Predicting and Understanding Drought

#ESoWC2019

UNIVERSITY OF OXFORD

To use **ECMWF/Copernicus open datasets** to evaluate machine learning (ML) techniques to **better predict one specific kind of an extreme weather event**, e.g. drought or hurricanes; provide templates for future ML work

ECMWF

#ESoWC2019

Project Overview

Robust, useful pipeline

Predictive System for Agricultural Drought

Interpretable Machine Learning

Communicate Results

Robust, useful pipeline

Export / download → Engineering → Model Training → Analysis

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Agricultural Drought

Target: Vegetation Health Index (VHI) at end of Season

Inputs: Meteorology (Temp, Precip), Hydrology (Soil Moisture), Climate Vars (Nino3.4), Static Variables (Orography)

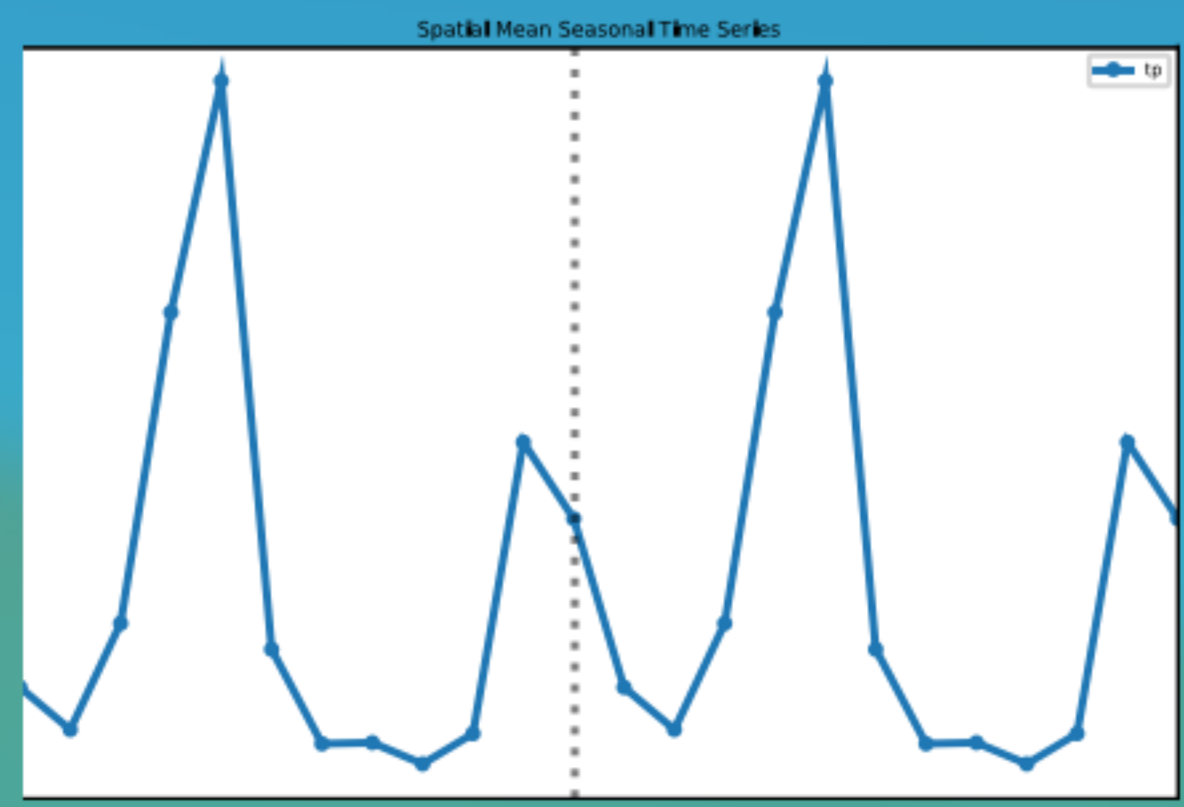
Challenges: Encoding spatial-temporal information, how to utilise climate variables, masking, Vegetation health observed from satellites?

