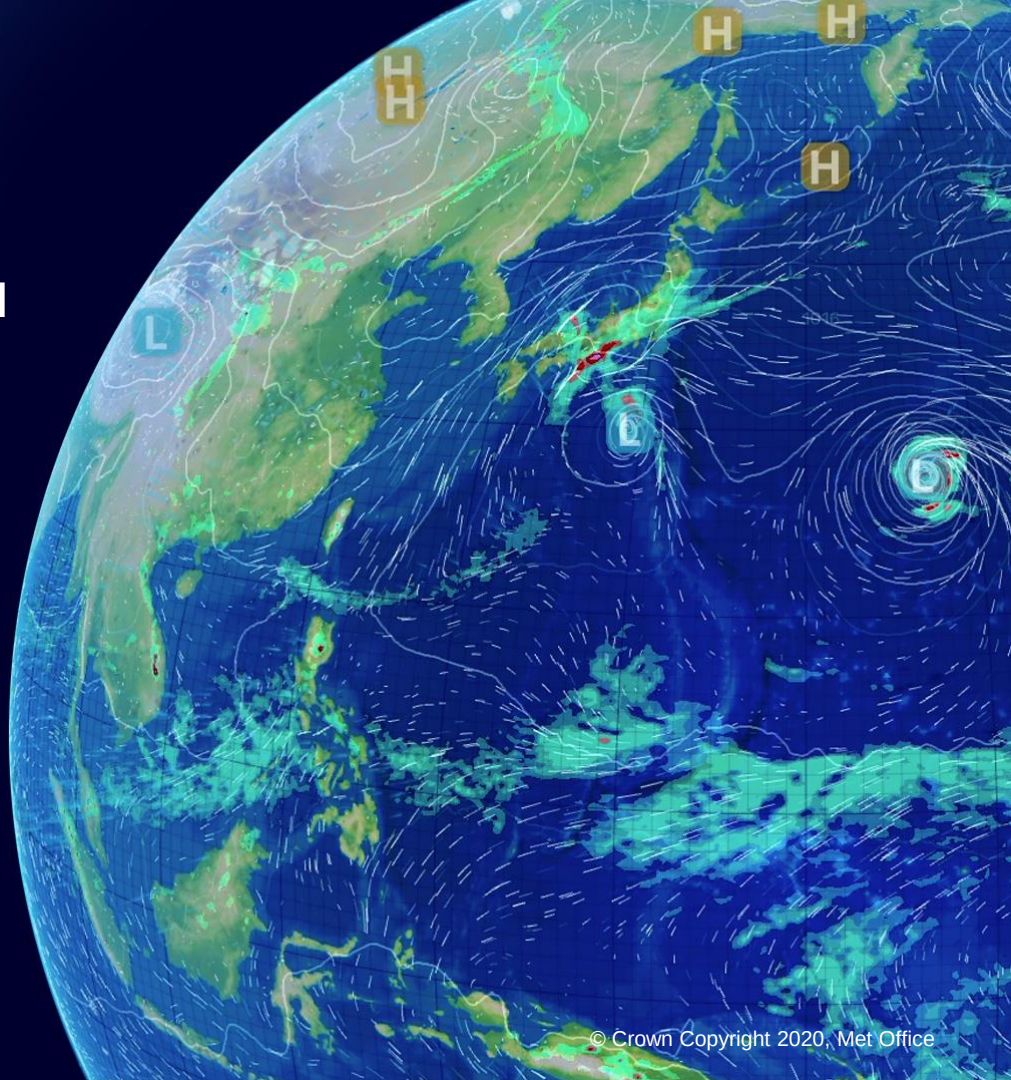
Bias correction of ocean wave model forecasts using Random Forests

Nefeli Makrygianni

SIG challenger meeting

September 2022



The importance of the ocean wave predictions

- Ocean waves can be a hazard for:
  - Economic sectors (renewable energy, shipping, coastal defences etc)
  - Population

