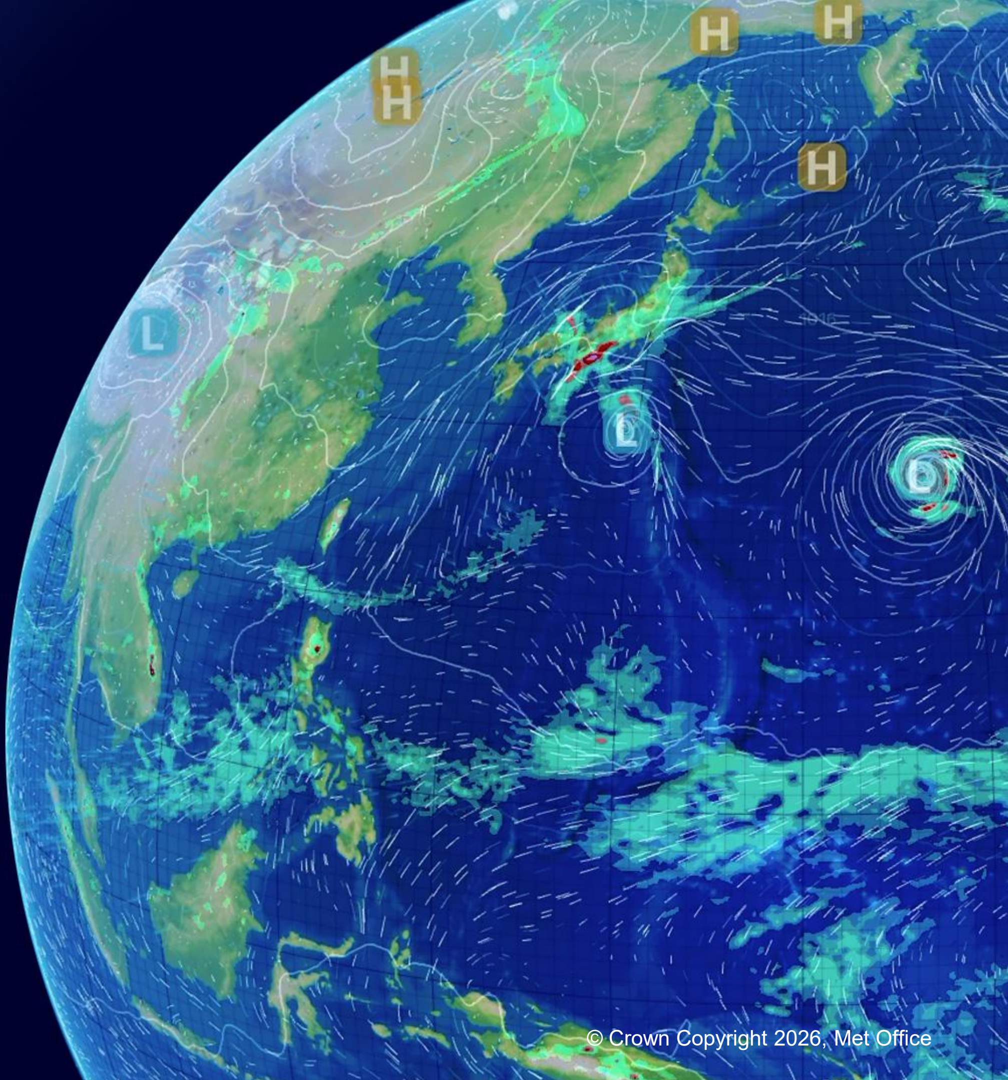
Met Office

Integrating 'Trustworthy AI' Principles into Machine Learning for Aviation Weather Forecasting

Thomas Chitson, Dr Piers Buchanan,
Dr James Fallon, James Shapland

Supporting material for poster

EGU26-7283



Motivation

Global aviation relies on accurate weather forecasting to safely operate hundreds of thousands of daily flights. Advances in weather forecasting using machine learning (ML) methods are leading to forecasting products that can achieve greater accuracy with reduced processing times and resource usage. However, ML methods must still meet stringent aviation safety and regulatory requirements before widespread industry adoption can be achieved.

The Aviation Applications research group at the Met Office has been exploring ways in which “Trustworthy AI” principles can be used in ML research to develop new products and services fit for the aviation industry.

