Reducing Errors Of Sea Level Height Forecasts Using Machine Learning

T. Xirouchaki, E. Steele, J. Amies, B. Gomez, C. O'Neill

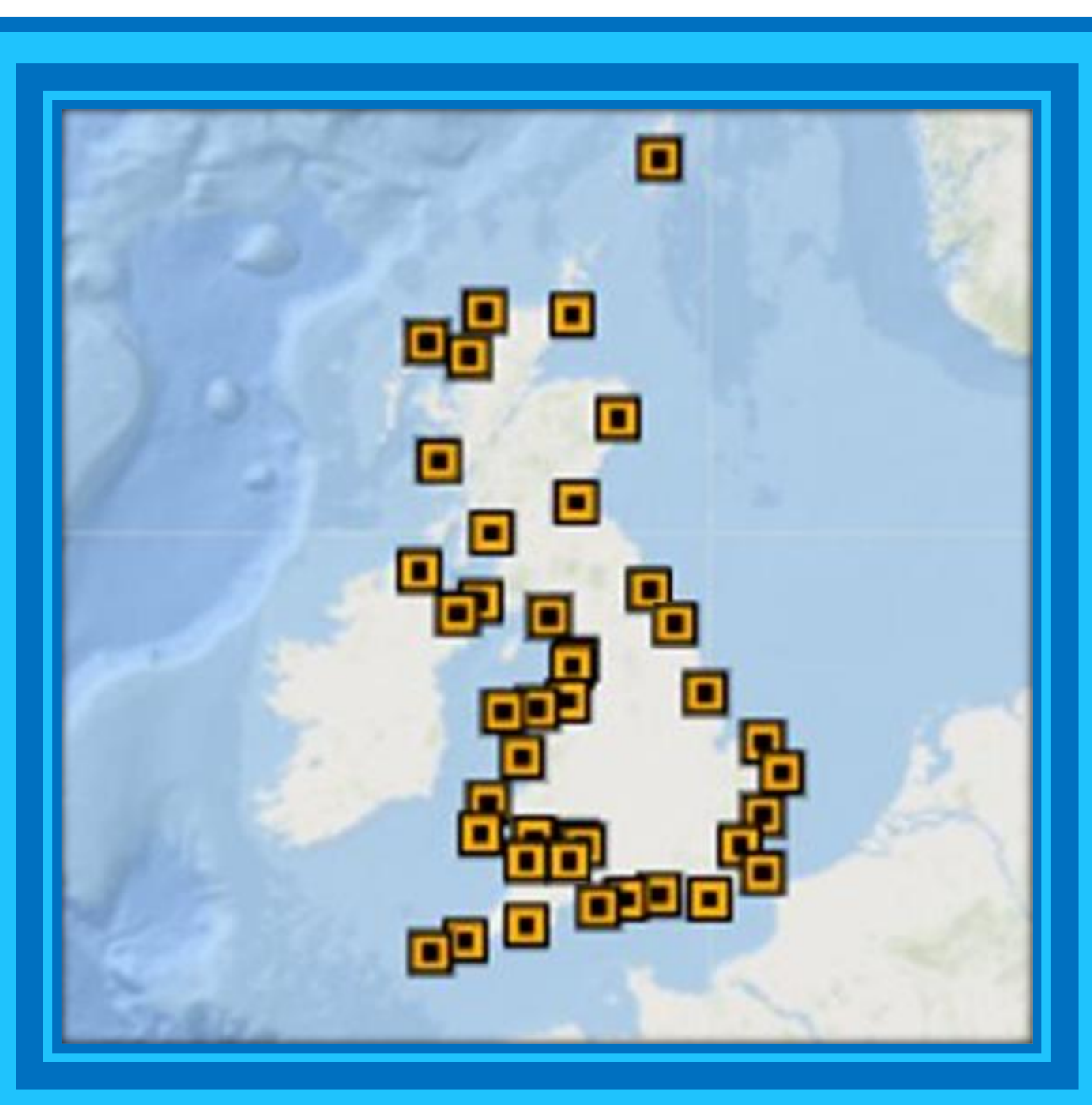

Fig. 1: Map of key locations the surge residual is forecast for in the UK

Context

Sea level is determined by two components, the **tide** and the **surge residual**. The **tide** is caused by forces on a planetary level, and can be estimated mathematically far in advance. The **surge residual** is a result of atmospheric conditions such as pressure and wind, and can only be forecast days to weeks ahead.

The Met Office produces sea level forecasts for ports across the UK. These are utilised by several stakeholders, such as the Environment Agency and the London Port Authority. They can be used to issue warnings, as well as manage anti-flooding measures like the **Thames Barrier**.

