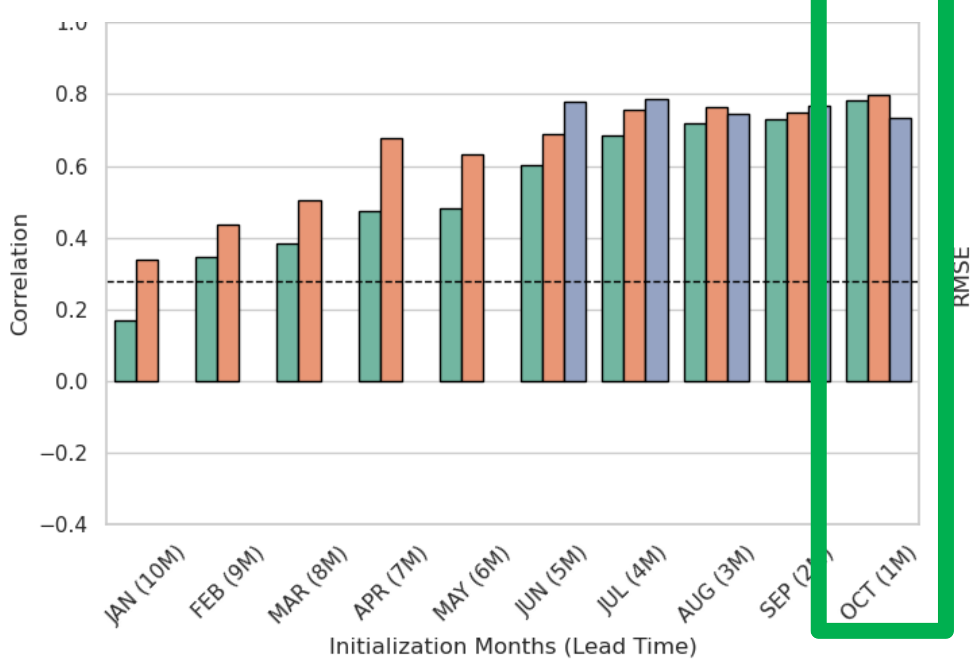
Initialization: Oct
Period: 1981-2016

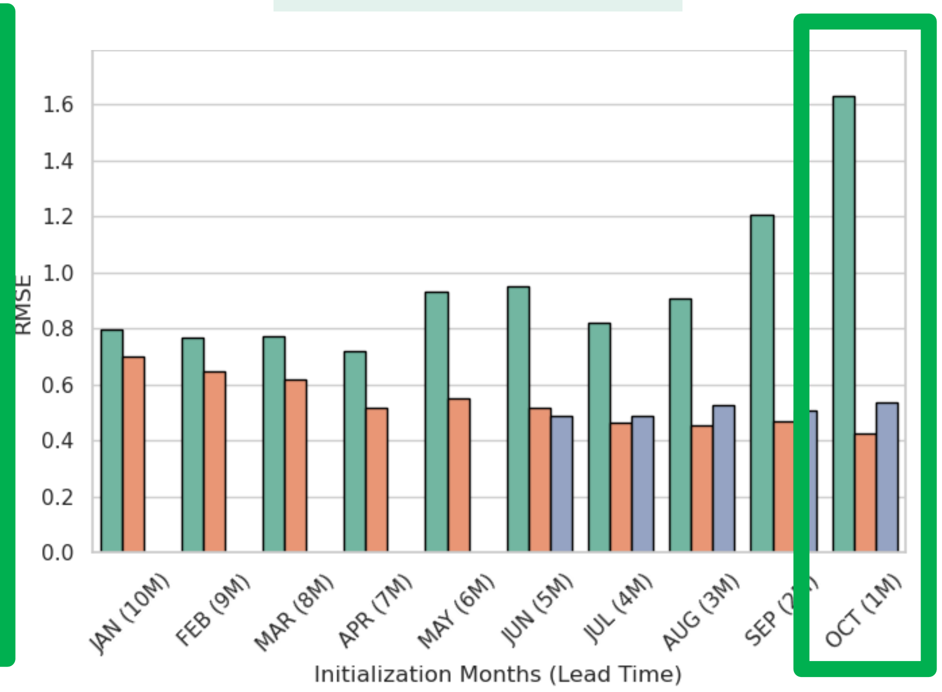
- Warm Pool Index.

