13.4 WIND PREDICTION AT AIRPORT DURING TROPICAL CYCLONES USING XGBOOST MACHINE LEARNING MODEL

Yin-lam NG[1,*], Mang-hin KOK[1], On-pong CHEUNG[1], Cheuk-fung FONG[2]

[1] *Hong Kong Observatory, Hong Kong, China*
[2] *Chinese University of Hong Kong*

1. INTRODUCTION

Chan (2023) studied historical data of the China Meteorological Administration (CMA) for tropical cyclones (TCs) making landfall in China during 1949–2019 and found that average number of TCs making landfall along the South China coast (Guangdong, Guangxi and Hainan provinces) is about 5. As Hong Kong lies within this popular landfall region, many studies looked into low level wind shear and turbulence at the Hong Kong International Airport (HKIA) caused by TCs (Wong and Tse 2016; Chan and Hon 2016; Chan 2019). In additional to these high impact phenomena, wind forecasting itself at the HKIA during TCs is a critical task for ensuring aviation safety and operational efficiency. To reduce risks associated with high winds during TCs, precise and timely forecasts are required to manage air traffic flow, ground operations, and resource allocation effectively. However, traditional forecasting methods, such as climatology, statistical regression, and analogue techniques, often lack the accuracy and reliability needed to address rapidly changing conditions during TCs.

To address these challenges and noting machine learning model could improve overall accuracy of TC intensity (Chan et al. 2021), this study applies eXtreme Gradient Boosting (XGBoost) (Chen and Guestrin 2016), a machine learning model known for its scalability and accuracy, to predict wind speed and direction at HKIA during TCs. The model integrates diverse datasets, including real-time observations, TC track and intensity forecasts, and post-processed numerical weather prediction (NWP) outputs such as the Objective Consensus Forecast (OCF) (Chan and Lee 2013). This study aims to evaluate the performance of the XGBoost model and its ability to enhance decision-making processes for aviation operations during TCs.

2. DATA AND METHODOLOGY

2.1. Data Sources

The XGBoost wind prediction models were trained using comprehensive datasets:

- Observation Data: aviation routine weather report issued at hourly intervals (METAR) during 1998-2020.
- TC Forecasts: Tropical cyclone track and intensity forecasts issued by the Hong Kong Observatory. Best tracks have been used for 1998-2020 and data included TCs within 500 km radius of HKIA.
- NWP Model Outputs: Post-processed multi-model consensus forecast (OCF) of winds at HKIA, which integrates predictions from various global and regional NWP models, during 2013-2020.

These corresponding data in 2021 and 2022 have been used as testing dataset.

2.2. Data Pre-processing

Key pre-processing techniques were applied to ensure the quality and usability of the data:

- Trigonometric Transformations: Wind direction was transformed into sine and cosine components to account for its periodic nature and avoid discontinuities at $0°/360°$.
- Feature Engineering: Additional variables, such as TC proximity to HKIA, changes in TC intensity, and local pressure gradients, were derived to capture relevant meteorological factors.
- Handling Missing Data: Missing values in observational and forecast datasets were imputed using interpolation, nearest-neighbour techniques or the XGBoost built-in branch directions learnt during model training.

2.3. Model Development

The XGBoost model was trained separately for three forecast ranges:

- Immediate (T+0 hour): Real-time wind predictions based on current observations.
- Short-term (T+1 to T+12 or T+24 hours): Predictions incorporating short-term METAR, TC forecasts and OCF outputs.
- Extended (T+1 to T+36 hours): Longer-range predictions leveraging OCF outputs.

Hyperparameter tuning was performed using a grid search approach provided by the python package, Optuna (Akiba et al. 2019), to optimise the model's performance. The loss function was set to minimise Mean Absolute Error (MAE).

[*] *Corresponding author address:* Danice YL Ng, Hong Kong Observatory, 134A Nathan Road, Kowloon, Hong Kong, China.;
E-mail: ylng@hko.gov.hk.