The aim of this project is to **reduce the predictive error** of the surge residual forecast.

Methodology – Architecture

The method used is **Gradient Boosted Decision Trees**.

A **Decision Tree** is a structure created by progressively splitting a dataset into smaller subgroups based on algorithmically determined criteria. When presented with a new data point, it is tested against the same criteria, allocated a subgroup, and the final prediction is determined by the values of the training data in that subgroup, also known as a leaf.

Gradient Boosting refers to gradually reducing the error with each iteration. For Decision Trees, that means the first tree aims to predict the physical model's error, the second tree aims to predict the first tree's error, and so on. In this project, up to **50 trees** were used per location.

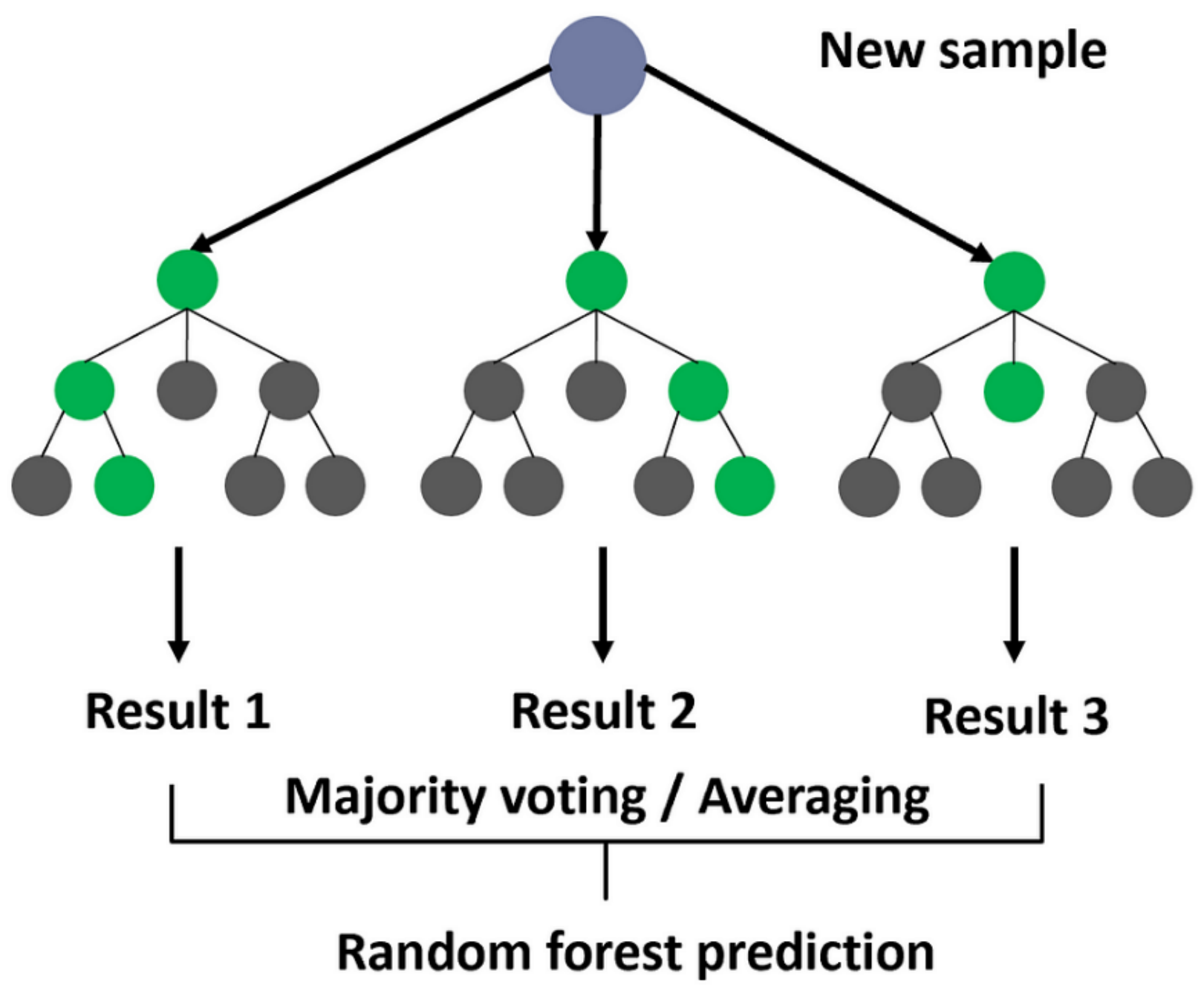

Fig. 2: A collection of Decision Trees forms a Random Forest [1]