Correlation

RMSE



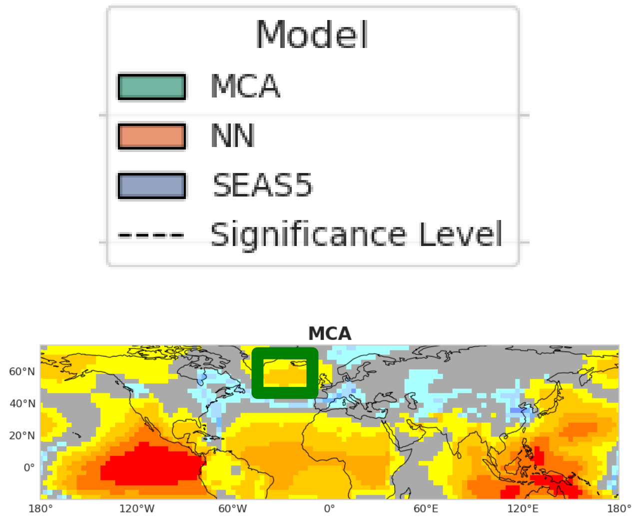




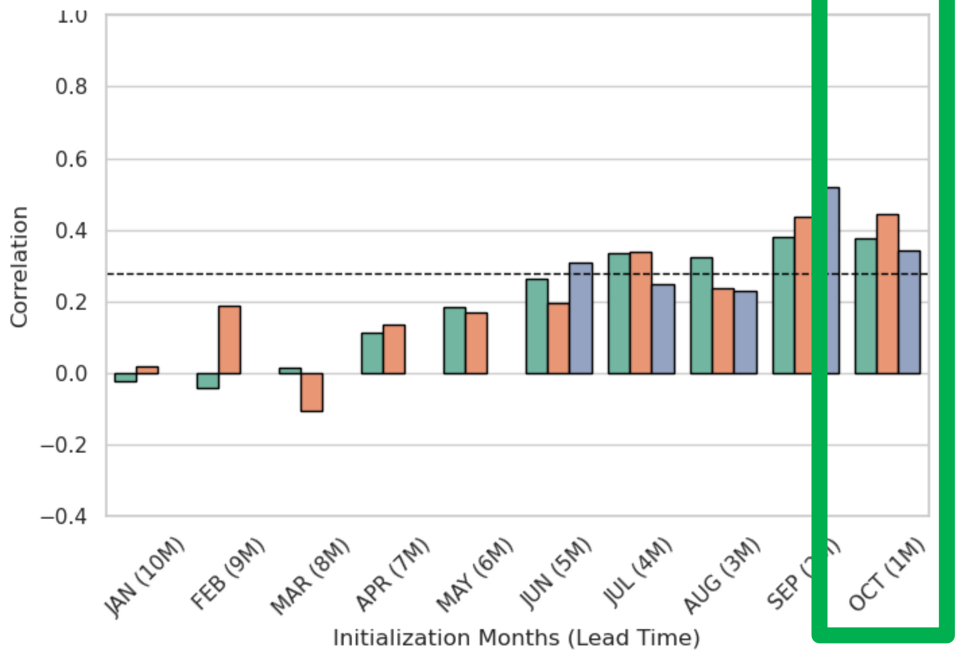
3. Application: **Models Comparison**

Initialization: Oct
Period: 1981-2016

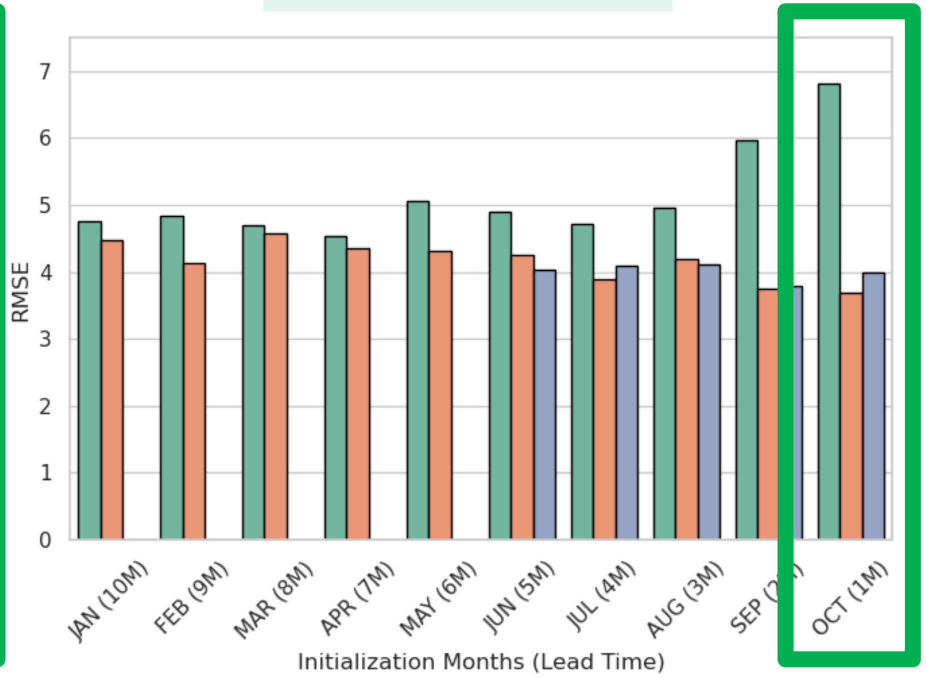
❖ North Atlantic Index



Correlation



RMSE

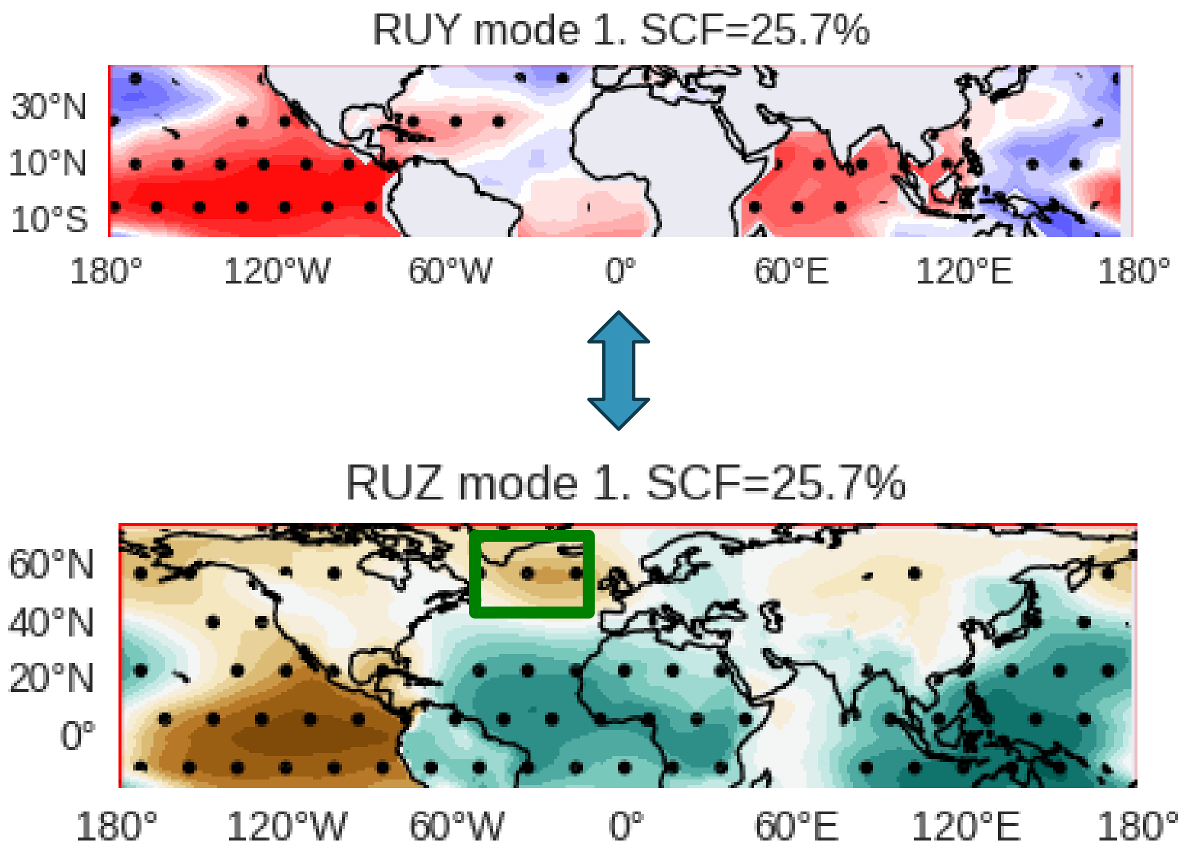


3. Application: **MCA Model Attributions**

Initialization: Oct
Period: 1981-2016

- Coupled modes of covariability between SST and SLP anomalies.