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System



ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Decision Trees

Gradient Boosted Trees: XGBoost
Random Forest: Scikit-learn

Neural Networks

RNNs: PyTorch
Linear Models: PyTorch
Segmentation Models: PyTorch

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Aim 1: Forecast

From preceding conditions can we forecast a value ahead of time? This requires us to learn the relationship between the previous conditions and the current conditions.

Goal: Predictive Accuracy

Challenge: Are we reproducing / competing with physically based models?

UNIVERSITY OF OXFORD

Predictive System

Aim 2: Identify Correlations

Can we use the high dimensional fields (SST, SLP) to identify connections between remote regions in space and time.

Goal: Identify climate drivers of variability from data

Challenge: Crazy high dimensions (atmospheric levels [z], geographic regions [x,y], lagged in time, multiple variables - 5D space)

UNIVERSITY OF OXFORD

Predictive System

Aim 1: Forecast

Operationalise

$$f(X_{train}) =$$

$$f(X_{SEAS5}) =$$

where

precip temp soil M

$X =$

t-1
t-2
...
t-n

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Aim 2: Identify Correlations

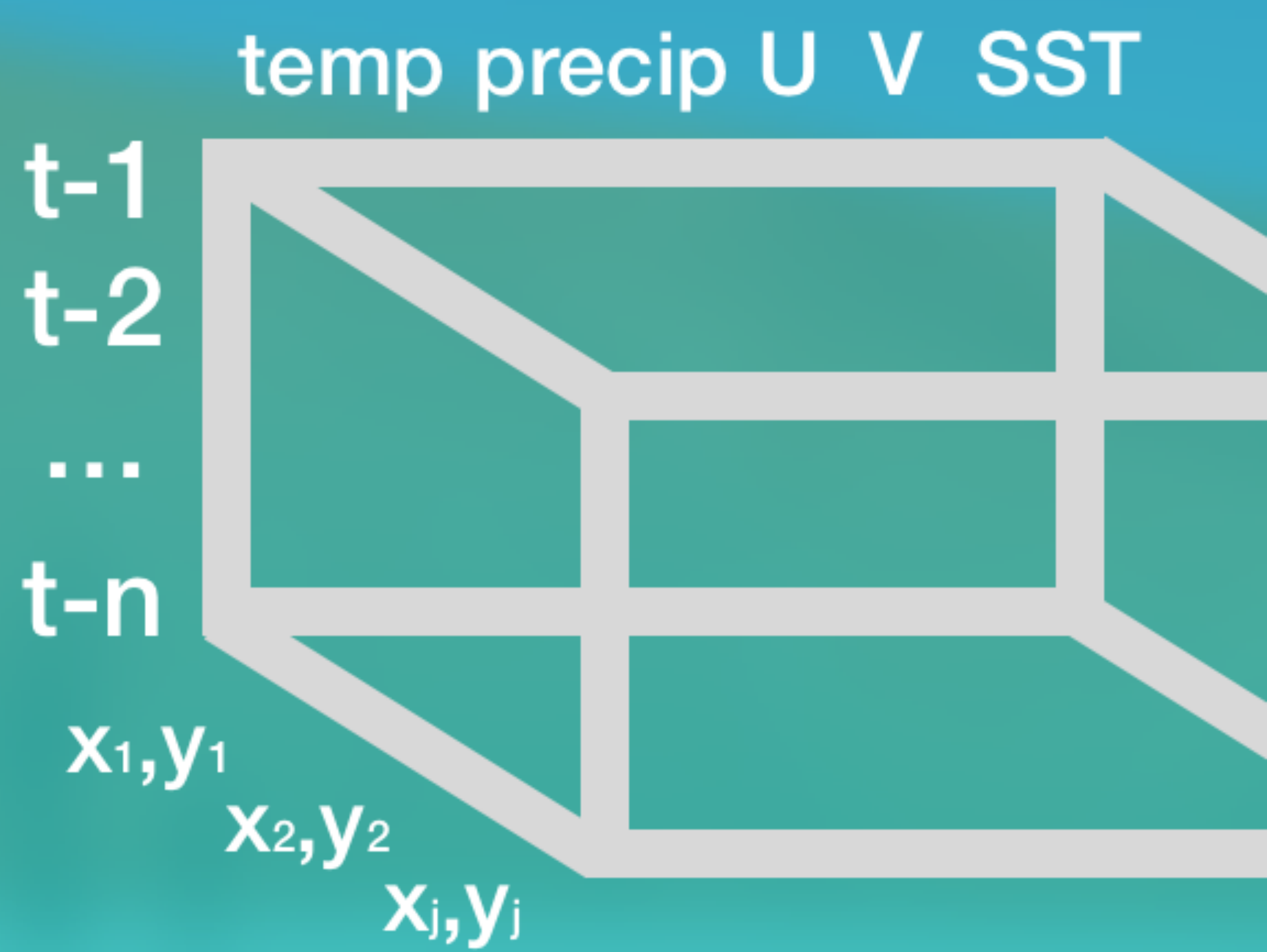
Operationalise

$$f\,(\,X_{train}\,) =$$

$$f\,(\,X_{now}\,) =$$

where

$X =$



ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

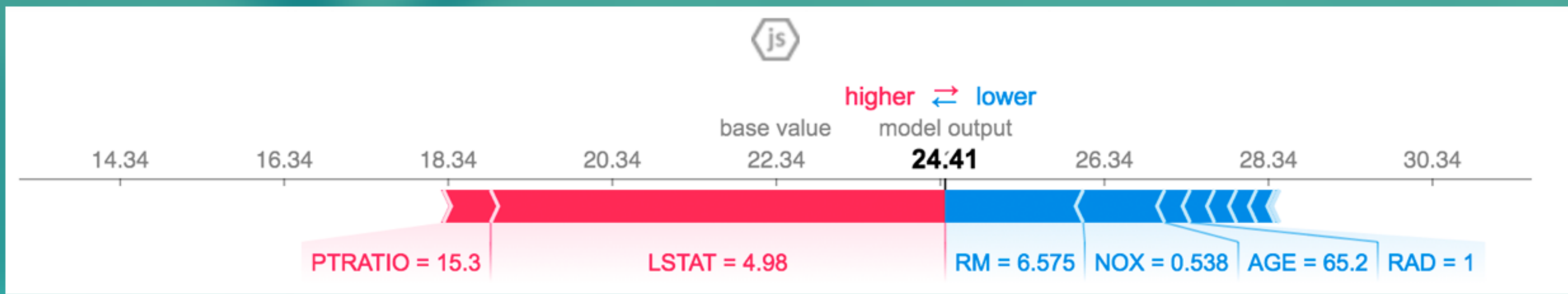
Interpret Models

SHAP Values

A method for assigning payouts to players (features) depending on their contribution to the total payout (predictive accuracy).

$$\text{importance of } j = f\ (\text{with } j) - f\ (\text{without } j)$$

Visualisations:



sum of all SHAP values = difference of prediction - base
(24.41 - 22.34)

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Interpret Models

Where do we find 'predictive skill' (geographically)?

How do we compare with physically-based forecasts?

Which of our features give us the best improvement in skill?

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Experiments

- choice of thresholds
- choice of definition
- choice of variable

Drought metrics vs. Drought impacts

Identify Teleconnections in SST/SLP data

Does soil moisture offer predictability for rainfall?

Can we quantify the human element in drought risk?

Combine ML Vegtation with SEAS5 Precipitation

UNIVERSITY OF OXFORD

Communicate Results

Flexible Pipeline

An experimental pipeline for easily exchanging models, definitions and input features.

Extensible: Python classes and funtions

Well Documented: Example notebooks, test suite

ECMWF

#ESoWC2019

Communicate Results

Blog Posts

Blog posts throughout the process outlining our thinking and our implementations.

