Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-8, Issue-3, March 2026
DOI: 10.53469/jrse.2026.08(03).14

Optimization Study of Low-Altitude Turbulence Intensity Modeling Based On TKE-XGBoost

Xi Gong

Business School, University of Shanghai for Science and Technology, Shanghai 200093, China

Abstract: ***To improve the accuracy of turbulence identification in low-altitude flight safety monitoring, a turbulence intensity modeling and optimization method based on Turbulent Kinetic Energy theory and the XGBoost model is proposed. Firstly, atmospheric stability is determined using the Richardson number. Subsequently, turbulence intensity is calculated by combining different stability conditions and Turbulent Kinetic Energy theory. The data for constructing physical model a originates from observations obtained by wind profile radar and microwave radiometer. Considering that temperature and humidity detection equipment like microwave radiometers are not always available in real-world scenarios, and in observation scenarios relying solely on wind profile radar, a Gradient Boosting Decision Tree algorithm is further introduced to construct a turbulence inversion model based exclusively on wind profile radar data. This model fully exploits the nonlinear relationships between multi-dimensional observational features such as radial velocity, velocity spectrum width, and signal-to-noise ratio of the radar and turbulence intensity. It is trained using the output of the benchmark model. Experimental results indicate that after optimizing the turbulence intensity calculation model b (based solely on wind profile radar) with model a, the model's MSE decreases by 0.11, MAE decreases by 0.13, and the R^2 value increases by 0.28. This optimization process reduces reliance on auxiliary temperature and humidity data and effectively addresses the challenge of turbulence identification under limited observational information.***

Keywords: Low-altitude turbulence, Wind profile radar, Microwave radiometer, Machine learning.

1. Introduction

Driven by a new wave of technological revolution and industrial transformation, the low-altitude economy, as a strategically important emerging industry and a key representative of new quality productive forces, is experiencing unprecedented development opportunities due to its immense growth potential and industrial enabling effects. In 2024, the "low-altitude economy" was included in China's Government Work Report for the first time, and nearly 30 provinces have subsequently introduced supportive policies, marking that the low-altitude economy has entered a new phase of large-scale, systematic development, with the industrial ecosystem rapidly taking shape [1].

However, low-altitude flight activities are characterized by small aircraft size, low flight altitude, and complex operational scenarios. Flight areas often cover various underlying surfaces such as cities, forests and mountains, and are significantly influenced by local topography, building clusters, and uneven heating. This leads to frequent occurrences of low-altitude atmospheric turbulence and wind shear, posing serious threats to flight safety [2]. Turbulence, as a typical hazardous weather phenomenon for aviation, has complex and diverse formation mechanisms, including thermal turbulence caused by thermal effects and mechanical turbulence induced by external forces like terrain blocking and surface friction. The scales of such turbulence are variable, and especially when the scale of turbulent eddies approaches the size of the aircraft, it can easily lead to attitude loss of control, crashes, or course deviations.

Currently, China still faces numerous technical challenges in low-altitude turbulence monitoring. Existing detection equipment, such as automatic weather stations, wind profile radars, microwave radiometers, and Doppler weather radars, have limitations in detection modes, spatiotemporal resolution, and detection range, making it difficult to directly and accurately capture the fine structure of low-altitude turbulence [3]. Furthermore, inconsistencies in data formats, detection ranges, and spatiotemporal coverage among different devices also restrict the ability to assess and warn about turbulence intensity [4]. Therefore, developing turbulence intensity evaluation indicators suitable for low-altitude flight characteristics and establishing high spatiotemporal resolution turbulence monitoring and warning technologies under current detection conditions has become a critical scientific issue and urgent task for enhancing low-altitude meteorological support capabilities and supporting the safe development of the low-altitude economy.