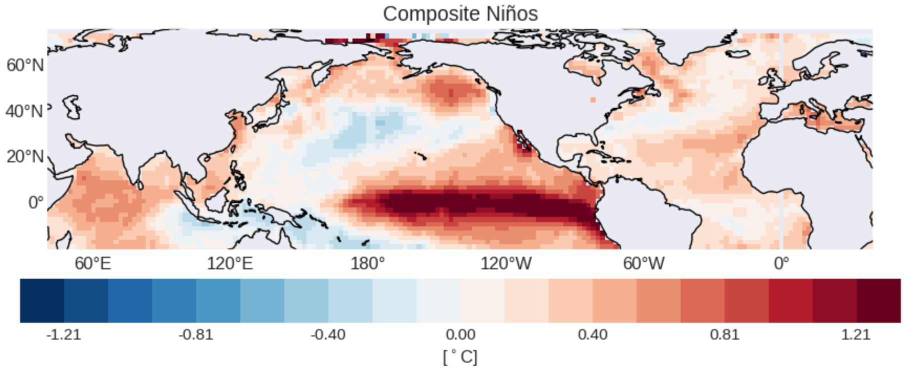
- ENSO mode that impact in the North Atlantic region.

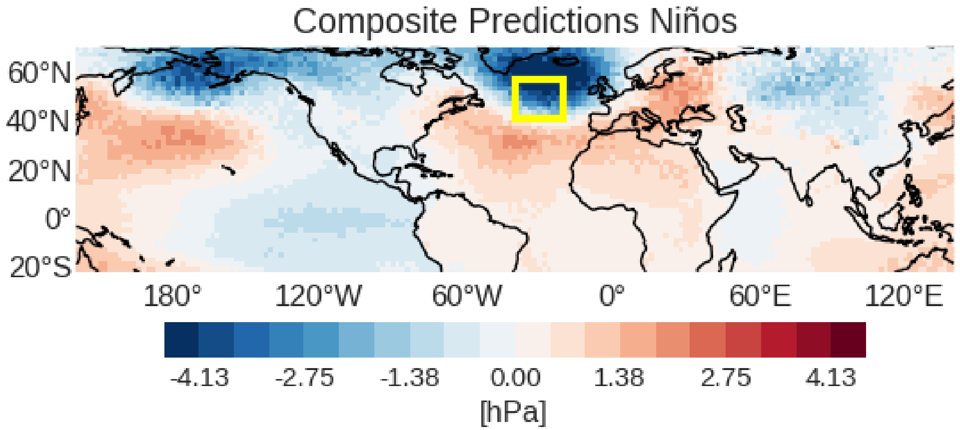


3. Application: **NN Model Attributions**

Initialization: Oct
Period: 1981-2016

- North Atlantic Index.
- **Niño years composite** (Niño 3.4 index).
- **Model attributions** (discover predictable signals)**.**
- Adnan et al. (2021), Joshi et al. (2021).

