Volume 3, Issue 3, 2025

Print ISSN: 2959-9865
Online ISSN: 2959-9873

WORLD JOURNAL OF ENGINEERING RESEARCH

World Journal of Engineering Research

Volume 3, Issue 3, 2025

Published by Upubscience Publisher

World Journal of Engineering Research
Print ISSN: 2959-9865 Online ISSN: 2959-9873
Email: info@upubscience.com
Website: http://www.upubscience.com/

Table of Content

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3032

MEASUREMENT OF THE CENTER OF GRAVITY OF A WATER BASED UAV BASED ON THE TWO-POINT SUSPENSION METHOD

ShuHang Peng*, Hang Zhang
Wuhan R&D Center, China Special Vehicle Research Institute, Wuhan 430070, Hubei, China.
Corresponding Author: ShuHang Peng, Email: pengsh006@avic.com

Abstract: Addressing the limitations of traditional jack suspension methods, which are unsuitable for water based UAV without landing gear, and the challenges of obtaining the center of gravity height position of water based UAV using the two-point suspension method, this paper proposes reading the force sensor readings at two suspension points under different small pitch angles and simplifying the torque equation through coordinate transformations. Based on this, the optimal estimates of the center of gravity position are calculated using the Fermat method and the least squares method, respectively. The results indicate that, while maintaining the same experimental costs and operational difficulty, the method proposed in this paper reduces experimental errors and improves the effectiveness and accuracy of center of gravity calculations for UAVs.
Keywords: Two-point suspension method; Water based UAV; Coordinate transformation; Fermat's method; Least squares method

1 INTRODUCTION

In the design process of unmanned aerial vehicles (UAVs), the center of gravity (COG) is a critical focus for designers [1-2], spanning the entire research, development, production, testing, and operational phases [3-7]. Due to various unfavorable factors such as theoretical weight errors during design, manufacturing errors in parts production, and assembly errors during final assembly, the actual center of gravity position of an UAV often deviates from the designed center of gravity position after it rolls off the production line. However, the center of gravity position has a significant impact on flight control program development, flight safety, and quality [8-13]. Therefore, the center of gravity must be calculated and calibrated before UAV test flights and delivery. Currently, commonly used methods for calculating the center of gravity of aircraft include the jack measurement method and the suspension method [4]. The hull of a water based UAV consists of the upper hull above the waterline and the lower hull below the waterline, with the lower hull having a “V”-shaped cross-section. The jack measurement method cannot be directly applied, so the suspension method is generally used to calculate the weight center of gravity of water based UAVs.
This paper employs the two-point suspension method to obtain force sensor readings at different pitch angles, simplifies the torque equation by changing the reference frame, and uses different fitting methods to calculate the optimal estimate of the center of gravity position. It explores a low-cost, easy-to-operate method for calculating the center of gravity of unmanned aerial vehicles and processing experimental data.

2 CALCULATION PRINCIPLES AND ERROR CORRECTION

2.1 Calculation Principle of Center of Gravity Calculation Method

As shown in Figure 1, this is the force situation of the UAV in the absolute coordinate system $x_gO_gy_g$ when the pitch angle is α (in the situation shown in the figure, the pitch angle is positive, so $\alpha > 0$).

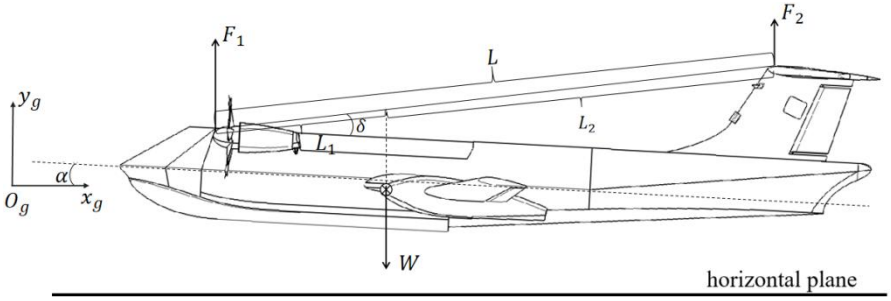

Figure 1 Force Diagram of the UAV

In Figure 1, F_1 and F_2 represent the readings of the tension gauges at the front and rear suspension points, respectively. W represents the weight of the UAV, L represents the distance between the front and rear suspension points, δ represents the angle between the line connecting the front and rear suspension points and the aircraft's centerline, L_1 and L_2 represent the distances from the front and rear suspension points on the suspension line to the line of gravity, respectively. In an absolute coordinate system $x_gO_gy_g$, the equilibrium equation can be expressed as:

$$\begin{cases} F_1 + F_2 = W \\ F_1 \cdot L_1 = F_2 \cdot L_2 \\ L_1 + L_2 = L \end{cases} \tag{1}$$