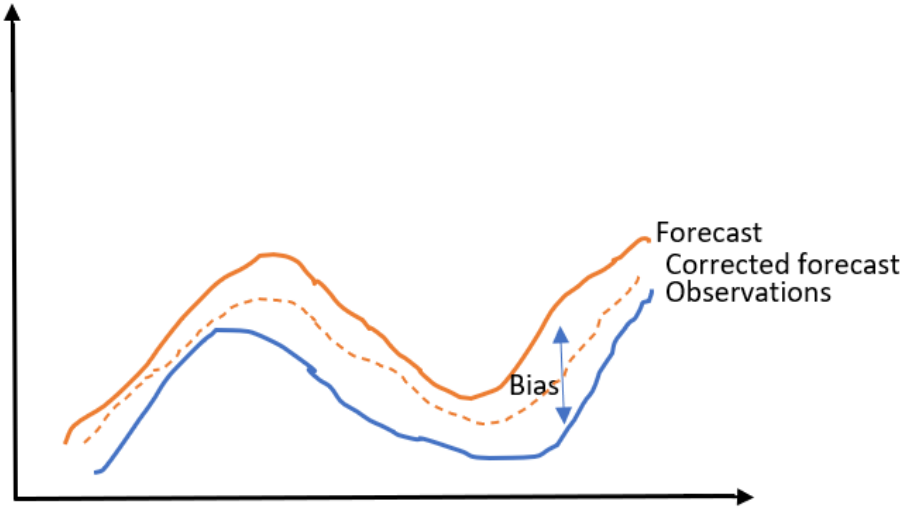
Wave forecasting: Issues and possible improvements

- Wave predictions rely on physics-based computational models which can hold many errors (initial conditions, non-linear processes etc)

- Statistical post-processing techniques do not replace, but can help on the improvement of the model outputs

Why Bias correction?

- Post processing technique, predicting the bias between ofc and obs and aims to correct the forecast
- Simple concept that has been found working well