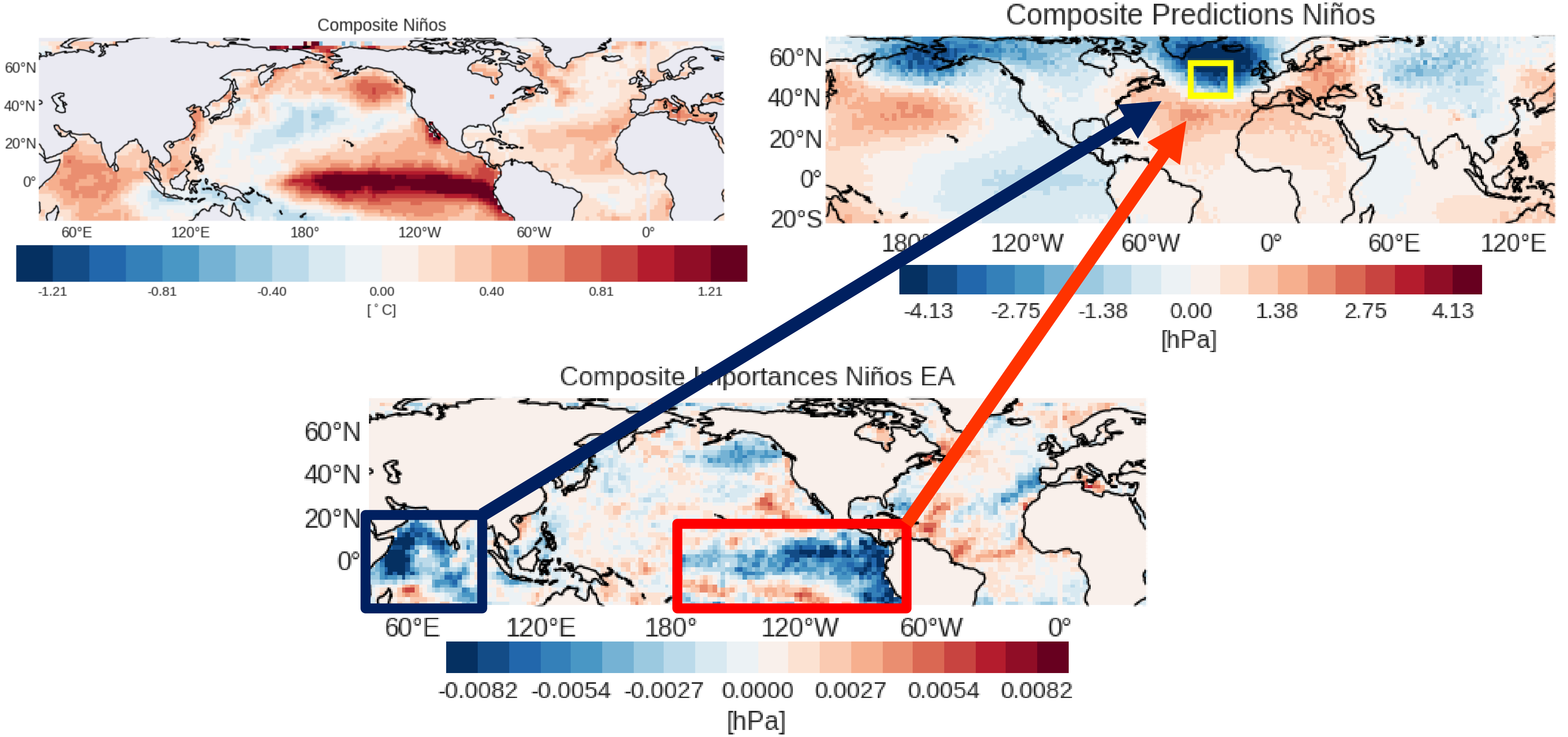




3. Application: **NN Model Attributions**

Initialization: Oct
Period: 1981-2016

- North Atlantic Index.
- **Niño years composite** (Niño 3.4 index).
- **Model attributions** (discover predictable signals).
- Adnan et al. (2021), Joshi et al. (2021).



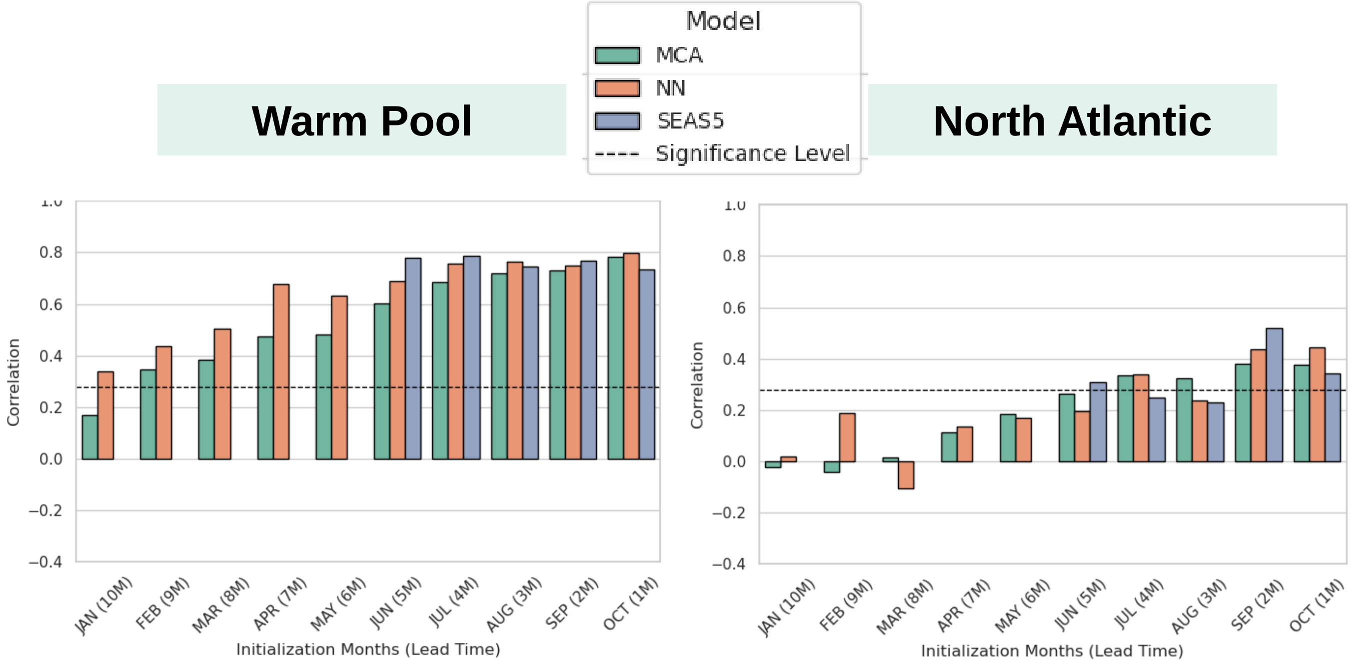
4. Conclusions:

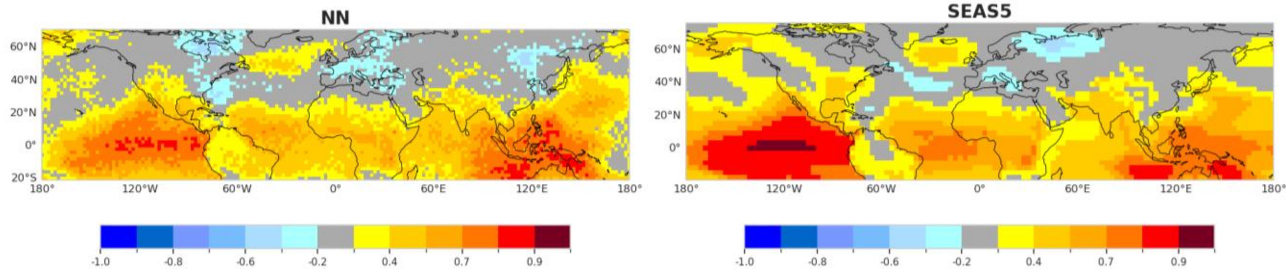
I. Statistical models' ability to capture non-stationary teleconnections.