More polished blog posts to follow towards the end of the ESoWC Project

#ESoWC2019

UNIVERSITY OF OXFORD

Communicate Results

Academic Papers

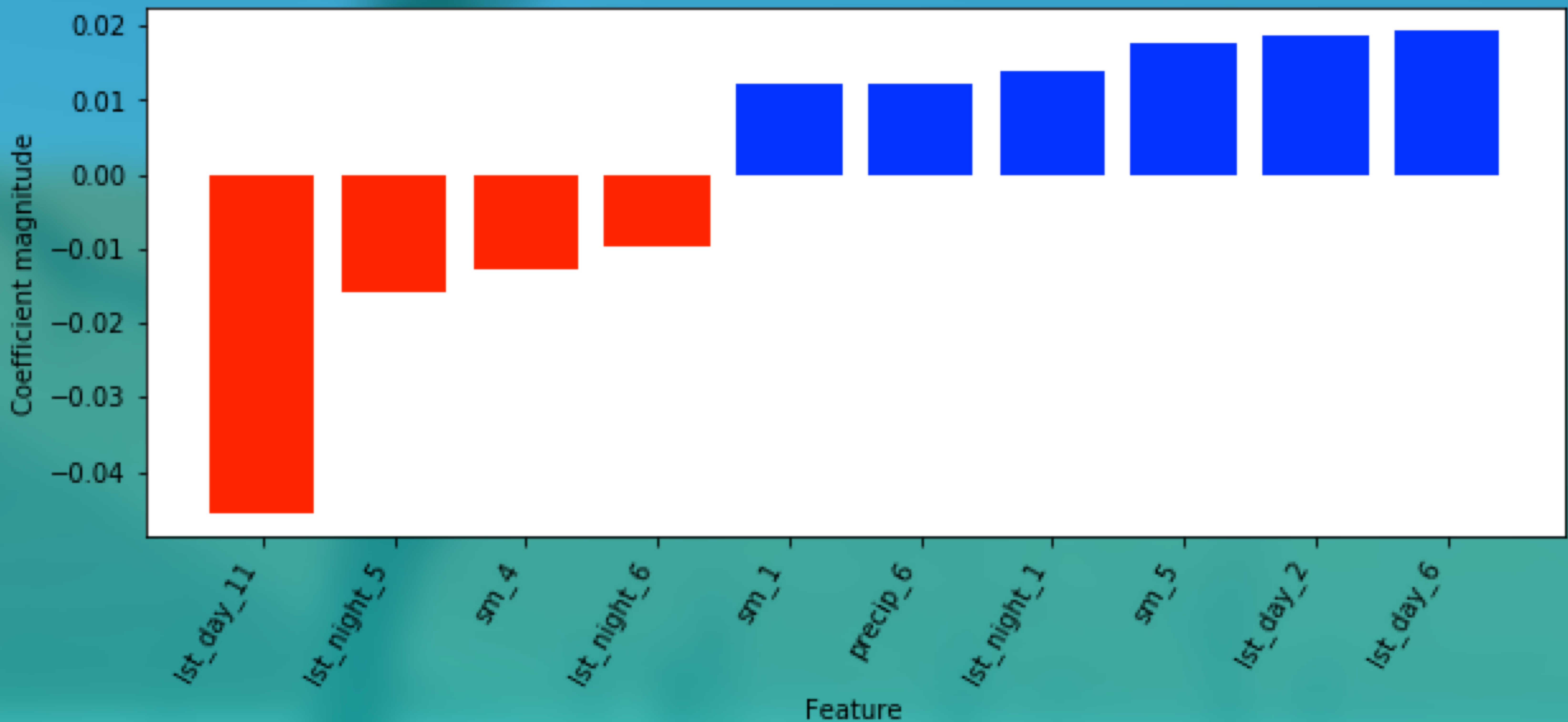
- Overview of the pipeline and functionality (JOSS, GMD).
- Overview of predictive accuracy and the algorithmic approach.
- Overview of scientific insights (e.g. SSTs in Western Pacific Warm Pool)