3. RESULTS AND PERFORMANCE EVALUATION

3.1. Metrics for Evaluation

The performance of the XGBoost models was evaluated using the following metrics:

- Mean Square Error (MSE): Measures the average squared difference between predicted and actual wind speed.
- Mean Absolute Error (MAE): Quantifies the average magnitude of wind speed forecast errors.
- R-squared (R^2): Assesses the proportion of wind speed variance explained by the model.
- Root Mean Squared Error (RMSE): Measures the average squared difference between predicted and actual wind direction.

These metrics were compared with estimations solely based on climatology and OCF outputs, which are the main guidance that forecasters at HKO would currently make reference to. Estimations for wind speed and direction solely based on climatology, given as minimum, mean and maximum, and resepective TC locations and intensities are updated on a yearly basis for the Hong Kong Flight Information Region. The grid resolution to overlay the TC forecast track for climate estimations is $1^\circ \times 1^\circ$.

A quick study has been done to compare XGBoost with three other machine learning models that are available at scikit-learn package (Pedregosa et al. 2011) for the immediate (T+0 hour) wind speed predictions. Three latest available OCF outputs, denoted as "OCF with 12-23 hour lead-time", "OCF with 24-35 hour lead-time", and "OCF with 36-47 hour lead-time" were included as machine learning model features for models to pick up any forecast trend changes. OCF with < 12 hours were not included due to spin-up time.

The following table, Table 1, showed the MSE, MAE and R^2 for ElasticNet Regression, XGBoost, Decision Tree and Random Forest, against climate estimations and OCF outputs with different lead times.

	Speed MSE	Speed MAE	Speed R^2	Direction RSME
XGBoost	19.02	3.41	0.48	31.64
ElasticNet	20.35	3.56	0.45	32.49
Decision Tree	35.22	4.644	0.041	58.52
Random Forest	19.72	3.49	0.46	34.50
	Speed MSE	Speed MAE	Speed R^2	Direction RSME
climate estimations (mean)	30.31	4.33	0.18	33.142
climate estimations (min)	99.75	8.40	−1.70	69.06
climate estimations (max)	187.46	10.02	−4.07	80.77
	Speed MSE	Speed MAE	Speed R^2	Direction RSME
OCF with 12-23 hour lead-time	21.04	3.53	0.43	37.87
OCF with 24-35 hour lead-time	23.54	3.73	0.36	47.48
OCF with 36-47 hour lead-time	28.23	4.05	0.24	51.83

TABLE 1. Model performances based on testing dataset for wind speed and direction forecast at the HKIA.

Among the four machine learning models, XGBoost has the lowest speed MSE, speed MAE and direction RSME. Therefore, XGBoost has been chosen to be used in the study. At this stage, no tuning of the hyperparameters have been made yet. In the following sections, XGBoost forecast has been improved by fine-tuning hyperparameter using the Optuna package.

3.2. Model Performance

The fine-tuned XGBoost models outperformed baseline models, including the OCF, in all forecast ranges.

1) IMMEDIATE RANGE (T+0 HOUR)

The fine-tuned XGBoost model achieved the lowest MSE and MAE, providing more accurate real-time wind predictions.

The information of wind structure given as radii (maximum or in 4 quadrants) of strong ($41-62$ km/h), gale ($63-87$ km/h), storm ($88-117$ km/h), hurricane (≥ 118 km/h) force winds improved MSE, MAE and R^2 only slightly (Table 2). Therefore, wind structure was not included in the model.

	Speed MSE	Speed MAE	Speed R^2	Direction RSME
No wind radii	17.48	3.24	0.52	32.82
Wind radii in 4 quadrants	17.28	3.22	0.53	34.63
Wind radii (maximum)	17.41	3.24	0.04	33.33

TABLE 2. Model performances based on testing dataset for wind speed and direction forecast at the HKIA.

Another experiment was done to see how much TC forecast data around the real-time wind predictions would be optimal. Past and future movement for the 3/6/9/12/15/18/24-hour TC position and intensity have been included as model features to reflect TC's movement and the performance of models with 3/6/9/12-hour TC forecast (positions and intensities) as model features is given in Figure 1. Having 6/9-hour TC forecast as model features provided the best evaluation metrics.