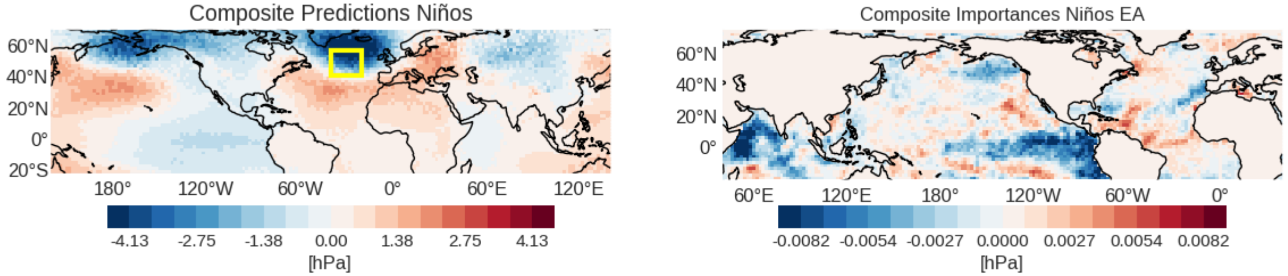
II. Statistically significant skill from previous winter for tropical regions

III. Similar results compared with dynamical based model.

IV. Model attributions point out regions of predictable signals in accordance with recent studies.







THANK YOU FOR YOUR ATTENTION!

QUESTIONS ?

Víctor Galván Fraile, Marta Martín del Rey, Irene Polo, Belén Rodríguez Fonseca, Magdalena Alonso Balmaseda, Esteban Rodríguez Guisado, María Moreno García.

Email direction: vgalva01@ucm.es

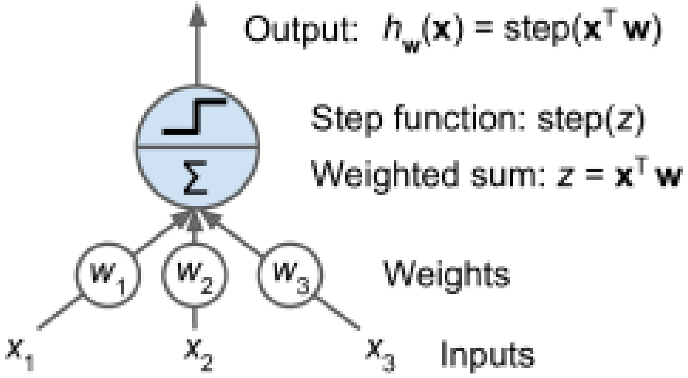
TROPA UCM

UNIVERSIDAD COMPLUTENSE MADRID

GOBIERNO DE ESPAÑA | MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

2. Methodology: Deep Learning (NN4CAST)

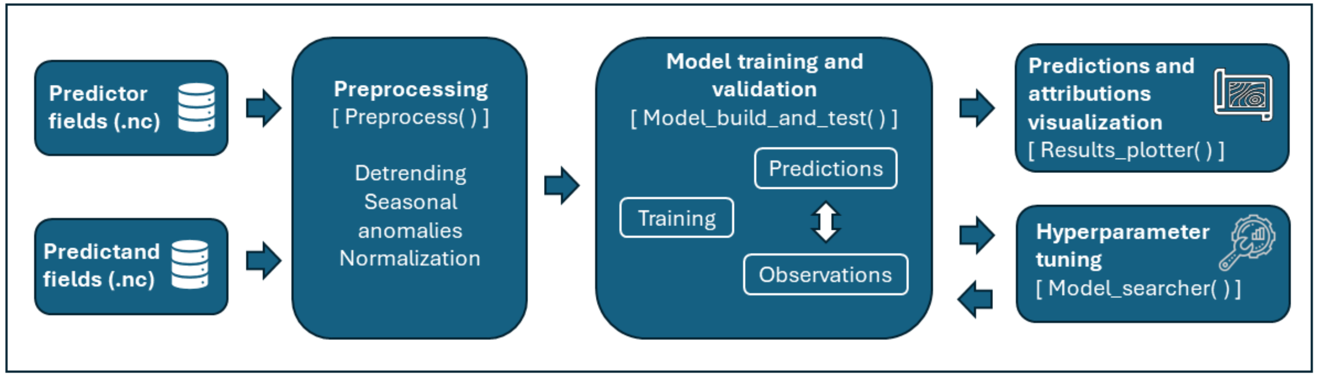
- Types of Machine Learning:
  - **Supervised learning.**
  - Unsupervised learning.
  - Reinforcement learning.



Source: Géron, A. (2022)

- **Artificial Neural Networks** (NNs):
  - Linear combination of inputs.
  - Non-linear activation function.
  - Input, output and hidden layers.
  - Universal function approximator.

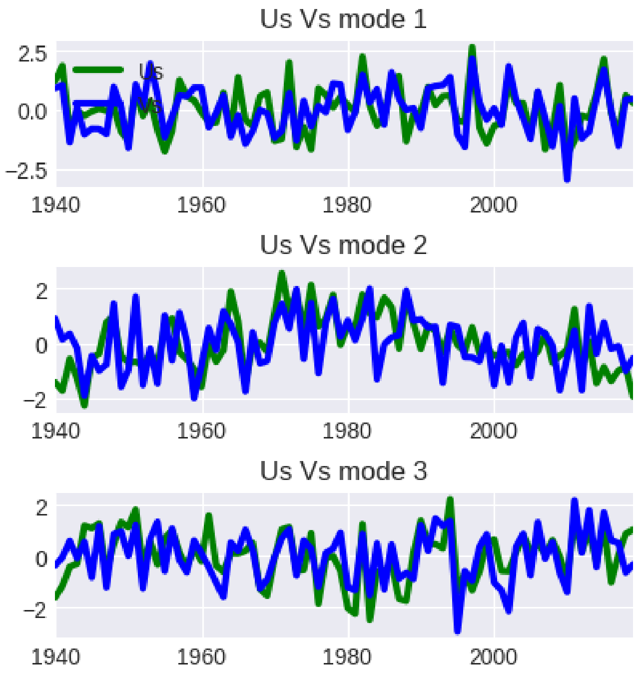
- **Objective**: **Minimize** the mean squared error during training.