Met Office

Met Office and the Civil Aviation Authority

The UK's aviation regulator, the Civil Aviation Authority (CAA), is responsible for ensuring weather forecasting provisions are fit for UK airspace requirements, allowing the safe operation of flights through the UK's skies.

The CAA promotes the development of AI technology in UK airspace (including weather forecasting provisions) but must ensure that innovation is safe, secure and promotes public trust, as laid out in the CAA AI Innovation Strategy (Artificial Intelligence | UK Civil Aviation Authority).

The principles of "Trustworthy AI" developed by the Aviation Applications group at the Met Office will enable ML aviation weather forecasting developments to meet the CAA's principles for implementation of AI in aviation.

Trustworthy AI

The Aviation Applications research group has developed a set of key considerations for building user-confidence in ML products, referred to as the principles of “Trustworthy AI”.

These principles include making sure ML methods are explainable and data used for training is free from bias but also extends to making sure new ML products and services are useful and can deliver benefit in a consistent and timely manner.

The next slide shows a summary of “Trustworthy AI” principles, and the following four slides explain how these principles apply to ML models used for aviation weather forecasting.

Met Office

Trustworthy AI

Data Assured

- Free from bias
- Good availability
- Density of Observations

Representative

- Follows laws of physics
- Captures extreme events
- Uncertainty quantification

Explainable

- Trusted by end-users
- Tested under different situations
- Interpretable methods

Operational

- Research to operations
- Clear data pipelines
- Useful, usable and used

Trustworthy AI: Data Assured

Data assured models use training and validation data that has good spatial and temporal availability, is free from bias, and has a good density of observations. In aviation weather forecasting this is important because:

- Some hazards (fog, turbulence, icing) are rare. Training data must ensure these events are captured.
- Aviation observations are sparse. Often there is only limited data available (e.g. spatially over flight routes, or temporally at airport opening hours). Models trained using these datasets are only representative of the coverage used.

Trustworthy AI: Explainable

Explainable AI refers to models that use methods to make their processes and outputs interpretable to users. This means their behaviour is well understood and can be trusted by end-users, as well as being reliable and predictable. This is important in aviation weather forecasting because:

- Aviation is a safety and regulation driven industry. If an incorrect weather prediction is made that could lead to a near-miss or accident, the model must be well understood to explain incorrect forecasts (this doesn’t mean models must be perfect!).
- End-users, such as airlines, airports and meteorologists, must be confident to be able to use ML methods to make decisions.

Trustworthy AI: Representative

ML models must be representative of the physical processes they are trying to predict. In aviation, it is vital that the predictions being made by models are physically possible and capture extreme events, as well as quantifying uncertainty. This is important in aviation weather forecasting because:

- Aviation weather forecasting services are used operationally to make decisions impacting aircraft safety. They must represent the physical world accurately.
- Aviation weather phenomena that leads to near-misses or accidents are often extreme events (convective cloud or severe turbulence). Models must capture extreme events well to be useful for hazard forecasting.
- Meteorologists make decisions based on probabilities of hazards occurring so ML models must quantify their certainty in decisions.

Trustworthy AI: Operational

The Met Office produces scientific products that are used operationally. End-users rely on products that are on-time and have clear data-pipelines. This is important in aviation weather forecasting because:

- Aviation weather forecasting is a global service requiring updates around the clock. ML models must be able to use the latest input data and deliver predictions to end-users quickly and reliably.
- Models must be useful, usable and used. ML models offer great results in certain scenarios (such as in nowcasting) but research must have defined benefits for end-users.

Met Office

Aviation Applications ML Research Projects

Cumulonimbus (Cb) Cloud Detection – Data and Method

- CNN detection model of Cb clouds near to airfields.
- Trained on data for 3 airfields but can be applied anywhere with suitable input data.
- Developed as an alternative to a radar and lightning detector system which suffers from high false-alarm rates.

Inputs	Outputs
Circular radar image centred around airfield.	Probability of Cb cloud
CAPE forecast ‘image’	Probability of towering-cumulus cloud (TCu)
METAR report (for truth labelling)	

Cumulonimbus (Cb) Cloud Detection – Key Results

- CNN method significantly reduces the number of false alarms.
- Operationally, requires similar input data and can make fast predictions.