2. Methodology

2.1 Problem Description

Low-altitude aircraft, constrained by factors such as low flight altitude and small size, are highly sensitive to low-altitude turbulence, which can easily lead to flight instability or even accidents. Existing turbulence monitoring methods face issues such as diverse detection mechanisms, inconsistent spatiotemporal data resolution, and heterogeneous observational features, limiting the accurate identification and warning capability for low-altitude turbulence. To this end, this paper established a turbulence intensity calculation model integrating data from wind profile radar and microwave radiometer (Model a) and a turbulence intensity calculation model using only wind profile radar data (Model b). The output of Model a will serve as a benchmark to validate and optimize Model b. Low-altitude turbulence is influenced by both thermal and dynamic processes. Wind profile radar can capture dynamic wind field information, while microwave radiometers provide thermal parameters like temperature and humidity.

2.2 Model Assumptions

(1) Hydrostatic Approximation Assumption

It is assumed that the atmosphere in the study area is approximately in a hydrostatic equilibrium state, and within a certain time window, wind shear and thermal stratification can be considered approximately stable.

(2) Applicability of Richardson Number Criterion

It is assumed that the Richardson number R_i can effectively describe atmospheric stability in the low-altitude layer, and the condition $R_i < 0.25$ can serve as a criterion for turbulence occurrence [5].

(3) Statistical Stability of Observation Errors Assumption

It is assumed that the systematic errors of different observation instruments remain relatively stable during the study period and can be represented by weights and uncertainty indicators.

(4) Monin-Obukhov Similarity Theory Assumption

It is assumed that all turbulence statistics are determined solely by height Z_0, friction velocity u_*, Obukhov length L, and buoyancy parameter (g/T).

2.3 Symbol Description

Table 1: Symbol Description

Symbol	Meaning	Unit
T_z	Actual air temperature at height z	K
q_z	Specific humidity at height z	kg/kg
U_z	Horizontal average wind speed at height z	m/s
V_z	Vertical average wind speed at height z	m/s
W_z	Average depth-wise wind speed at height z	m/s
Z_0	Height	m
R_i	Richardson number	—
TKE	Turbulent Kinetic Energy	m^2/s^2
$T(x,y,z,t)$	Three-dimensional turbulence intensity field	—
V_r	Radial wind speed	m/s

3. Model of Low-Altitude Turbulence Intensity Model Based on TKE-XGBoost

First, utilizing observational data from both wind profile radar and microwave radiometer, a reliable benchmark model for turbulence intensity (Model a) is established based on Turbulent Kinetic Energy (TKE) theory, considering comprehensive factors. Second, targeting practical application scenarios where only wind profile radar data are available, multi-dimensional features such as wind field radial velocity and velocity spectrum width are extracted, and an eXtreme Gradient Boosting (XGBoost) model is constructed. Finally, using the calculation results of Model a as learning targets, Model b is trained and optimized through supervision, thereby establishing a turbulence intensity calculation model based on a single wind profile radar that has low dependence on external temperature and humidity data, high result accuracy, and strong practicality

3.1 Dual-Source Data-Driven Benchmark Model for Turbulence Intensity (Model a)

To obtain a reliable benchmark for turbulence intensity, a mathematical model combining wind profile radar and microwave radiometer data is established. The wind profile radar can accurately measure the vertical structure of the wind field, while the microwave radiometer provides vertical profiles of atmospheric temperature and humidity. Combining both allows for turbulence intensity calculation that considers dynamic factors and incorporates thermal factors for correction. This paper selects Turbulent Kinetic Energy (TKE) as the core indicator for measuring turbulence intensity. TKE refers to the kinetic energy per unit mass of fluid due to turbulent motion, can reasonably describe turbulence intensity, and reflects the intensity of turbulent motion [6]. The modeling steps for Model a are as follows:

(1) Virtual Temperature Calculation

Water vapor in the atmosphere reduces air density and enhances buoyancy-driven turbulence. Therefore, humidity data provided by the microwave radiometer affects air buoyancy through virtual temperature T_v. The formula for T_v is:

$$T_v = T \times (1 + 0.61q) \tag{1}$$

where T is the actual temperature, 0.61 is the coefficient reflecting the influence of water vapor on density, and q is the specific humidity.