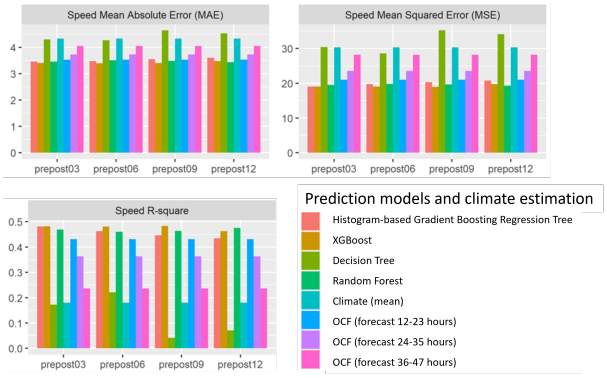


FIG. 1. MSE, MAE and R^2 of AI models with 3/6/9/12-hour TC forecast as model features.

The mean absolute value of SHAP (SHapley Additive exPlanations) (Lundberg 2017) for the best preforming XGBoost model for real-time wind predictions is given in Figure 2. In additional to OCF wind speed forecast, TC track information and METAR's QNH (mean sea level pressure measured at airport with adjustment made to suit aeronautical use) have the stronger influence on the wind speed prediction.

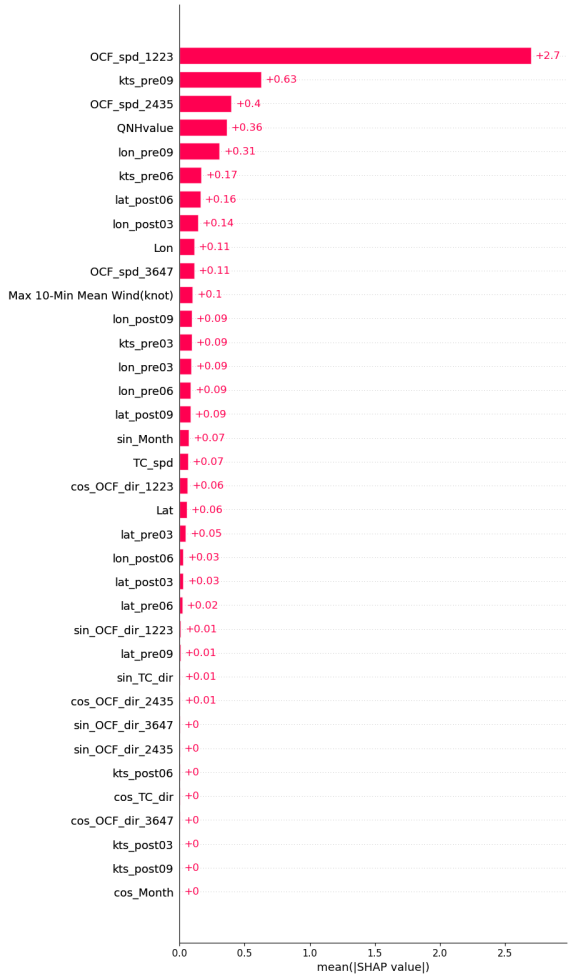


FIG. 2. The mean absolute value of SHAP for the best preforming XGBoost model for real-time wind predictions.

An example of XGBoost model for real-time wind predictions is given in Figure 3.

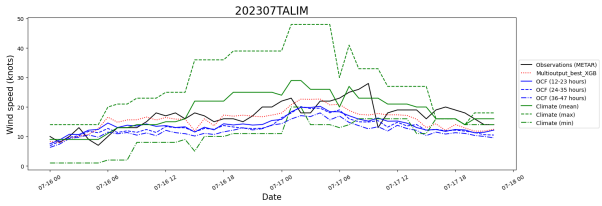


FIG. 3. Wind forecasts and observation for TC Talim in 2023. Green lines give the climate estimates (max, mean, min) and blue lines are OCF with different forecast hours. Red dotted line is the forecast by the best performing XGBoost model which is closer to the black line that indicates observed wind speed.

An XGBoost model has been built to forecast the probability of gusts. The "G" in METAR code signifies that there are wind gusts. TC position with gusts reports in 2013-2022 is given in Figure 4.