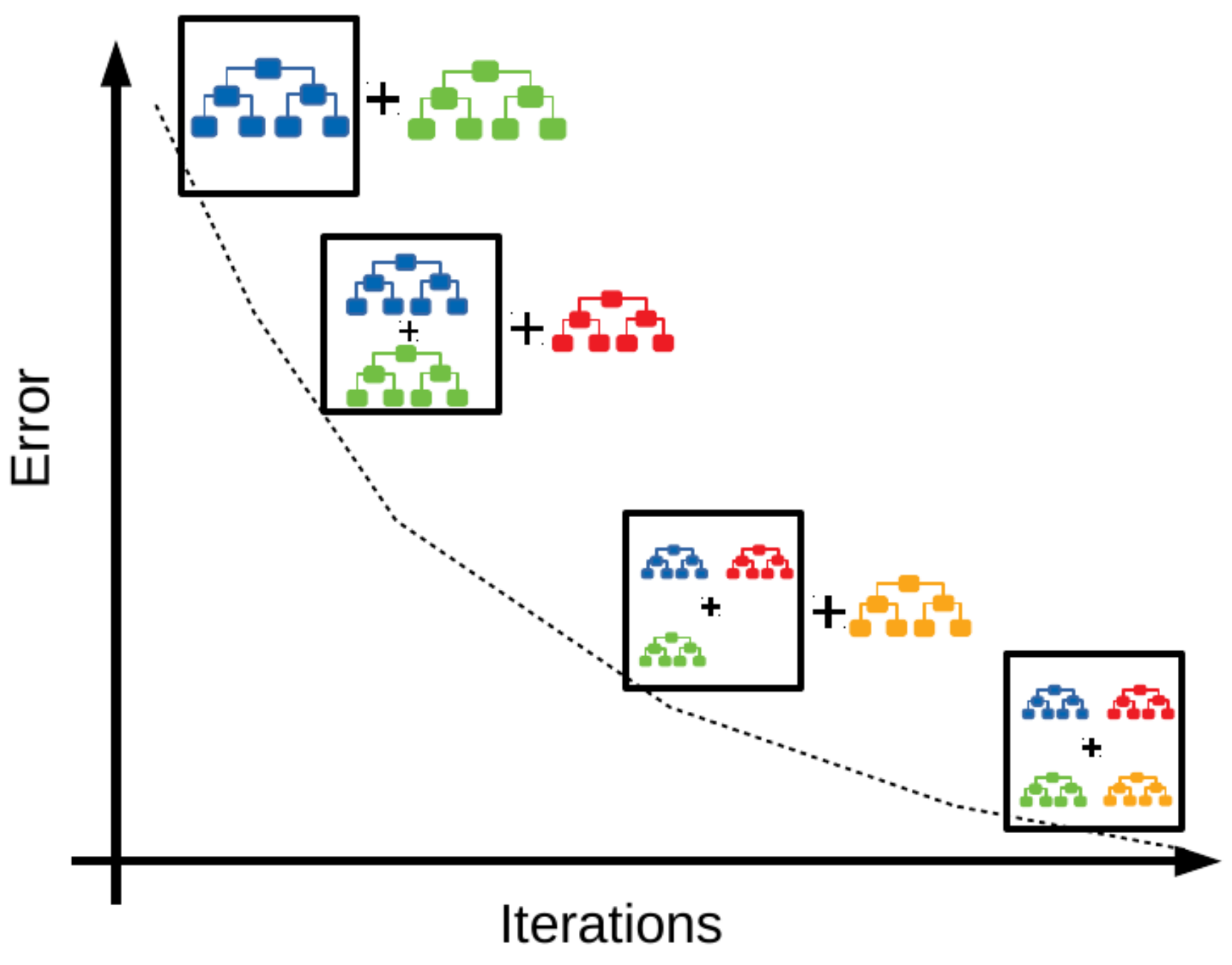

Fig. 3: Using Gradient Boosting to incrementally reduce the error [2]

Methodology - Data

Data from **August 2016** to **December 2020** is used to train the Machine Learning model, and data from **2021** is used to test its performance.

The physical model that forecasts the surge residual (NEMO) is an **ensemble model**, which means the model runs multiple times, to give a distribution of possible outcomes. To condense that into data the Machine Learning model can use, without losing the spread of the ensemble, we used the 10th, 50th, and 90th percentiles of the ensemble forecast.