Method	Hits	False Alarms
Conventional	257	1614
CNN radar only	259	1216
CNN radar & CAPE	259	858

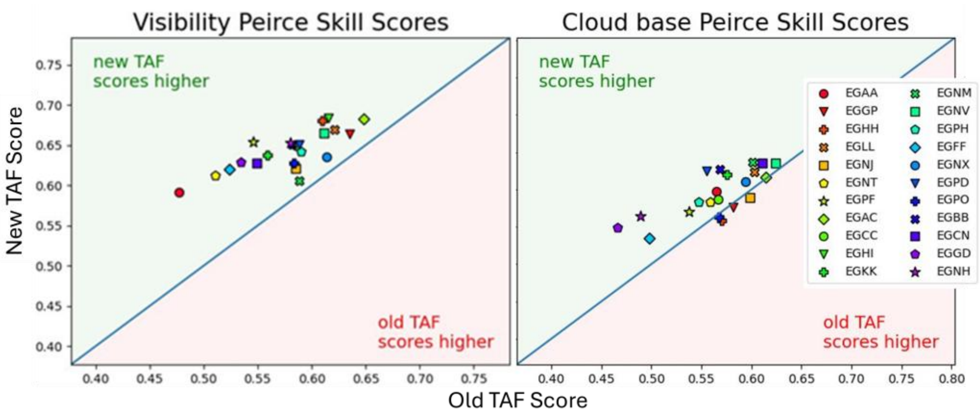
AutoTAF Nudging – Data and Method

- Met Office uses post-processing methods to convert model data into automatic terminal aerodrome forecasts (AutoTAFs). These are text-based forecasts for airfields, used globally.
- AutoTAFs are susceptible to expiring due to being over-specific (going "bust").
- XGBoost and Random Forest classifiers used to predict how AutoTAFs will go bust, then correct the input data to generate AutoTAFs less susceptible to going bust.
- Uses 21 meteorological features from IMPROVER forecast and trains on visibility and cloud observations from METARs.

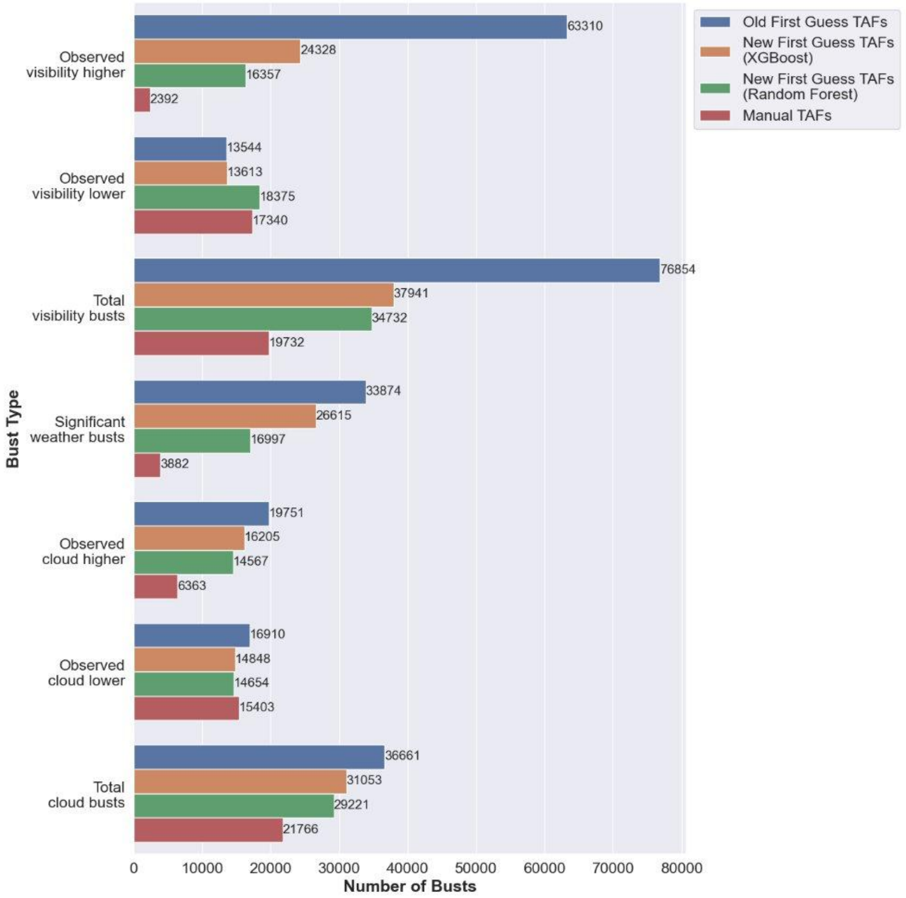
AutoTAF Nudging – Method

- Generate conventional AutoTAFs using unadjusted IMPROVER data (**Old TAFs**).
- Predict bust labels (e.g. higher than observed visibility, lower than observed visibility, no bust…)
- Use ML predictions to adjust IMPORVER input data (repeated 5 times to allow for multiple category changes).
- Generate new AutoTAF using adjusted IMPROVER data (**New TAFs**).
- Compare to observations (METARs).