ECMWF

#ESoWC2019

(VERY) initial results

Input variables (linear model)

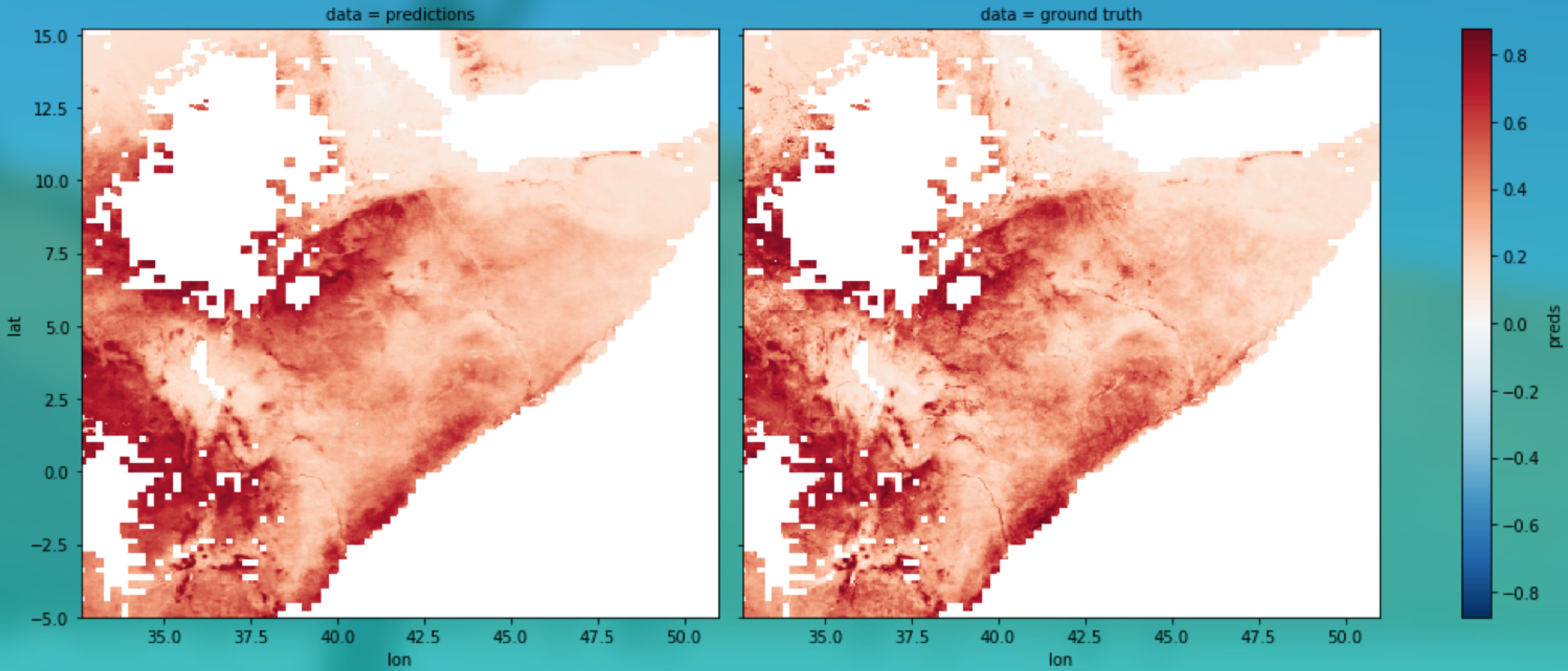


ECMWF

#ESoWC2019

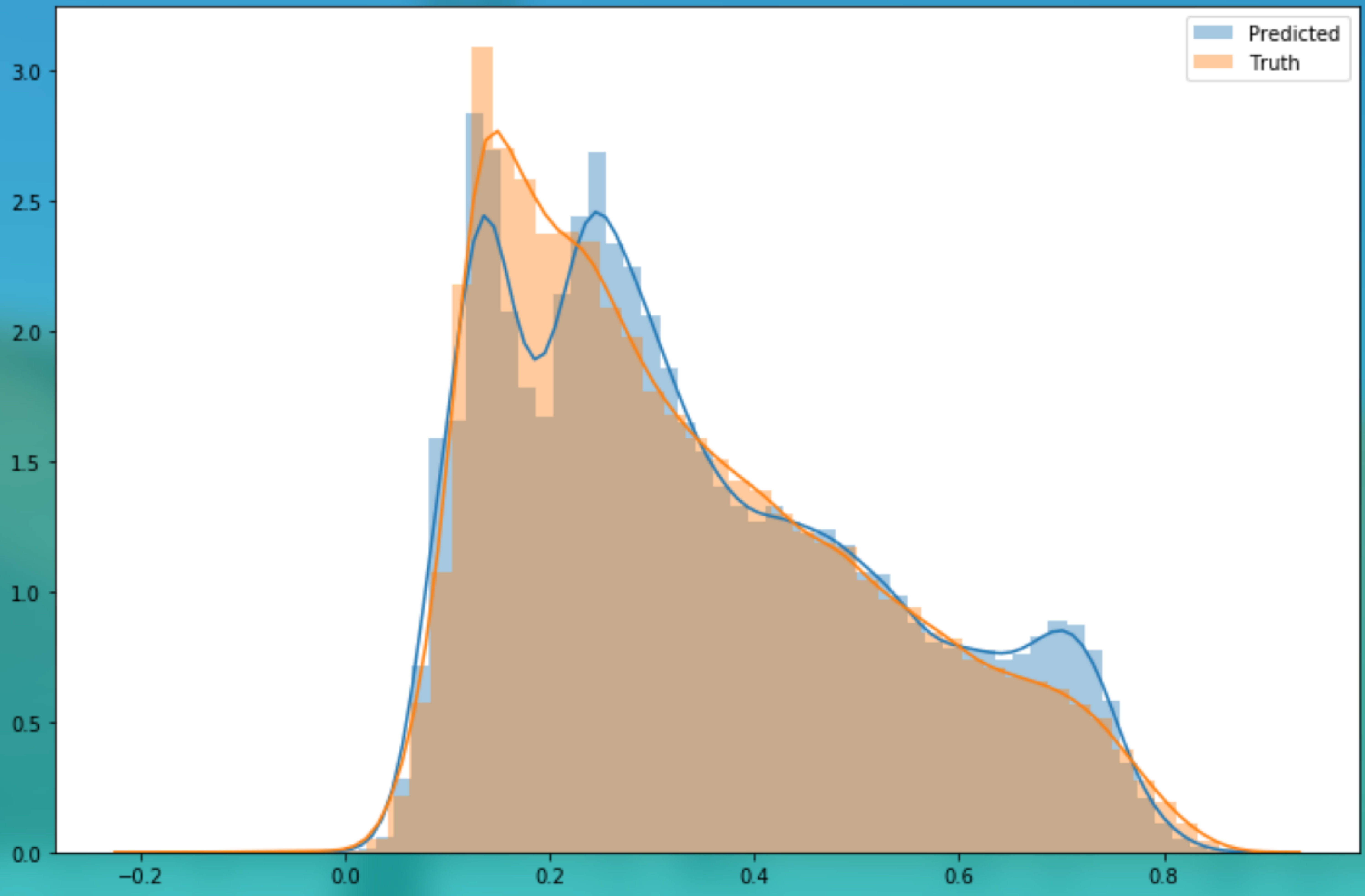
(VERY) initial results

Predictions vs. Truth



#ESoWC2019

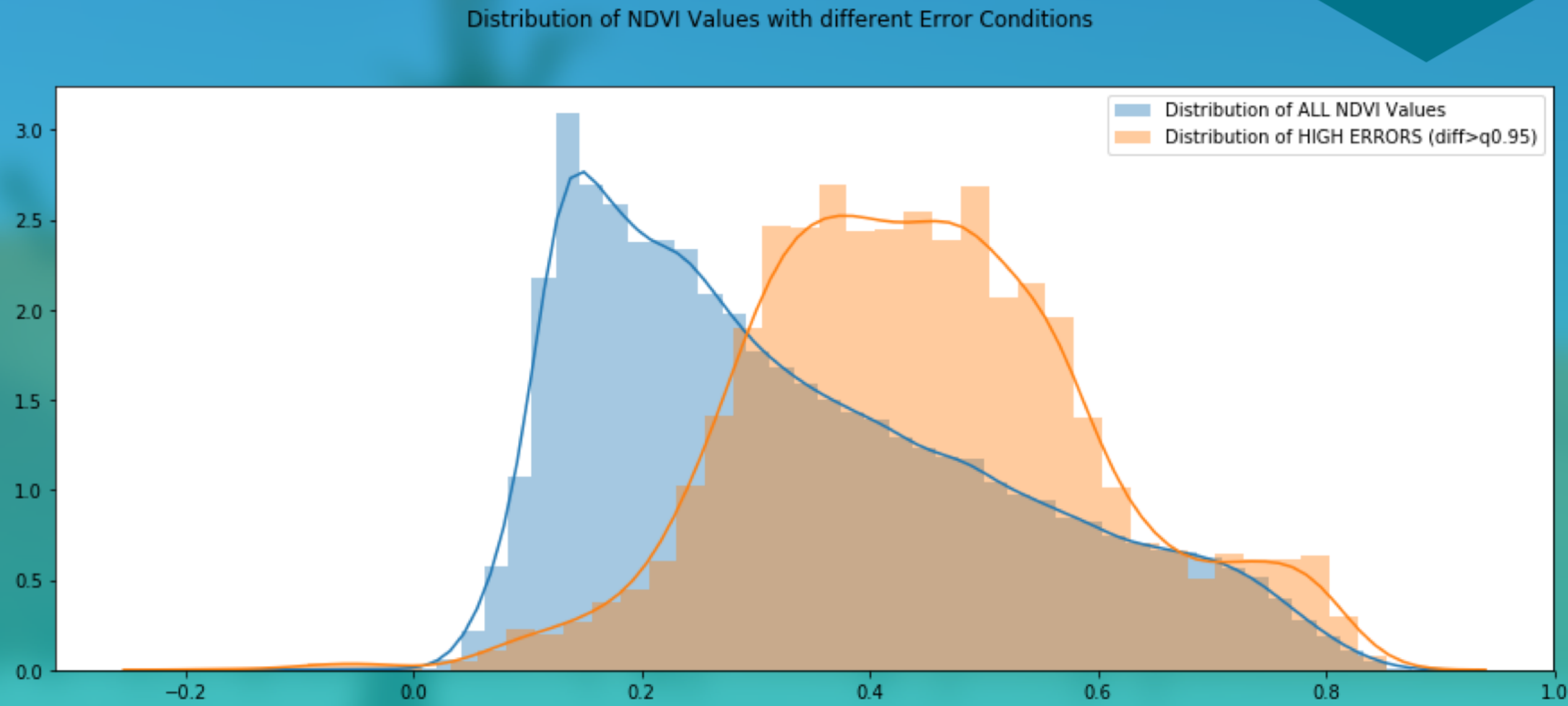
(VERY) initial results



ECMWF

#ESoWC2019

(VERY) initial results



ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

https://github.com/esowc/ml_drought

@tommylees112

ECMWF

#ESoWC2019

Appendix

UNIVERSITY OF OXFORD

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Hydrological Drought

Target: Soil Moisture Time Series

Inputs: Meteorology (Temp, Precip), Vegetation Health (NDVI, VHI), Climate Vars (Nino3.4), Static Variables (Orography, Soil Type)

Challenges: Encoding spatial-temporal information, how to utilise climate variables, masking, is the data correct?