Other data used as input includes:

- the harmonic tide
- time since last tidal peak
- wind
- sea level pressure
- moon phase
- prevailing weather regime

The error correction, also known as bias correction, is applied on the **mean of the ensemble** produced by the physical model, and accuracy is assessed by comparing to observational data from **tide gauges**.

Results

The Machine Learning model reduces the physical model's Mean Absolute Error by **20.5%** as well as reducing how often the error is more than 20cm by **37%**. As shown by Fig. 5, it also recentres the error distribution, such that the mean is reduced by **6.8 cm**, meaning the model is less likely to underestimate high surges.

As can be seen in Fig. 4, the model performs especially well in the south-east. Overall, only three locations in the UK show negative results.

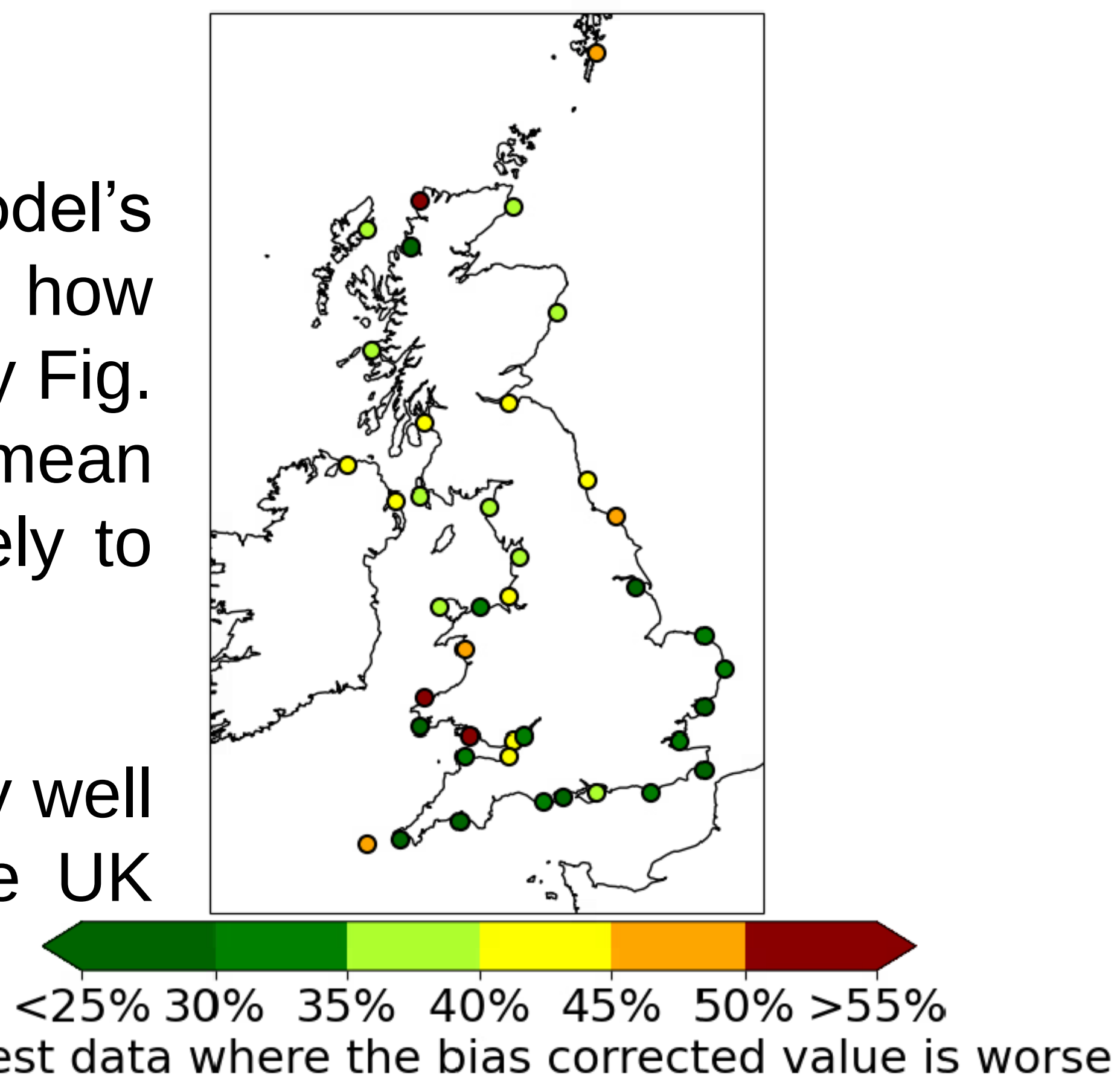