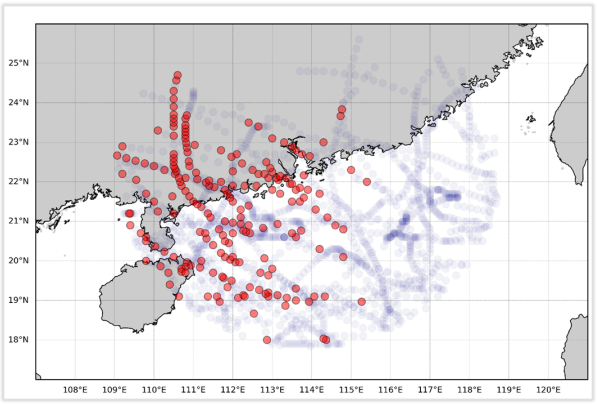


FIG. 4. The red dots indicated TC position that associated with gusts reports and blue dots are TC locations with no gusts observed in 2013-2022 at the HKIA.

Best threshold for gust probability given by XGBoost model is chosen by balancing false positives and false negatives in an AUC-ROC curve (Figure 5). The AUC score is 0.87 with best threshold being 0.07.

Using a threshold of 0.07, the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) values are 0.79, 0.61 and 0.35 respectively. The contingency table is given in Table 3.

An example of XGBoost model for gust probability predictions is given in Figure 6.

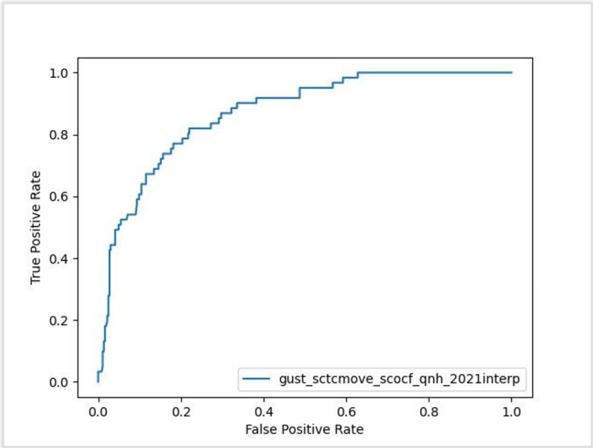


FIG. 5. AUC-ROC curve for gust probability given by XGBoost.

	Observations (METAR)	
XGBoost model forecast	No gusts observed	Gusts reported
No gusts forecast	287	13
Gusts forecast	76	48

TABLE 3. Contingency table for gust and no gust classification.

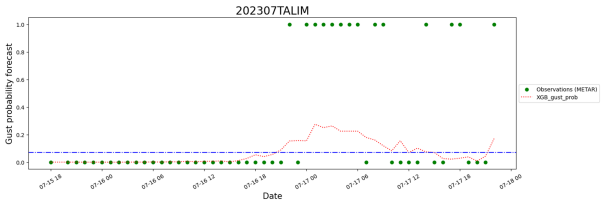


FIG. 6. Gust probability forecasts and observation for TC Talim in 2023. Green dots were the hourly METAR observations with 0: no gust and 1: gust observed. XGBoost probability is plotted as red dotted line and the blue horizontal line marked the best threshold of 0.07.

2) SHORT-TERM RANGE (T+1 TO T+12 HOURS)

XGBoost demonstrated stronger predictive skill over the latest three OCF outputs. In additional to other model features, the XGBoost model post-processed the OCF outputs and provided better forecast. Model evaluation metrics is given in Table 4.

	MSE	MAE	R^2
XGBoost	15.957	2.986	0.578
OCF with 12-23 hour lead-time	22.72	3.63	0.40
OCF with 24-35 hour lead-time	26.38	3.83	0.34
OCF with 36-47 hour lead-time	31.94	3.96	0.27

TABLE 4. Model performances based on testing dataset for wind speed T+1 to T+12 hours forecast at the HKIA.

3) EXTENDED RANGE (T+1 TO T+36 HOURS)

While the model's accuracy decreased slightly at longer lead times, it still outperformed baseline models. It was found that using three XGBoost models for T+1 to T+12 hours, T+13 to T+24 hours and T+25 to T+36 hours forecasts respectively could provide more accurate forecast than using one XGBoost model for T+1 to T+36 hours forecast.