(2) Stability Determination

The gradient Richardson number $Ri_g$is calculated based on vertical gradients to determine stability.

$$Ri_g = \frac{g}{T_{v0}} \times \frac{\Delta T_v/\Delta z}{(\Delta U/\Delta z)^2} \tag{2}$$

where

$$\Delta T_v = T_{v(Z_1)} - T_{v(Z_2)} \tag{3}$$

$$\Delta Z = Z_2 - Z_1, Z_2 > Z_1 \tag{4}$$

$$\Delta U = U_{(Z_2)} - U_{(Z_1)} \tag{5}$$

Here, $\Delta T_v/\Delta z$ is the vertical gradient of virtual temperature, $\Delta U/\Delta z$ is the vertical gradient of wind speed, g is the acceleration due to gravity (9.8 m/s²), and T_{v0} is the virtual temperature at a reference layer (taken as the average of the observed layers to reduce errors). Generally, when Ri_g is less than the critical Richardson number $Ri_c \approx 0.25$, the atmospheric stratification is unstable, and turbulence is likely to occur. The smaller the Ri_g value, the stronger the dynamic effect, and the more intense the turbulence.

(3) Parameterization of Turbulent Exchange Coefficients

Based on Monin-Obukhov Similarity Theory (MOST), the momentum coefficient K_m and the eddy diffusivity for heat K_h are parameterized. The Monin-Obukhov length L and friction velocity u_* are core parameters. The Monin-Obukhov length L is a dimensionless quantitative indicator of stability, positively correlated with Ri_g. According to previous research results [7], its calculation formula is:

$$L = \frac{{u_*}^2 T_v}{kg\overline{w'T'}} \tag{6}$$

where k is the von Kármán constant, taken as 0.4 in this paper; T_v is the virtual temperature; T' is the virtual temperature fluctuation.

$$u_*{}^2 = \left(\overline{u'w'^2} + \overline{v'w'^2}\right)^{1/2} \tag{7}$$

where u'、v'、w' are wind speed fluctuations.

Based on MOST, K_m and K_h can be expressed as:

$$K_m = \frac{kzu_*}{\phi_m(\zeta)} \tag{8}$$

$$K_h = \frac{kzu_*}{\phi_h(\zeta)} \tag{9}$$

where $\zeta = z/L$, ζ is the dimensionless height.

$$\phi_m(\zeta) = \begin{cases} (1-16\zeta)^{-1/4}, & \zeta < 0 \\ 1, & \zeta = 0 \\ 1 + 4.7\zeta, & \zeta > 0 \end{cases} \tag{10}$$

$$\phi_h(\zeta) = \begin{cases} (1-16\zeta)^{-1/2}, & \zeta < 0 \\ 1, & \zeta = 0 \\ 1 + 6.8\zeta, & \zeta > 0 \end{cases} \tag{11}$$

(4) Turbulent Kinetic Energy Calculation

Turbulent Kinetic Energy TKE consists of the shear production term P_s and the buoyancy production term P_b. Their expressions are as follows:

$$TKE = C \times (P_s + P_b) \tag{12}$$

Where

$$P_s = K_m \times \left(\frac{\Delta U}{\Delta z}\right)^2 \tag{13}$$

$$P_\beta = -K_h \times \frac{g}{T_{v0}} \times \frac{\Delta T_v}{\Delta z} \tag{14}$$

C is an empirical constant, $P_s \geqslant 0$, when buoyancy suppresses turbulence, $P_\beta < 0$.

(5) Turbulence Intensity Calculation

Turbulence intensity is the ratio of the fluctuating wind speed amplitude to the mean wind speed, reflecting the relative intensity of turbulence. It is often represented by the three-dimensional turbulence intensity I_3D, with the formula:

$$I_3D = \frac{\sqrt{\frac{2}{3}TKE}}{U} \tag{15}$$

where U is the mean wind speed at the reference height. A larger I_3D indicates more intense wind speed fluctuations.