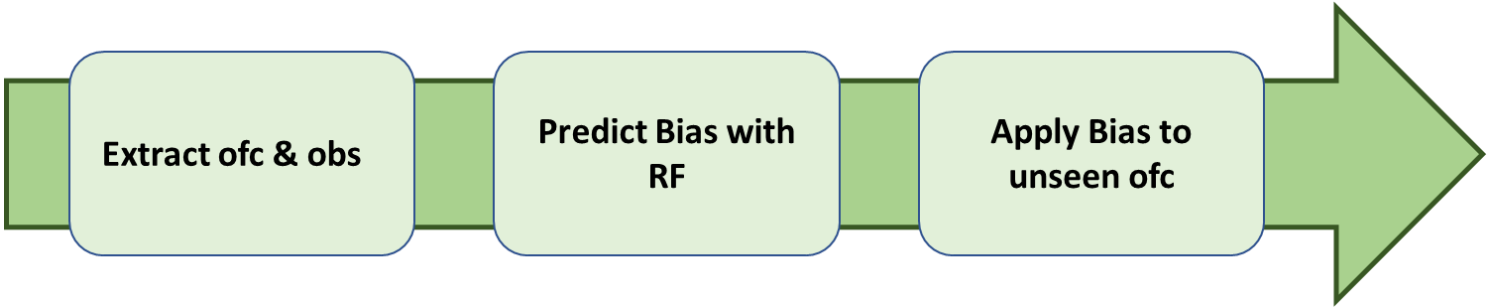
BIAS = forecast – observations

Machine Learning (ML) for bias correction

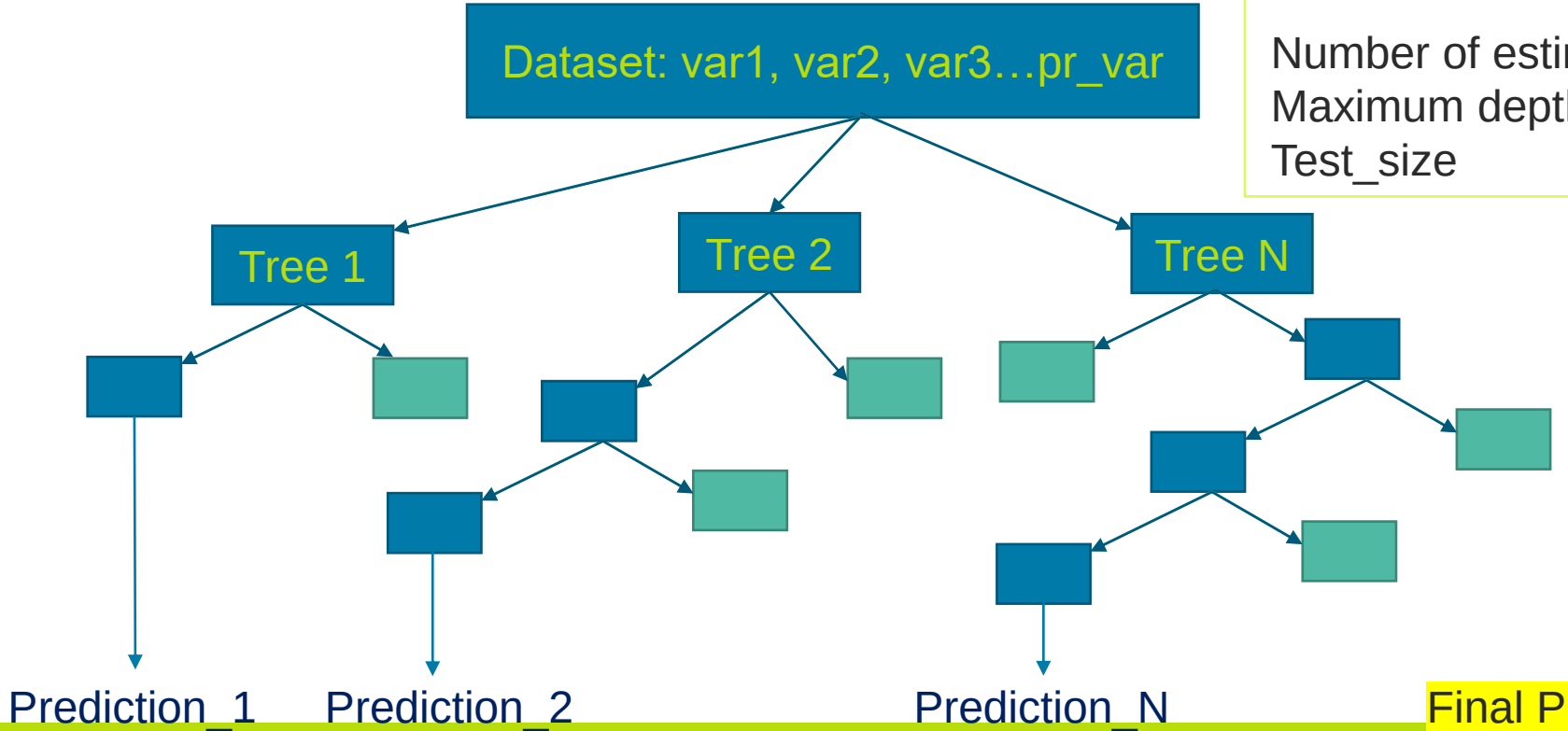
- For many years bias correction was performed manually
  - Time consuming
- Recently, machine learning techniques have been introduced as an automate process
- Random Forests (RF) technique has been found to work well for regression purposes (Chen et al., 2021)

Project Aim

- RF: initial test case
- Develop a RF regression model which predicts significant wave height bias based on observations.
- Correcting new/unseen forecasts from the model and test its success.



Methods: ML Random Forests regression



Number of estimators
Maximum depth
Test_size

Methods: ML Random Forests

- We here use
  - 100 number of estimators
  - Maximum depth 20
  - 80% for training 20% for testing
  - New unseen forecasts
- Tools:
  - Python provides Scikit Learn toolbox which includes RF

Data

- Available Data:
  - Operational forecast data from Met Office archive
  - Observations from offshore locations
- Variables:
  - significant wave height (Hs),
  - wind speed (U10),
  - wind direction (wdir),
  - wave direction (dir),
  - peak period (Tp),
  - wave directional spread (Spr)
  - Predicted variable: bias