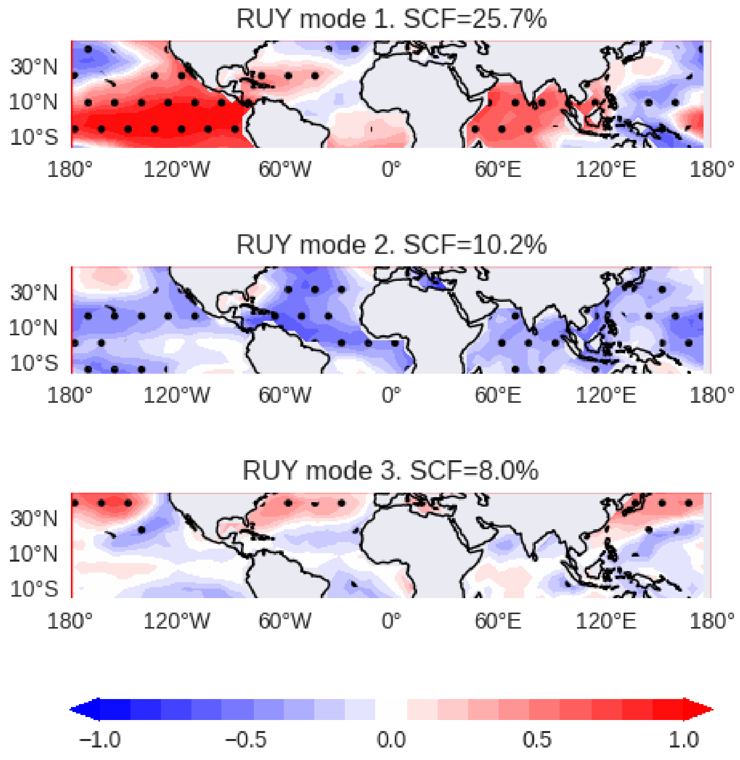
Source: Galván-Fraile, V. et al. (2024)

3. Application: **MCA** prediction of ND SLP

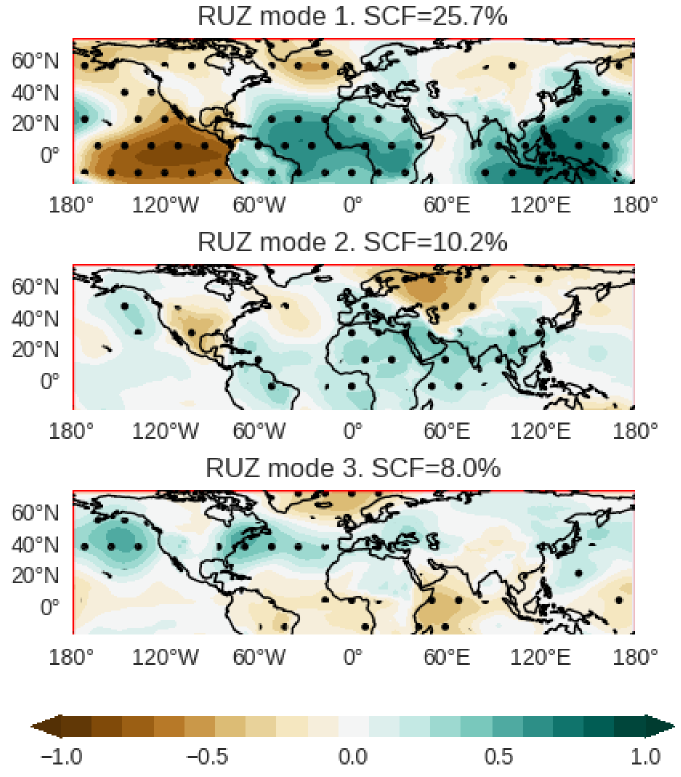
- Model hyperparameters (default):
  - Number of modes = 3.
- **Leave-One-Out Cross-Validation** method: period [**1940-2019**].