3.2 Single Wind Profile Radar Turbulence Intensity Calculation Model (Model b)

In scenarios driven solely by wind profile radar data, the accuracy of turbulence intensity detection results can be affected due to the lack of temperature data [8]. Therefore, this paper aims to employ data-driven machine learning methods to uncover the deep-seated relationships between wind profile radar data and turbulence intensity [9]. The construction and optimization of Model b are divided into two stages: feature extraction and model training.

(1) Feature Extraction

In the feature extraction stage, feature vectors X_i capable of reflecting the turbulence state are extracted from the wind profile radar data. For each height level Z_i, the feature vector X_i is constructed as follows:

$$X_i = \left[w_s(z_i), u(z_i), v(z_i), \frac{\partial u}{\partial z}\Big|_i, \frac{\partial v}{\partial z}\Big|_i, \frac{\partial^2 u}{\partial z^2}\Big|_i, \frac{\partial^2 v}{\partial z^2}\Big|_i, \dots\right] \tag{16}$$

This vector includes the velocity spectrum width w_s measured directly by the radar, the horizontal wind speed components u, v and the first and second-order wind field gradients obtained through calculation.

(2) Model Training

In the model training stage, the goal is to find a nonlinear mapping function f such that $\hat{y}_i = f(X_i)$, where $\hat{y}_i$ is an accurate estimate of the true turbulence intensity y_i. This paper selects the eXtreme Gradient Boosting (XGBoost) model to construct this mapping. The XGBoost model is an ensemble of K decision trees f_k:

$$\hat{y}_i = f(X_i) = \sum_k f_k(X_i) \tag{17}$$

Model training aims to minimize an objective function L that includes a loss function and a regularization term:

$$L = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \tag{18}$$

where $l(y_i, \hat{y}_i)$ is the loss function measuring the difference between predicted and true values, and $\Omega(f_k)$ is the regularization term penalizing model complexity.

4. Results and Discussion

4.1 Experimental Setup

To validate the performance of the final model, turbulence data from wind profile radars at points a and b were divided into training and test sets in an 80% and 20% ratio, respectively. Prediction accuracy was quantified by calculating statistical indicators such as Mean Squared Error MSE, Mean Absolute Error MAE, and Coefficient of Determination R^2, and compared against linear regression and Random Forest algorithms.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{19}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{20}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{21}$$

4.2 Results and Discussion

Using the methods described above, the experimental results are shown in Table 2.

Table 2: Comparative Experimental Results of Turbulence Intensity Calculation Using Single Wind Profile Radar

Model	MSE	MAE	R^2
LR	0.303269	0.262628	0.303100
RF	0.191934	0.234761	0.581642
XGBoost(before)	0.136276	0.154659	0.686843
XGBoost(after)	0.002670	0.020964	0.965688

The comparative experimental results indicate that the Random Forest algorithm performs poorly when handling high-dimensional sparse data and requires improvement based on different dataset types [10]. In contrast, XGBoost demonstrates higher accuracy and efficiency in processing sparse data, leading to its widespread application in fields like hydrometeorology and atmospheric dynamics. The experimental results show that after optimization using Model a, Model b's MSE decreased by 0.11, MAE decreased by 0.13, and the R^2 value increased by 0.28, indicating that the accuracy and data fit of Model b's turbulence calculation were improved to a certain extent through multi-source data optimization.

To further illustrate the turbulence intensity calculated by the optimized XGBoost model, we compared the turbulence intensity from the optimized Model b with that calculated by Model a, obtaining the result, as Figure 1 shows.

The wind profile radar and microwave radiometer data from points a and b were matched and aligned in time and vertical height, missing values were interpolated, and outliers in the velocity spectrum width from the wind profile radar were removed. The obtained wind speed and temperature data were categorized according to four time points from 00:00 to 00:24 at 8-minute intervals. The height and Richardson number for each time point were calculated, and the resulting vertical profiles of turbulence parameters are shown in Figure 4.2. The figure shows that at point a, the R_i values between 100-1000m are all above 0.25, indicating overall stable conditions in the low-altitude layer. At point a, TKE exhibits significant fluctuations within the 1000m-2000m height range. This result is consistent with the meteorological background of atmospheric instability during the early morning hours. Turbulence at point b is more dispersed compared to point a, with unstable airflow present in the 1000m-1500m and 2000-2500m ranges. Caution should be exercised for low-altitude operations during these periods.