	MSE	MAE	R^2
XGBoost (T+1 to T+12 hours)	15.96	2.99	0.58
XGBoost (T+13 to T+24 hours)	20.01	3.29	0.50
XGBoost (T+25 to T+36 hours)	30.23	3.82	0.31
XGBoost (T+1 to T+36 hours)	MSE	MAE	R^2
XGBoost as 3 separate models	23.80	3.51	0.44
XGBoost as one model	24.55	3.51	0.43
OCF	MSE	MAE	R^2
OCF (T+1 to T+12 hours)	22.72	3.63	0.40
OCF (T+13 to T+24 hours)	26.38	3.83	0.34
OCF (T+25 to T+36 hours)	31.94	3.96	0.27
OCF (T+1 to T+24 hours)	24.84	3.77	0.36
OCF (T+1 to T+36 hours)	28.74	3.97	0.33

TABLE 5. Model performances based on testing dataset for wind speed T+1 to T+36 hours forecast at the HKIA.

With hourly observation available at HKIA, XGBoost model for extended range wind predictions could be updated at an hourly basis. An example of XGBoost model for extended range wind predictions is given in Figure 7. Although the highest wind speed observed was at 09Z on 17 July 2023, XGBoost models provided a heads-up for winds could be close to 25 knots at the HKIA, while OCF generally underestimated the wind speed. Then at a later basetime as shown in Figure 8, XGBoost model gave an improved estimate of wind speed especially for T+1 to T+12 hours and OCF still underestimated the wind speed.

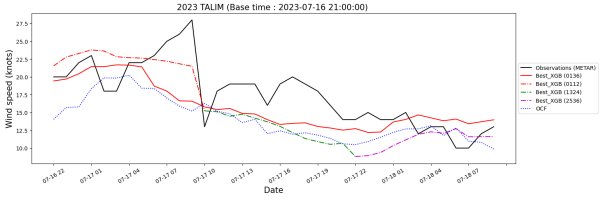


FIG. 7. Wind forecasts and observation for TC Talim with basetime 21Z on 16 July 2023. The black line indicates observed wind speed. Red, green and purple gave the XGBoost forecasts for T+1 to T+12 hours, T+13 to T+24 hours and T+25 to T+36 hours forecasts respectively. The blue dotted line is the OCF forecast.

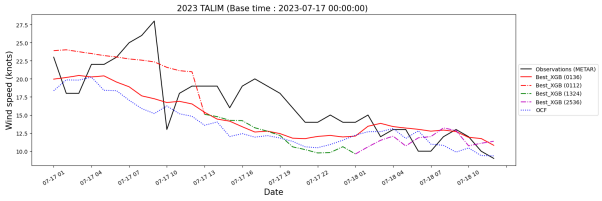


FIG. 8. Wind forecasts and observation for TC Talim with basetime 00Z on 17 July 2023. The black line indicates observed wind speed. Red, green and purple gave the XGBoost forecasts for T+1 to T+12 hours, T+13 to T+24 hours and T+25 to T+36 hours forecasts respectively. The blue dotted line is the OCF forecast.

4. OPERATIONAL BENEFITS

Another XGBoost model was trained with OCF not included as model features. This would allow forecasters to make amendment to the TC forecast (track and/or intensity) and this XGBoost model could provide forecast guidance based on the alternative track. This XGBoost model could also be applied to the Ensemble Prediction System (EPS) of the ECMWF, which comprises one control forecast (CNTL) plus 50 forecasts each with slightly altered initial conditions and slightly altered model physics. The spread of wind speed forecasts could allow forecasters to assess model uncertainty. An interactive webpage has been developed for trial by aviation forecasters at the Hong Kong Observatory.

4.1. Case Study: Typhoon Saola (2023)

TC Saola tracked close to Hong Kong in 2023 summer time. Aviation forecasters could change the track and intensity at the interactive webpage in Figure 9 to assess wind changes under different scenarios.