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Meteorological Drought

Target: Standardised Precipitation Index (SPI)

Inputs: Preceding Meteorology (Temp, Precip), Hydrology (Soil Moisture), Climate Vars (Nino3.4), Static Variables (Orography, Soil Type)

Challenges: Encoding spatial-temporal information, how to utilise climate variables, Performance vs. SEAS5

ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Predictive System

Hydrological Drought 2

Target: Streamflow

Inputs: Preceding Meteorology (Temp, Precip), Hydrology (Soil Moisture), Climate Vars (Nino3.4), Static Variables (Orography, Soil Type)

Challenges: Encoding spatial-temporal information, how to utilise climate variables, Performance vs. SEAS5

ECMWF

#ESoWC2019

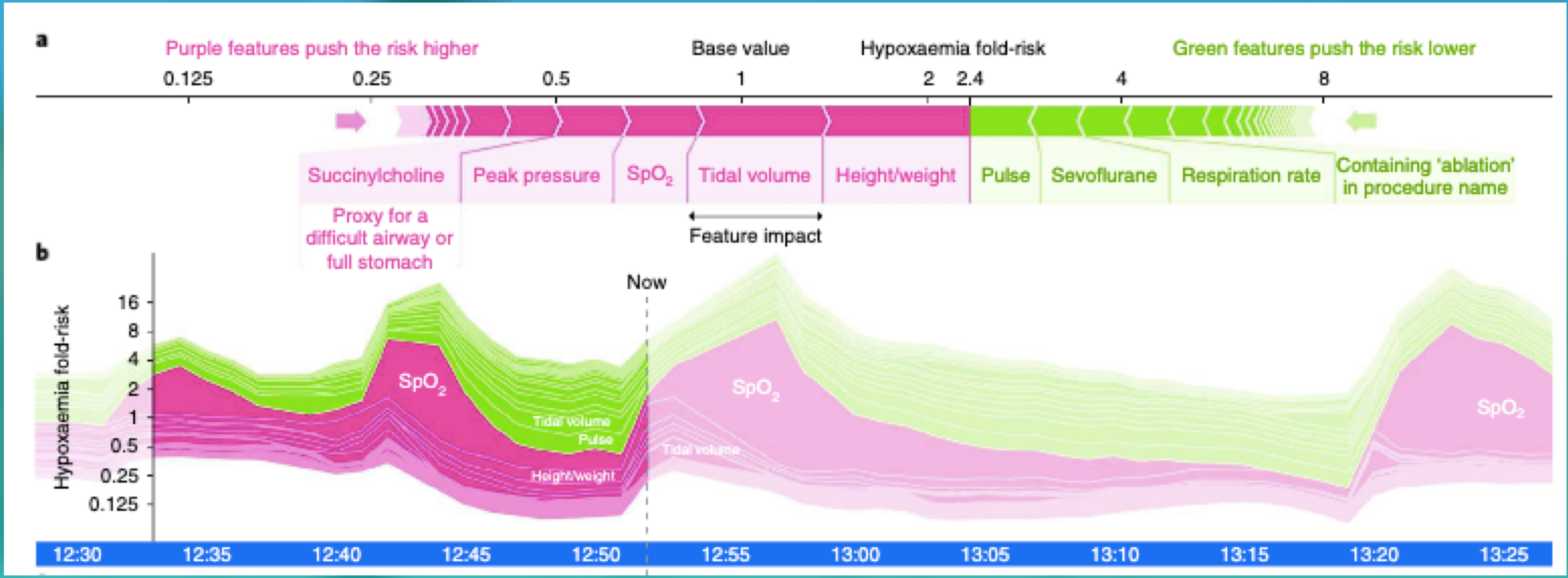
Interpret Models

SHAP Values

Package: SHAP (https://github.com/slundberg/shap)

Interpret Models

SHAP Values



ECMWF

#ESoWC2019

UNIVERSITY OF OXFORD

Interpret Models

SHAP Values

ECMWF

#ESoWC2019