In equation (1), the weight of the UAV is taken as the average value of N times horizontal measurements:

$$W = \frac{1}{N}\sum_{i=1}^{N}\left(F_{1,i} + F_{2,i}\right) \tag{2}$$

According to equation (1), we can conclude that:

$$L_1 = \frac{F_2 \cdot L}{W}, L_2 = \frac{F_1 \cdot L}{W} \tag{3}$$

From equation (3), we can determine the position of the center of gravity on the fuselage axis, but we cannot determine its position in terms of fuselage height.

To further determine the position of the center of gravity in terms of fuselage height, multiple sets of data can be measured under various pitch angles, and the equations can be solved simultaneously. However, directly setting up equations in the absolute coordinate system $x_gO_gy_g$ results in nonlinear terms. To simplify the equations for easier solution, the reference frame can be transformed into a fixed reference frame centered at the front suspension point on the fuselage, as shown in Figure 2. The x-axis of this reference frame coincide with the line connecting the suspension points pointing backward, and the y-axis is perpendicular to the x-axis pointing upward. Let the coordinates of the origin O of the suspension point coordinate system xOy in the absolute coordinate system $x_gO_gy_g$ be (x_{GO}, y_{GO}), then the coordinate transformation rule between the absolute coordinate system $x_gO_gy_g$ and the suspension point coordinate system xOy is:

$$\begin{bmatrix} x_G \\ y_G \end{bmatrix} = \begin{bmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_{GO} \\ y_{GO} \end{bmatrix} \tag{4}$$

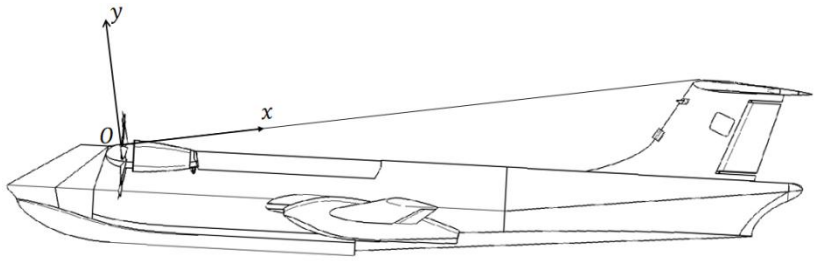

Figure 2 Diagram of Suspension Point Coordinate System

Figure 3 shows the force conditions of the UAV during two measurements in the suspension point coordinate system. When the suspension point coordinate system is used as the stationary reference system, the UAV remains stationary in the suspension point coordinate system when the aircraft pitch angle α is changed, while the angle θ between the direction of the tension gauge and the x-axis of the suspension point coordinate system changes with the pitch angle α. Their relationship is as follows:

$$\theta = 90° - \alpha - \delta \tag{5}$$

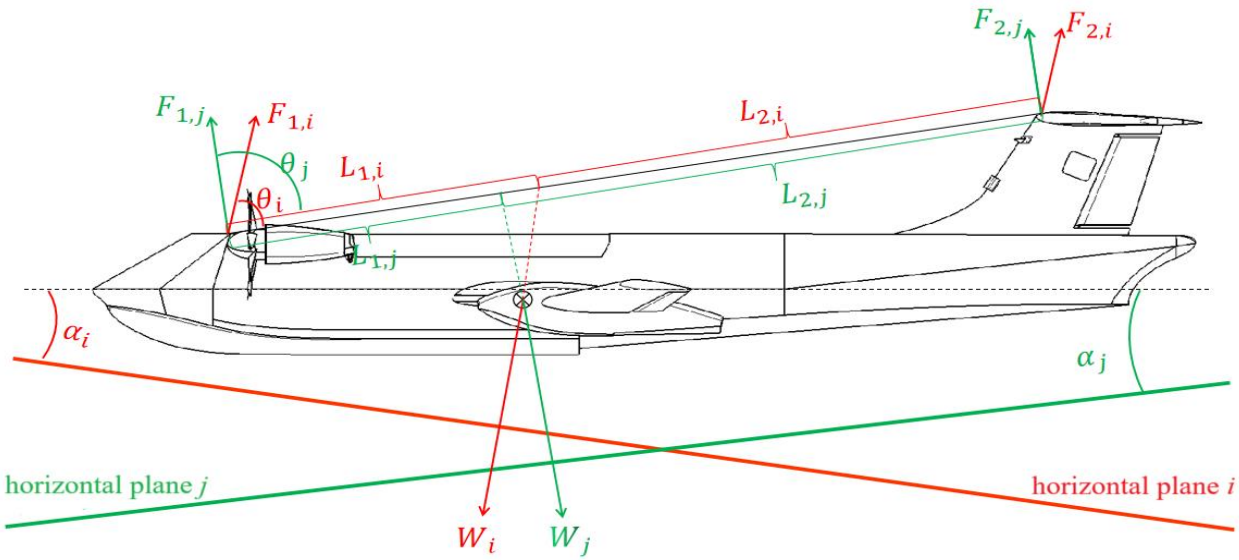

Figure 3 Force Conditions on the UAV during Two Measurements in the Suspension Point Coordinate System

Similarly to equation (3), the following relationship can be written:

$$\begin{cases} L_{1,i} = \dfrac{F_{2,i} \cdot L}{W}, L_{2,i} = \dfrac{F_{1,i} \cdot L}{W} \\ L_{1,j} = \dfrac{F_{2,j} \cdot L}{W}, L_{2,j} = \dfrac{F_{1,j} \cdot L}{W} \end{cases} \tag{6}$$

Furthermore, as can be seen from Figure 3, in the suspension point coordinate system, since the UAV remains stationary relative to the suspension point coordinate system, the position of the center of gravity at each measurement is independent of the pitch angle at that time, i.e., the gravity lines obtained from each measurement intersect at the same point, and that point is the center of gravity.

Based on Figure 3 and Equation (6), the intercepts b_i and b_j of the gravity lines on the x-axis at the two measurement times are:

$$\begin{cases} b_i = \dfrac{F_{2,i} \cdot L}{W} \\ b_j = \dfrac{F_{2,j} \cdot L}{W} \end{cases} \tag{7}$$

From Figure 2, we can see that the slopes of the two gravity lines measured are $k_i = \tan\theta_i$ and $k_j = \tan\theta_j$, respectively. Therefore, we can write the linear equations of the two gravity lines:

$$\begin{cases} l_i = \left(x - \dfrac{F_{2,i} \cdot L}{W} \right) \cdot \tan\theta_i \\ l_j = \left(x - \dfrac{F_{2,j} \cdot L}{W} \right) \cdot \tan\theta_j \end{cases} \tag{8}$$

Solving the system of equations (8), we obtain the coordinates (x_C, y_C) of the center of gravity in the suspension point coordinate system as:

$$\begin{cases} x_C = \dfrac{L\left(F_{2,i}\tan\theta_i - F_{2,j}\tan\theta_j\right)}{W\left(\tan\theta_i - \tan\theta_j\right)} \\ y_C = \dfrac{L\left(F_{2,i} - F_{2,j}\right)\tan\theta_i \tan\theta_j}{W\left(\tan\theta_i - \tan\theta_j\right)} \end{cases} \tag{9}$$

Substituting equation (9) into equation (4) yields the coordinates $(x_{C,G}, y_{C,G})$ of the center of gravity in the absolute coordinate system:

$$\begin{bmatrix} x_{C,G} \\ y_{C,G} \end{bmatrix} = \begin{bmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{bmatrix} \begin{bmatrix} x_C \\ y_C \end{bmatrix} + \begin{bmatrix} x_{GO} \\ y_{GO} \end{bmatrix} \tag{10}$$

2.2 Correction of Raw Data

During the experiment, the UAV was suspended from a gantry frame by two suspension cables. Ideally, the two cables should be parallel to each other and perpendicular to the ground. However, during the actual experiment, due to changes

in the UAV's pitch angle, the horizontal distance between the two suspension points on the UAV also underwent minor changes, causing the two cables to neither be parallel to each other nor perpendicular to the ground. However, the angle deviation between the cables and the vertical direction was ignored in the aforementioned calculations. The current revision addresses the impact of the deviation angle of the suspension cables on experimental results when the pitch angle is non-zero.