Predictor: Pacific SST O



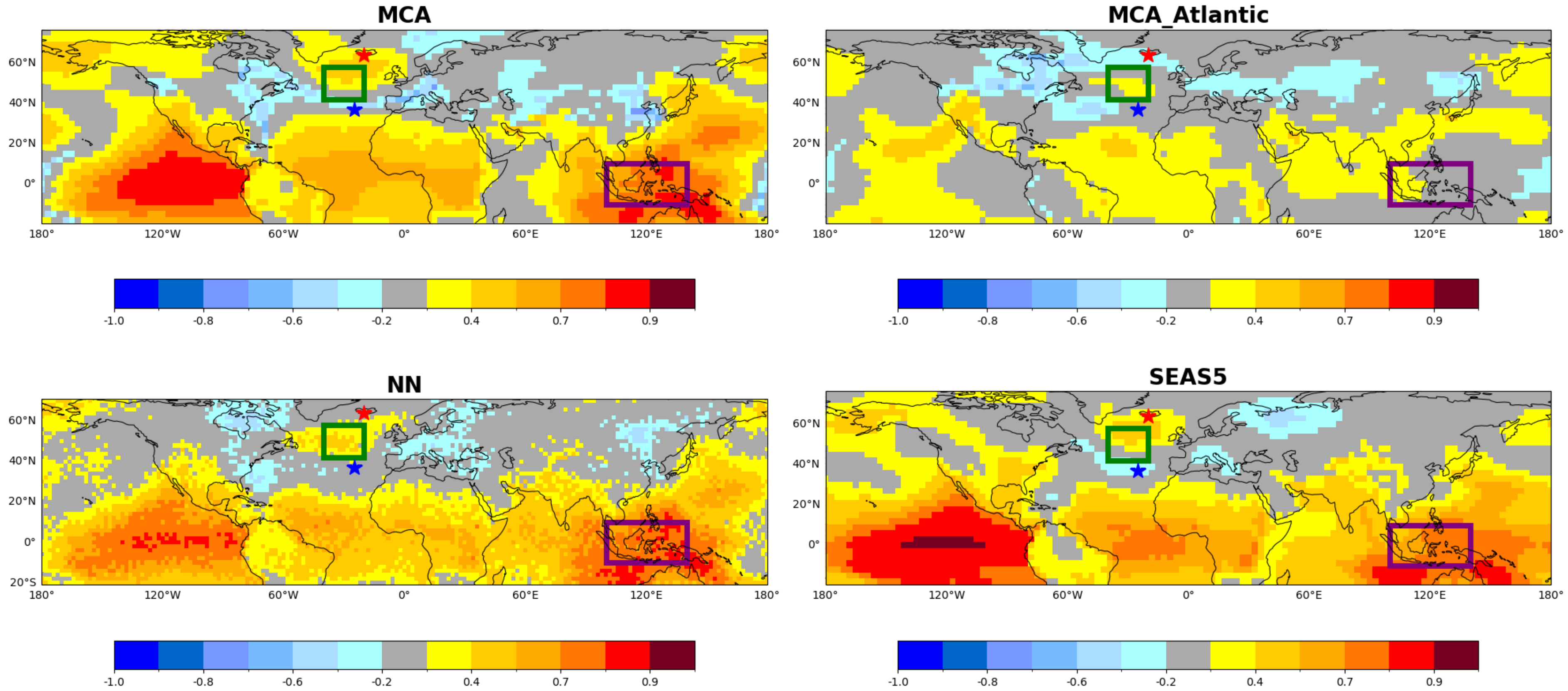
Predictand: Global SLP ND



3. Application: **Models Comparison**

Initialization: Oct
Period: 1981-2016

- Warm Pool Index.
- North Atlantic Index
- NAO Index



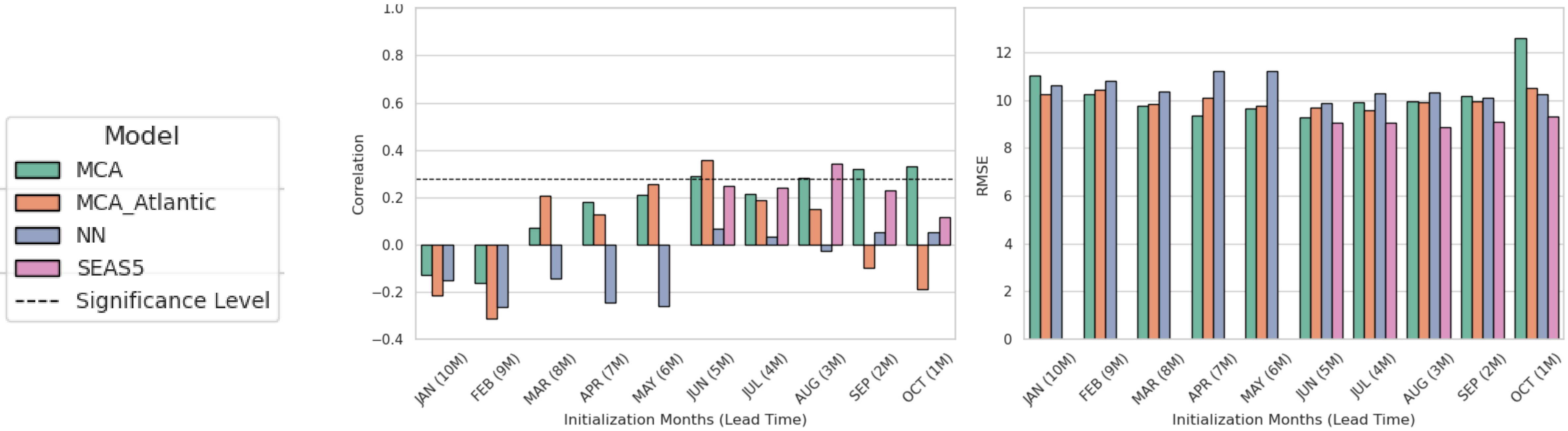
3. Application: **Models Comparison**

Initialization: Oct
Period: 1981-2016

- NAO Index ★ - ★

Correlation

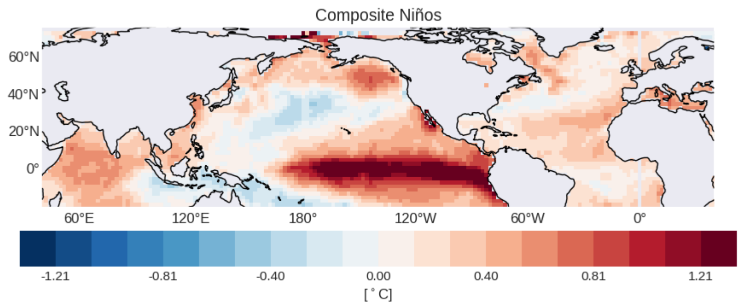
RMSE

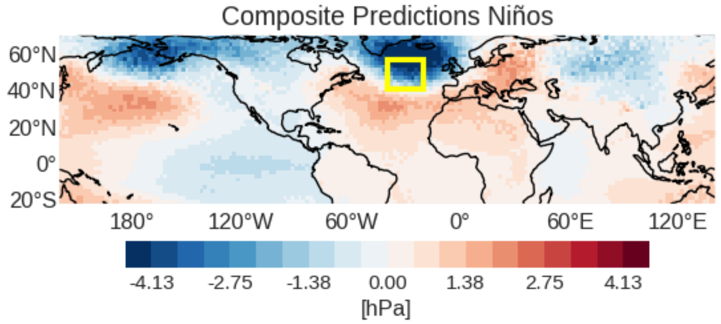


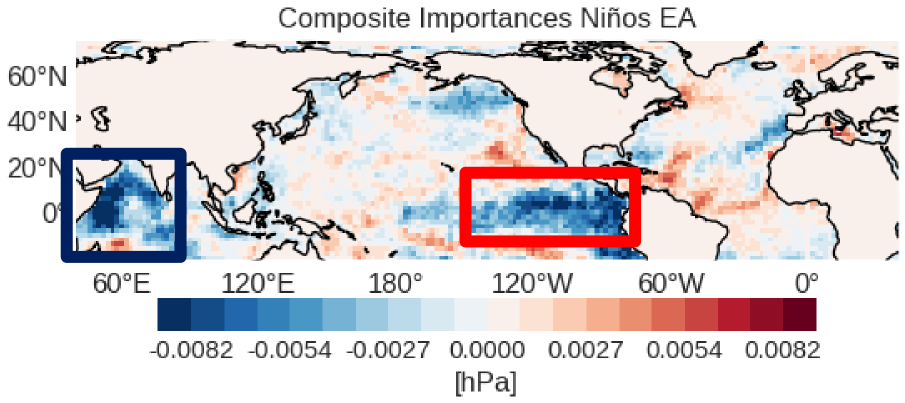
3. Application: **Model Attributions**

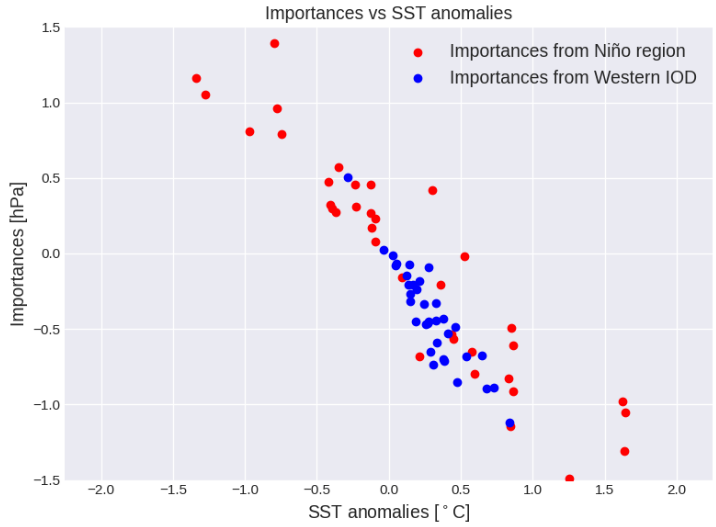
Initialization: Oct
Period: 1981-2016

- North Atlantic Index.
- **Niño years composite** (Niño 3.4 index).
- **Model attributions** (discover predictable signals).