After aviation forecasters have made amendment to the TC forecast, calculation using XGBoost model would be displayed as in Figure 10. XGBoost model predictions are given as dark blue lines for the shown TC track, while green shaded area gives the spread of wind speed forecast based on ECEPS and green line for the mean of the ECEPS forecast. OCF forecast is also displayed for easier reference. As this is a case study, the actual observations were included as grey lines. Crosswinds forecast and gust probability have been provided as crosswinds and gusts would impact aviation operations.

5. DISCUSSION AND APPLICATIONS

The XGBoost wind prediction models present significant advantages over traditional methods, including:

- The ability to capture non-linear relationships involving multiple factors.
- Improved accuracy in predicting wind speed during TCs considering both historical and forecast data which are critical for aviation safety.
- The potential to quantify forecast uncertainty and provide scenario-based guidance by adjusting TC track and/or intensity deviations.

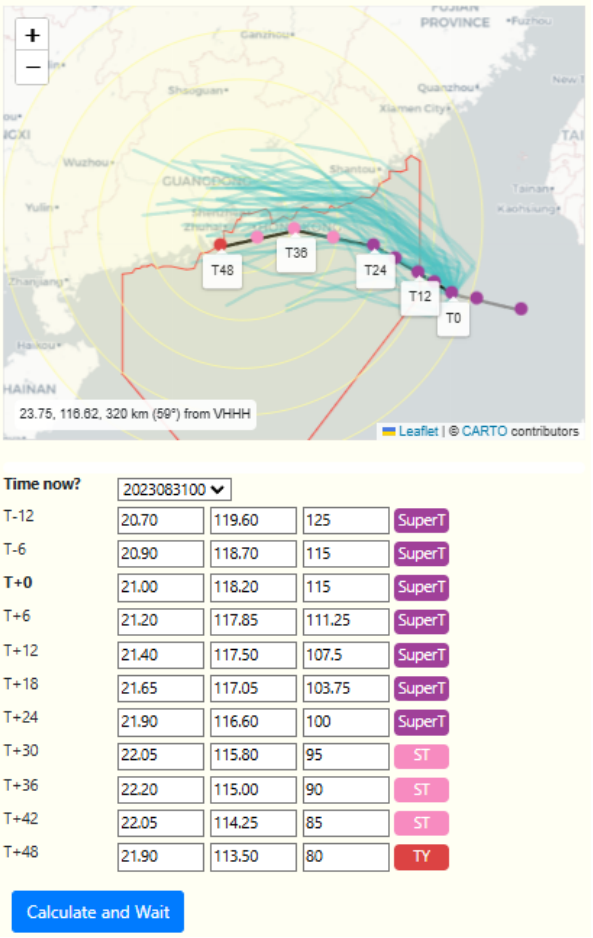


FIG. 9. An interactive webpage has been developed for trial. Example using TC Saola in 2023 is given. Track and intensity could be input by aviation forecasters and ECEPS tracks also plotted for reference.

These models can be integrated into an interactive tool for aviation forecasters, offering:

- Hourly Wind Forecasts: Detailed guidance for 3-hour Aerodrome Forecast (TAF).
- Scenario Analysis: Forecast updates based on alternative TC forecasts.
- Uncertainty Visualisation: Probabilistic forecasts to support risk-based decision-making.

6. CONCLUSION

The XGBoost wind prediction models developed in this study demonstrate the potential to enhance wind forecasting accuracy at HKIA during TCs. By leveraging diverse observational and NWP model-based OCF