When the UAV is at a non-zero pitch angle, its force diagram in the absolute coordinate system is shown in Figure 4. In Figure 4, γ_i indicates the deviation angle between F_i and the vertical direction at this time.

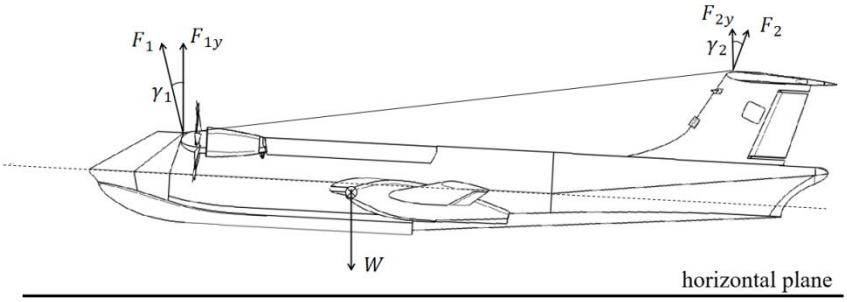

Figure 4 Force Diagram of a UAV at a Non-Zero Pitch Angle

$$\begin{cases} F_1\cos\gamma_1 + F_2\cos\gamma_2 = W \\ F_1\sin\gamma_1 = F_2\sin\gamma_2 \end{cases} \tag{11}$$

Based on the force diagram, the equilibrium equation can be listed as follows:

When the pitch angle is taken as 0°, substitute the measured W into equation (11).:

$$\cos\gamma_1 = \frac{F_1^2 + W^2 - F_2^2}{2F_1W}, \cos\gamma_2 = \frac{F_2^2 + W^2 - F_1^2}{2F_2W} \tag{12}$$

Then, calculate using the corrected original data $F_{1y} = F_1\cos\gamma_1, F_{2y} = F_2\cos\gamma_2$ and substitute it into equations (9) and (10).

3 EXPERIMENTAL DATA PROCESSING METHODS

If a total of N pitch angles are selected for weighing, a total of C_N^2 equations similar to Equation (8) can be obtained. Solving these equations yields C_N^2 sample points, which are denoted as $(x_k, y_k), k \in [1, C_N^2]$. Due to the existence of errors, these C_N^2 sample points do not coincide.

3.1 Calculating the Estimated Center of Gravity Position using the Feynman Method

Let the true value of the center of gravity coordinates be (x_0, y_0), and the distance from the k th sample point to the center of gravity be:

$$d_k(x_0, y_0) = \sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2} \tag{13}$$

The sum of the distances from each sample point to the center of gravity is:

$$D_1(x_0, y_0) = \sum_{k=1}^{C_N^2} d_k(x_0, y_0) = \sum_{k=1}^{C_N^2} \sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2} \tag{14}$$

According to Fermat's principle, the value of $(\hat{x}_0, \hat{y}_0)$ that minimizes D_1 is considered to be the best estimate of (x_0, y_0). Take the second-order partial derivative of the function $D_1(x_0, y_0)$ in equation (14):

$$\begin{cases} \dfrac{\partial^2 D_1}{\partial x_0^2} = \displaystyle\sum_{k=1}^{C_N^2} \dfrac{(y_0 - y_k)^2}{\left[\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}\right]^3} \\ \dfrac{\partial^2 D_1}{\partial x_0 \partial y_0} = \displaystyle\sum_{k=1}^{C_N^2} \dfrac{(x_0 - x_k)(y_0 - y_k)}{\left[\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}\right]^3} \\ \dfrac{\partial^2 D_1}{\partial y_0^2} = \displaystyle\sum_{k=1}^{C_N^2} \dfrac{(x_0 - x_k)^2}{\left[\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}\right]^3} \end{cases} \tag{15}$$

The critical point discriminant of $D_1(x_0, y_0)$ can be written from equation (15) as:

$$\left(\frac{\partial^2 D_1}{\partial x_0 \partial y_0}\right)^2 - \frac{\partial^2 D_1}{\partial x_0^2}\frac{\partial^2 D_1}{\partial y_0^2} = \left\{\sum_{k=1}^{C_N^2} \frac{(x_0 - x_k)(y_0 - y_k)}{[d_k(x_0, y_0)]^3}\right\}^2 - \sum_{k=1}^{C_N^2} \frac{(x_0 - x_k)^2}{[d_k(x_0, y_0)]^3} \cdot \sum_{k=1}^{C_N^2} \frac{(y_0 - y_k)^2}{[d_k(x_0, y_0)]^3} \tag{16}$$