- Adnan et al. (2021), Joshi et al. (2021).





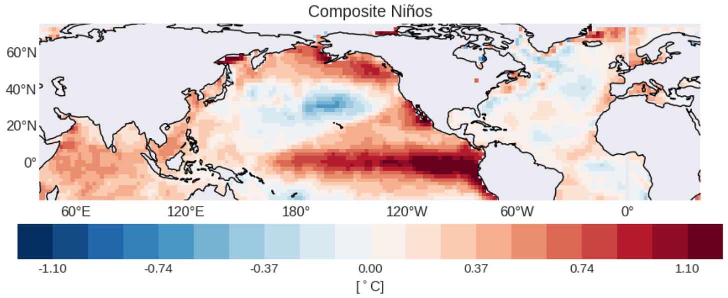


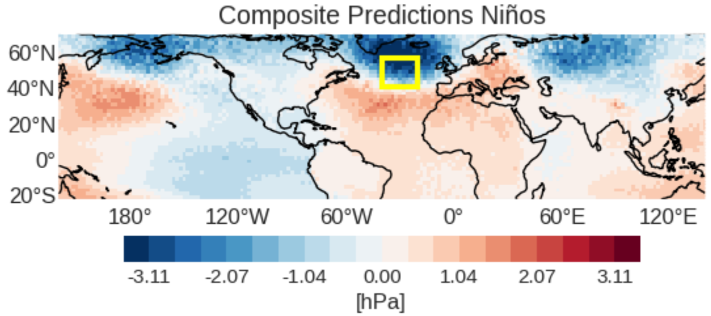


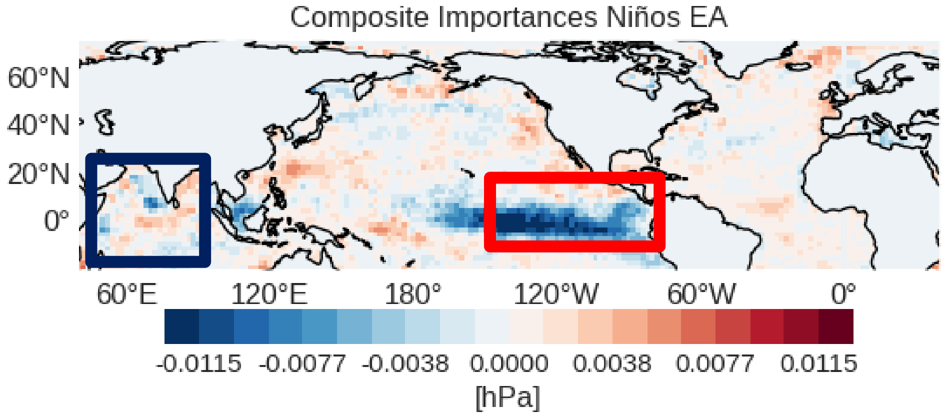
3. Application: **Model Attributions**

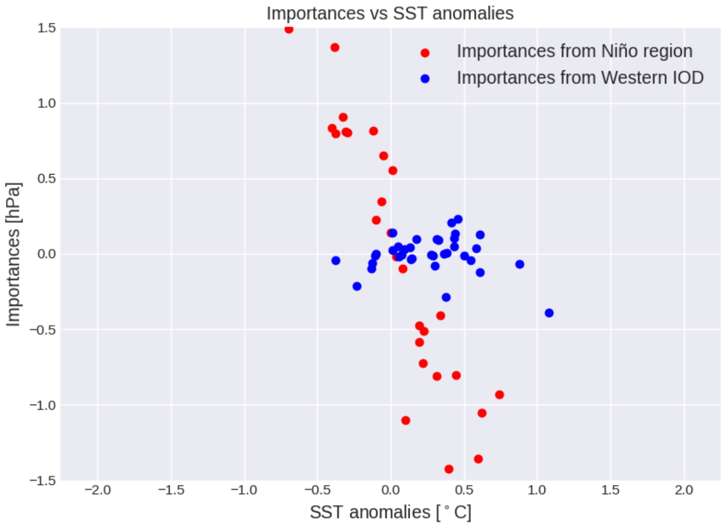
Initialization: JUN
Period: 1981-2016

- North Atlantic Index.
- **Niño years composite** (Niño 3.4 index).
- **Model attributions** (discover predictable signals).
- Adnan et al. (2021), Joshi et al. (2021).









3. Application: **Model Attributions**

Initialization: JUN
Period: 1981-2016

- North Atlantic Index.
- **Niño years composite** (Niño 3.4 index).
- **Model attributions** (discover predictable signals).
- Adnan et al. (2021), Joshi et al. (2021).