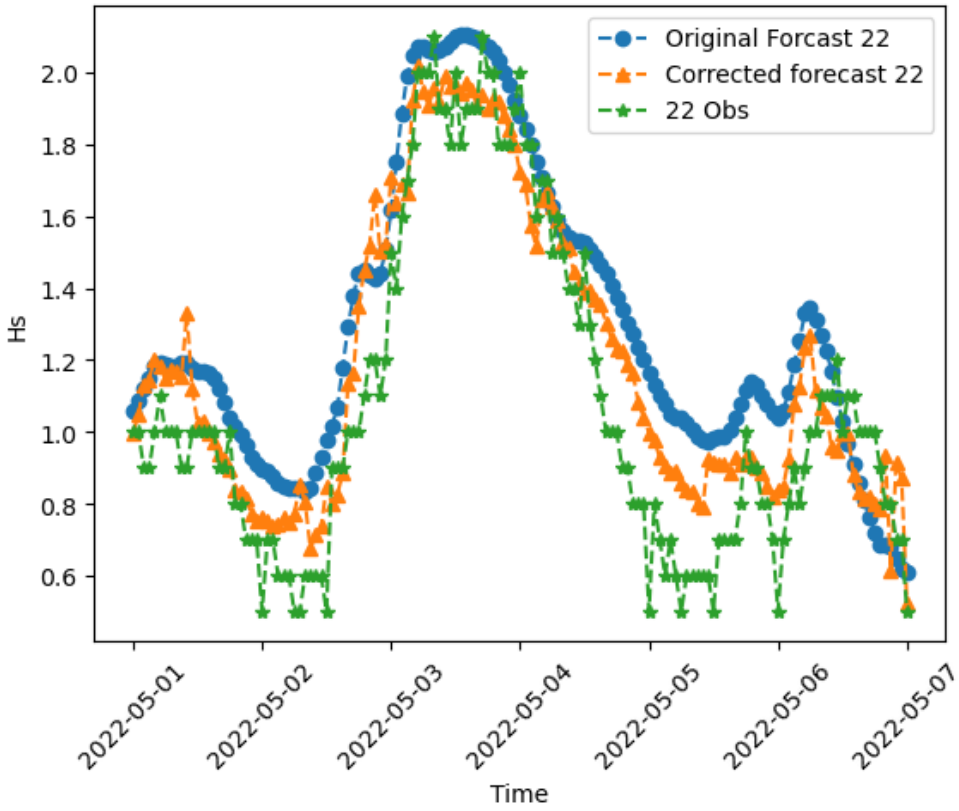
Data

- Forecasting period used:
  - 2020 - 2021 (train and test)
  - 2022 forecasts (validate)
- Pre-processed: 2D ML Matrix (dataframes in the right format for RF)

Preliminary Results

Single forecast example

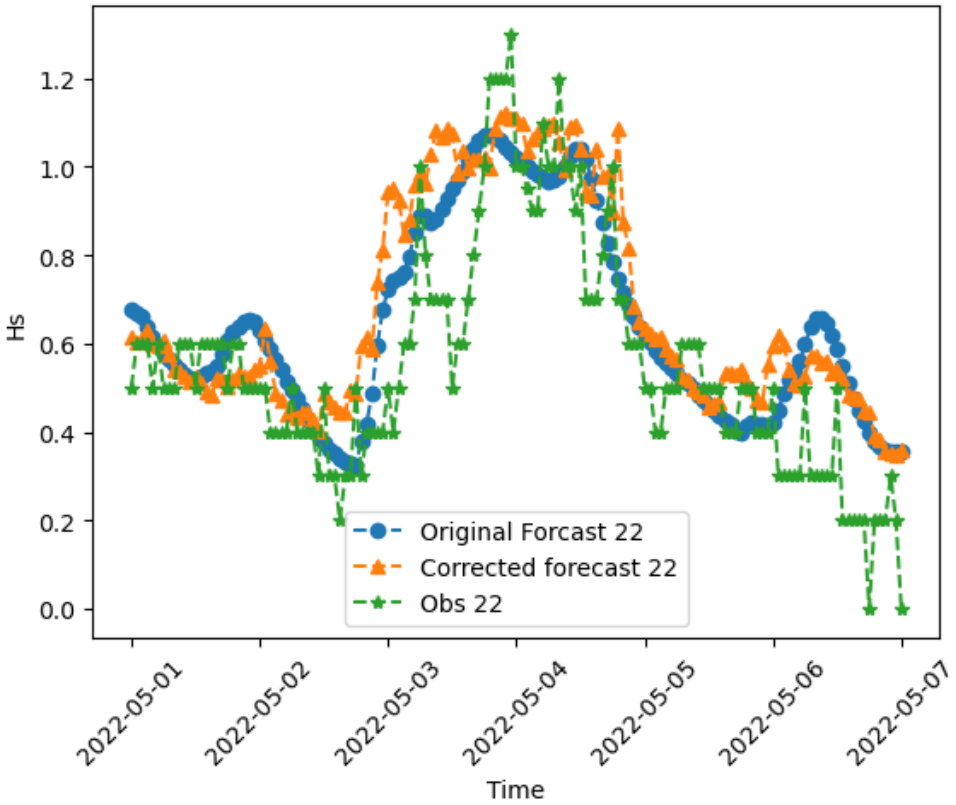
Northern NS



Preliminary Results

Single forecast example

Southern NS



Met Office

Summary and Future Work

- RF technique for bias correction, using MO ofc and NS obs
  - Promising initial results
  - Easy to use

- Very initial experiments
- Consider other machine learning techniques
- Longer datasets for training
- Apply bias correction technique for SA
- Long term validation statistics and more experiments
  - Single forecast used now means that different times will show different representatives