According to Cauchy's inequality, equation (16) is always less than 0. Considering that the second derivative is always greater than 0, the minimum value of $D_1(x_0, y_0)$ is obtained at the zero point of the first partial derivative.

$$\begin{cases} \dfrac{\partial D_1}{\partial x_0} = \displaystyle\sum_{k=1}^{C_N^2} \dfrac{x_0 - x_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}} = 0 \\ \dfrac{\partial D_1}{\partial y_0} = \displaystyle\sum_{k=1}^{C_N^2} \dfrac{y_0 - y_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}} = 0 \end{cases} \tag{17}$$

Perform an identity transformation on (17):

$$x_0 = \frac{\displaystyle\sum_{k=1}^{C_N^2} \frac{x_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}}}{\displaystyle\sum_{k=1}^{C_N^2} \frac{1}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}}}, y_0 = \frac{\displaystyle\sum_{k=1}^{C_N^2} \frac{y_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}}}{\displaystyle\sum_{k=1}^{C_N^2} \frac{1}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}}} \tag{18}$$

Equation (18) satisfies the typical form of general iterative methods and can be rewritten as the following iterative equation:

$$x_0^{(p+1)} = \varphi_x\left(x_0^{(p)}, y_0^{(p)}\right), y_0^{(p+1)} = \varphi_y\left(x_0^{(p)}, y_0^{(p)}\right) \tag{19}$$

Select appropriate initial value $\left(x_0^{(0)}, y_0^{(0)}\right)$ and control accuracy ε. When the accuracy requirements meet $\left(x_0^{(p)} - x_0^{(p+1)}\right)^2 + \left(y_0^{(p)} - y_0^{(p+1)}\right)^2 \le \varepsilon^2$, the center of gravity position $\left(x_0^{(p+1)}, y_0^{(p+1)}\right)$ is considered to be the optimal estimated coordinate in the suspension point coordinate system.

3.2 Calculating the Estimated Center of Gravity Position Using the Least Squares Method

From equation (13), the distance from the k th sample point to the center of gravity is $d_k(x_0, y_0) = \sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}$. The sum of the squares of the distances from each sample point to the center of gravity is:

$$D_2(x_0, y_0) = \sum_{k=1}^{C_N^2} d_k^2(x_0, y_0) = \sum_{k=1}^{C_N^2} \left[(x_0 - x_k)^2 + (y_0 - y_k)^2\right] \tag{20}$$

According to the principle of least squares, the value of $(\hat{x}_0, \hat{y}_0)$ that minimizes D_2 is considered to be the best estimate of (x_0, y_0). Calculate the second-order partial derivative of the function $D_2(x_0, y_0)$ in equation (20):

$$\begin{cases} \dfrac{\partial^2 D_1}{\partial x_0^2} = 2C_N^2 \\ \dfrac{\partial^2 D_1}{\partial x_0 \partial y_0} = 0 \\ \dfrac{\partial^2 D_1}{\partial y_0^2} = 2C_N^2 \end{cases} \tag{21}$$

Obviously, the discriminant of the extreme point of $D_2(x_0, y_0)$ is negative and the two second-order partial derivatives are greater than 0, then the zero point of the first-order derivative is the minimum point of the original function. Let:

$$\begin{cases} \dfrac{\partial D_1}{\partial x_0} = 2C_N^2 \cdot x_0 - 2\sum_{k=1}^{C_N^2} x_k = 0 \\ \dfrac{\partial D_1}{\partial y_0} = 2C_N^2 \cdot y_0 - 2\sum_{k=1}^{C_N^2} y_k = 0 \end{cases} \tag{22}$$

It can be seen that under the least squares criterion, the optimal estimate of the center of gravity is the arithmetic mean of each sample point:

$$x_0 = \frac{1}{C_N^2}\sum_{k=1}^{C_N^2} x_k, y_0 = \frac{1}{C_N^2}\sum_{k=1}^{C_N^2} y_k \tag{23}$$

4 PRACTICAL APPLICATIONS AND EXAMPLES

In the center of gravity calculation experiment for a certain type of surface unmanned aerial vehicle, the readings of the front suspension point force gauge F_1 and rear suspension point force gauge F_2 were taken within a pitch angle range of -3.0° to 3.0°, with readings recorded every 0.5°, resulting in N = 13 sets of data. Three experiments were conducted, and the average values obtained are shown in Table 1.