Fig. 4: Performance of the trained model across the UK

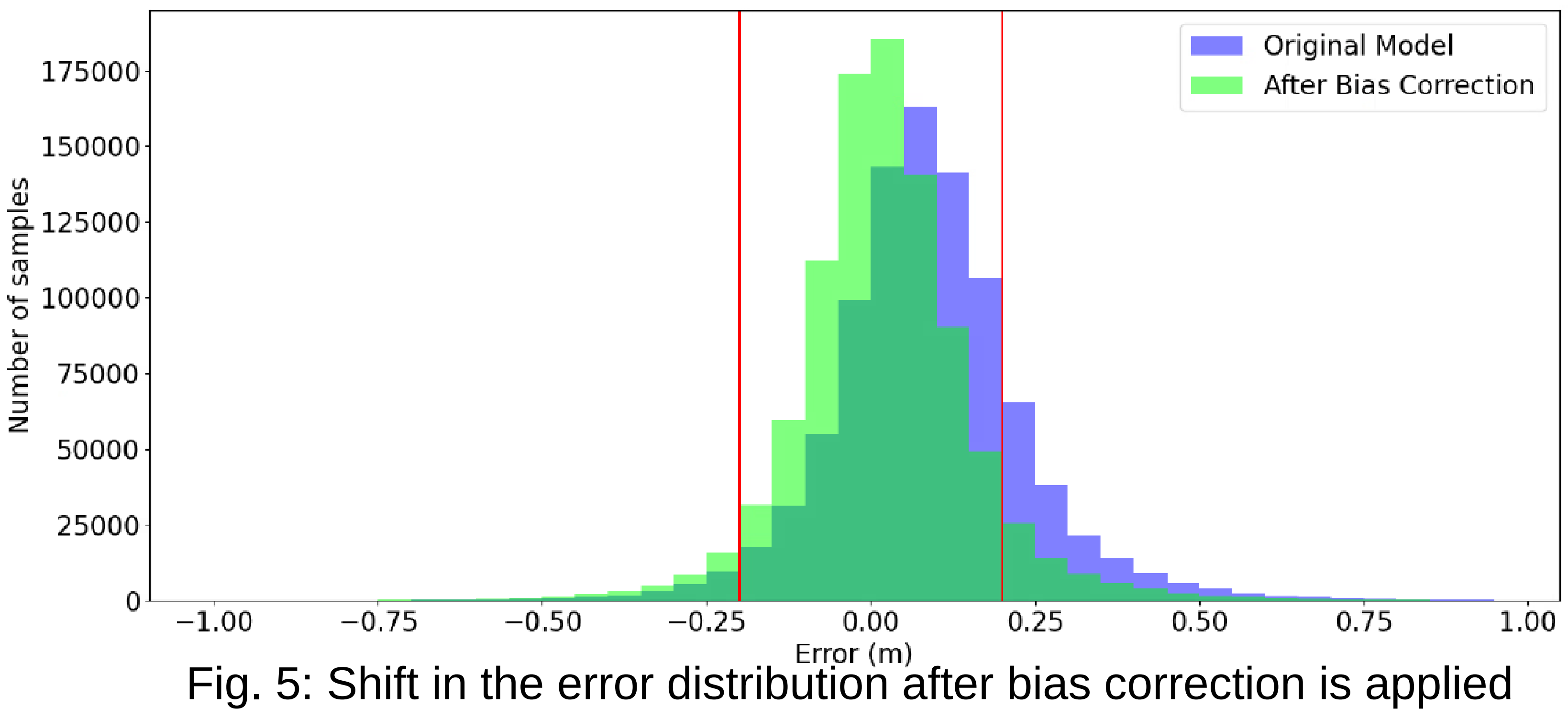

Fig. 5: Shift in the error distribution after bias correction is applied

Future Work

Next steps will include validating the performance of the model in **real time** and developing a method to apply it to the ensemble forecast without losing its probabilistic nature –such as applying the error correction to each **ensemble member,** while retaining their spread.

[1] Yehoshua, R. (2023) *Random forests*, *Medium*. Available at: https://medium.com/@roiyeho/random-forests-98892261dc49

[2] Pal, A. (2020) *Gradient boosting trees for classification: A beginner's guide*, *Medium*. Available at: https://medium.com/swlh/gradient-boosting-trees-for-classification-a-beginners-guide-596b594a14ea

Met Office FitzRoy Road, Exeter, Devon, EX1 3PB United Kingdom
Email: theano.xirouchaki@metoffice.gov.uk