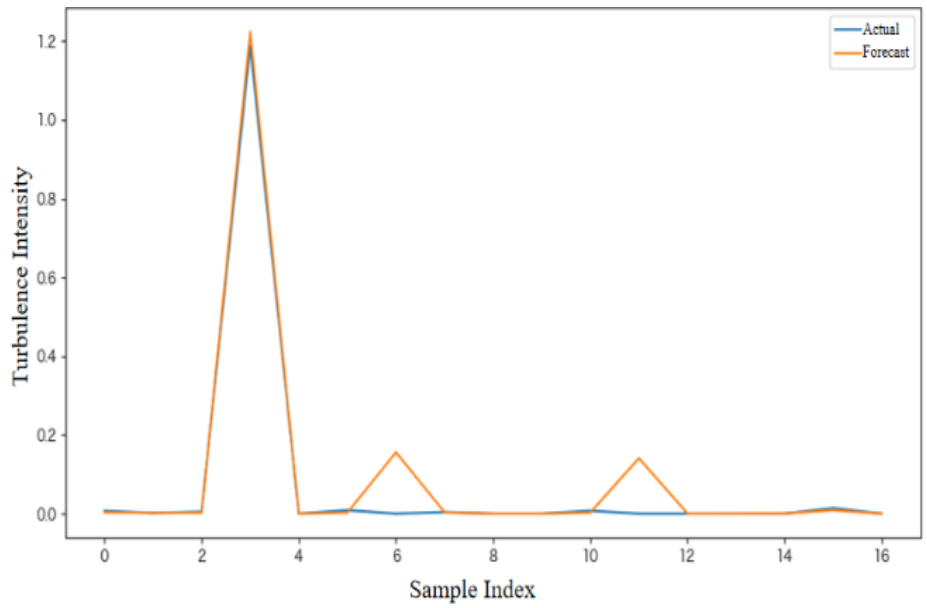
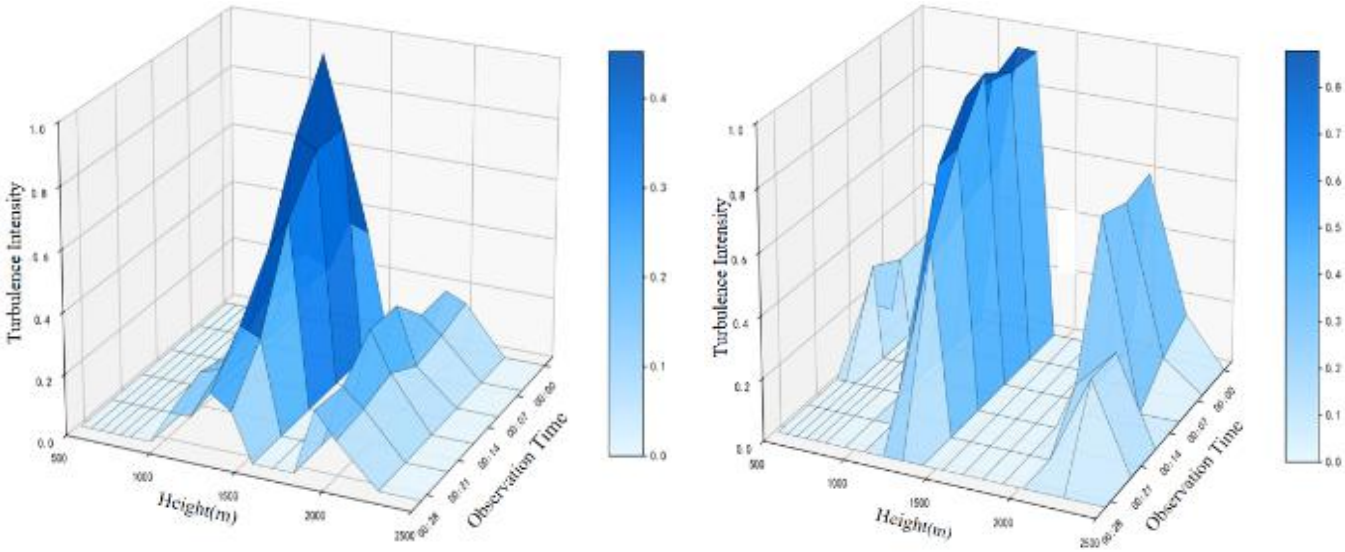