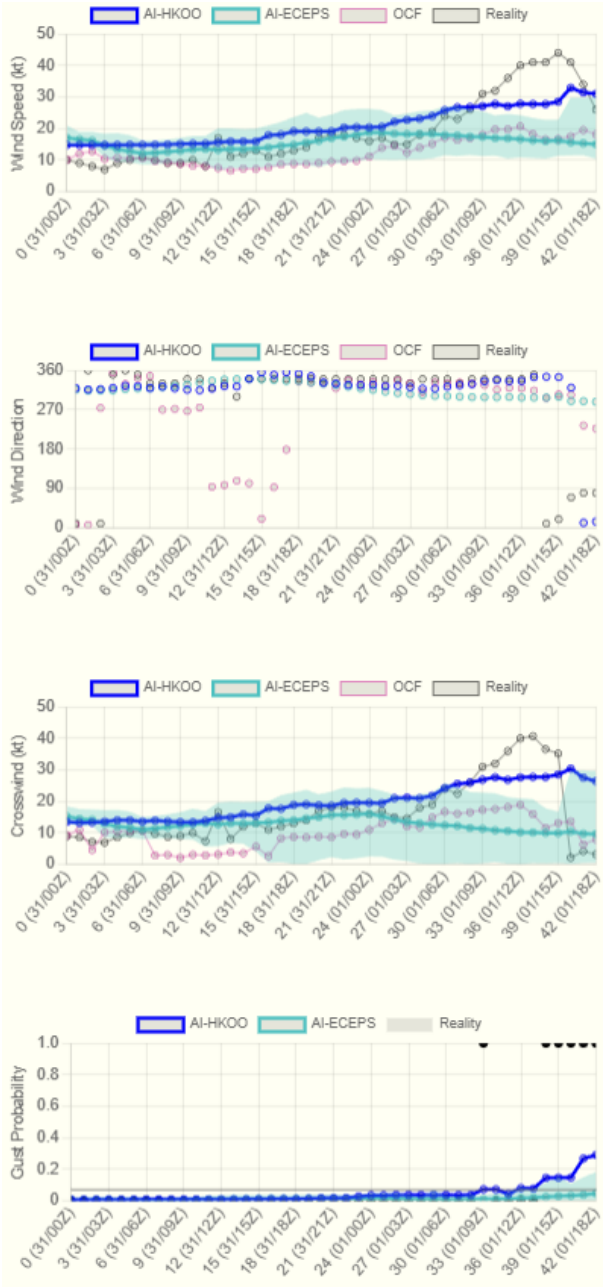


FIG. 10. Example using TC Saola in 2023 is given. Top panel gives wind speed forecast and second panel gives wind direction forecast. Third panel is wind speed forecast of forecast crosswinds and the bottom panel is gust probability.

forecast datasets, the models provide reliable predictions that support aviation safety and operational efficiency. Future work will focus on including wind direction forecast, integrating these models into operational workflows and exploring their applications in other high-impact weather scenarios. It is also planned to incorporate nearby airport observations and more pressure pattern metrics as model features for better forecast accuracy.

Acknowledgments. Special thanks to the staff of the International Aviation Meteorological Collaboration Division in HKO for their valuable comments and technical support. We would also like to extend our thanks to HKO's aviation forecasters at the Aeronautical Meteorological Office at the Hong Kong International Airport who have provided insights and suggestion on the trial interactive tool.

REFERENCES

Akiba, T., S. Sano, T. Yanase, T. Ohta, and M. Koyama, 2019: Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Chan, J. C., 2023: Frequency and intensity of landfalling tropical cyclones in east asia: past variations and future projections. *Meteorology*, **2 (2)**, 171–190.

Chan, M. H. K., W. K. Wong, and K. C. Au-Yeung, 2021: Machine learning in calibrating tropical cyclone intensity forecast of ecmwf eps. *Meteorological applications*, **28 (6)**, e2041.

Chan, P., 2019: Verification of tropical cyclone intensity forecast from lidar observations at hong kong international airport. *Meteorological Applications*, **26 (3)**, 397–404.

Chan, P., and K. Hon, 2016: Performance of super high resolution numerical weather prediction model in forecasting terrain-disrupted airflow at the hong kong international airport: Case studies. *Meteorological Applications*, **23 (1)**, 101–114.

Chan, S., and J. C. Lee, 2013: Objective consensus forecast: A site-specific multi-model consensus forecast in the hong kong observatory. *27th Guangdong-Hong Kong-Macao Seminar on Meteorological Science and Technology*.

Chen, T., and C. Guestrin, 2016: Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Lundberg, S., 2017: A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*.

Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.

Wong, W., and S. Tse, 2016: Development of a quantitative forecast system for tropical cyclone induced rainfall at the hong kong international airport. *Meteorological Applications*, **23 (3)**, 498–504.