AutoTAF Nudging – Key Results



- ML methods reduce the number of busts across cloud and visibility prediction.
- Achieve better skill scores for prediction (added bonus!).

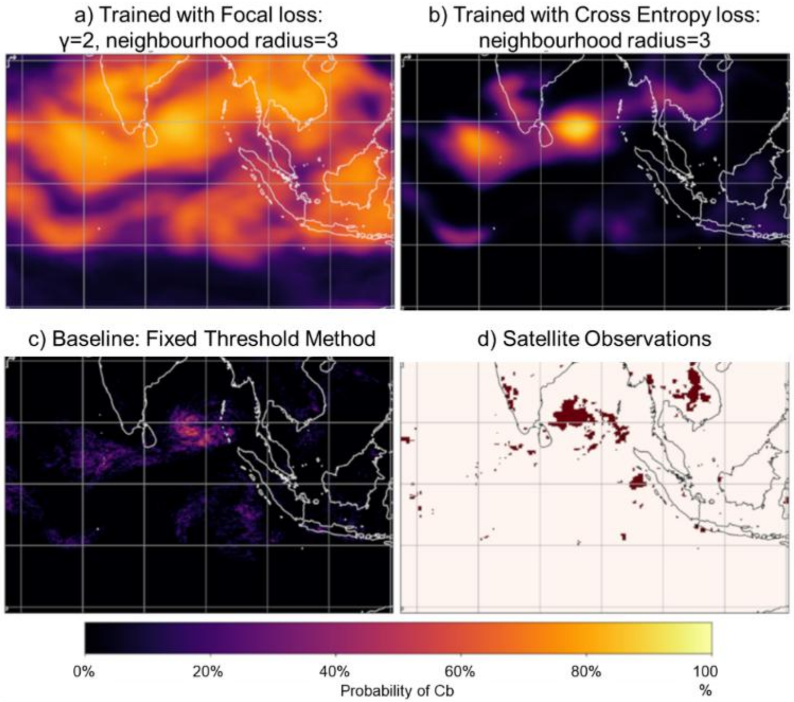


Global Hazardous Cloud Forecasting– Data and Method

- Current operational Met Office system only uses thresholds for MUCAPE and rainfall rate to identify Cb cloud risk in NWP forecasts.
- Train a CNN to learn what Cb clouds “look like” in NWP data (semantic image segmentation problem).
- Tested 4 loss-functions, 2 model architectures (U-Net and FCN) and 3 neighbourhood sizes = 24 model configurations.

Inputs	Outputs
16 MOGREPS-G parameters (3 months of training data, MJJ 2021)	Gridded Cb probability images
Satellite Cb Observations as truth	

Global Hazardous Cloud Forecasting– Key Results



- Model able to represent global convective cloud.
- 18 out of 24 model configurations showed skill in prediction.
- Excellent proof-of-concept showing image segmentation method can be applied to convective forecasting.

Case studies: Application of “Trustworthy AI”

Project	Trustworthy AI Methods
Convective Cloud Detection	• Filters out scenarios in which convective cloud not possible. • Built to capture rare events. Custom loss function rewards correct detection. • Outputs probability of detection.
AutoTAF Nudging	• Built on established and trusted operational system. • ML outputs are checked by forecaster before being issued. • Random forest method is conceptually easier to follow than other models.
Global Convective Forecasting	• Builds on Met Office expertise in utilising global model data. • Fills gaps in convective nowcasting capability.

Conclusions

- The Aviation Applications research group at the Met Office has **demonstrated ways to build trust in ML products for aviation** through several research projects.

- To operationalise forecasting products, the group has **worked with the CAA to consider aviation regulation requirements**. These align with the principals of **Trustworthy AI** that have been used by the group allowing future UK airspace forecasting requirements to be met with the benefits of ML methods.

Met Office

More questions?

For more information please contact

www.metoffice.gov.uk

thomas.chitson@metoffice.gov.uk

0370 900 0100