Figure 1: Comparison of Turbulence Intensity between Optimized Model b and Model a



(a) point a (b) point b

Figure 2: Time-Height Three-Dimensional Distribution of Turbulent Kinetic Energy

5. Conclusions

This paper constructed a turbulence intensity calculation model (Model a) using wind profile radar data and microwave radiometer observations. The output of Model a was then used as the standard turbulence intensity to optimize a single wind profile radar-based turbulence intensity calculation model (Model b) using XGBoost. Experimental results demonstrate that the optimized Model b exhibits significantly improved accuracy and fitting capability, validating the effectiveness and practicality of the proposed method. This provides an alternative solution for accurately calculating turbulence intensity at sites lacking microwave radiometer observation conditions.

However, the turbulence intensity calculation Model a, based on integrated Richardson number and TKE theory, relies heavily on observational data, requiring synchronized, high-quality observations from both wind profile radar and microwave radiometer. Furthermore, the sample coverage directly affects the model's generalizability. Overall, the optimization method proposed in this paper offers a feasible approach for accurately calculating turbulence intensity using a single wind profile radar. Nevertheless, subsequent efforts should focus on enhancing the method's robustness and applicability by expanding observational samples and introducing independent validation data.

References

[1] Song J, Bai H, Cheng B, et al. A review of large eddy simulation of aviation turbulence[J]. Acta Aerodynamica Sinica, 2023,41(8):26-43.
[2] Ji H, Chen R, Lu L, et al. Pilot workload investigation for rotorcraft operation in low-altitude atmospheric turbulence[J]. Aerospace Science and Technology, 2021, 111: 106567.
[3] Ren X, Liu Y, Cai Z, et al. Observations of Dynamic Turbulence in the Lower Stratosphere over Inner Mongolia Using a High-resolution Balloon Sensor Constant Temperature Anemometer[J]. Advances in Atmospheric Sciences, 2022, 39(3): 519-528.
[4] Kaiser E, Kutz J N, Brunton S L. Sparse identification of nonlinear dynamics for model predictive control in the low-data limit[J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2018, 474(2219).
[5] Banakh V A, Falits A V, Sherstobitov A M, et al. On estimation of the turbulent mixing layer altitude from the altitude-time distributions of the Richardson number[J]. Atmospheric and Oceanic Optics, 2023, 36(1): 30-40.
[6] Wu C, Deng J, Zhou X, et al. Turbulence affected by submerged aquatic vegetation under wind-induced flow[J]. Physics of Fluids, 2025, 37(1).
[7] Ding W, Shi Y, Liu L, et al. Research on vertical distribution characteristics and calculation method of turbulence intensity above urban canopy based on Beijing 325m Meteorological Tower[J]. Building Structure,2024,54(13):71-80.
[8] Sinha S, Regeena M L, Sarma T V C, et al. Doppler profile tracing using MPCF on MU radar and sodar: Performance analysis[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(4): 508-511.
[9] O'Connell M, Shi G, Shi X, et al. Neural-fly enables rapid learning for agile flight in strong winds[J]. Science Robotics, 2022, 7(66): eabm6597.
[10] Niazkar M, Menapace A, Brentan B, et al. Applications of XGBoost in water resources engineering: A systematic literature review [J]. Environmental Modelling & Software, 2024, 174: 105971.