Table 1 Weight Record Table

Item	1	2	3	4	5	6	7	8	9	10	11	12	14
$\alpha(°)$	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0
F_1/W	0.639	0.64	0.642	0.643	0.643	0.645	0.645	0.646	0.648	0.649	0.652	0.653	0.639
F_2/W	0.365	0.364	0.362	0.359	0.358	0.356	0.355	0.354	0.353	0.352	0.352	0.351	0.365

Based on the 13 sets of data in Table 1, $C_{13}^2 = 78$ sample points $(x_k, y_k), k \in [1,78]$ can be obtained using Equation (9).

Using the Fermat method to process the coordinate data of the 78 sample points, the estimated coordinates of the center of gravity in the suspension point coordinate system are obtained as $\hat{x}_0$ = 1226.670 mm and $\hat{y}_0$ = -560.231 mm. Substituting these estimated values into Equation (10), the estimated coordinates of the center of gravity in the absolute coordinate system are obtained as $\hat{x}_{C,G}$ = 2413.913 mm and $\hat{y}_{C,G}$ = -67.594 mm.

Using the least squares method to process the coordinate data of the 78 sample points, the estimated coordinates of the center of gravity in the suspension point coordinate system are obtained as $\hat{x}_0$ = 1227.668 mm and $\hat{y}_0$ = -545.851 mm. Substituting these estimated values into equation (10), the estimated coordinates of the center of gravity in the absolute coordinate system are obtained as $\hat{x}_{C,G}$ = 2412.494 mm and $\hat{y}_{C,G}$ = -53.249 mm.

Comparing the results of the two calculation methods, the estimated results for the horizontal coordinate show a small difference of 0.13% MAC, indicating good agreement; the estimated results for the vertical coordinate differ by 1.31% MAC, showing a larger difference compared to the horizontal coordinate.

The data in Table 1 is corrected according to $F_{1y} = F_1 \cos\gamma_1, F_{2y} = F_2 \cos\gamma_2$, the results are shown in Table 2:

Table 2 Corrected Weight Data

Item	1	2	3	4	5	6	7	8	9	10	11	12	14
α(°)	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0
F_{1y}/W	0.638	0.639	0.641	0.642	0.643	0.644	0.645	0.646	0.647	0.649	0.651	0.652	0.653
F_{2y}/W	0.362	0.361	0.359	0.359	0.357	0.356	0.355	0.354	0.353	0.351	0.349	0.348	0.347

Using the data from Table 2 in the calculation, the following results are obtained:
Using the Fermat method, the estimated coordinates of the center of gravity in the suspension point coordinate system are $\hat{x}_0$ =1234.385 mm and $\hat{y}_0$ =-513.925 mm; in the absolute coordinate system, the estimated values are $\hat{x}_{C,G}$ =2413.782 mm and $\hat{y}_{C,G}$ =-20.650 mm.
Using the least squares method, the estimated coordinates of the center of gravity in the suspension point coordinate system are $\hat{x}_0$ =1234.192 mm and $\hat{y}_0$ =-515.465 mm; in the absolute coordinate system, the estimated values are $\hat{x}_{C,G}$ =2413.850 mm and $\hat{y}_{C,G}$ =-22.200 mm.
After data correction, comparing the estimated values from the two calculation methods, the horizontal coordinate difference is 0.02% MAC, and the vertical coordinate difference is 0.14% MAC.
As shown in Table 3, the following are the calculated results of the center of gravity estimated values before and after data correction:

Table 3 Calculation Results of Center of Gravity Estimates Before and After Data Correction

	Coordinate	Fermat's method	Least squares method	Difference
Before correction	$\hat{x}_{C,G}$(mm)	2413.913	2412.494	0.13%MAC
	$\hat{y}_{C,G}$(mm)	-67.594	-53.249	1.31%MAC
After correction	$\hat{x}_{C,G}$(mm)	2413.782	2413.850	0.02%MAC
	$\hat{y}_{C,G}$(mm)	-20.650	-22.200	0.14%MAC

Comparing the data in Table 3, there is no significant difference between the two methods in terms of the horizontal coordinate values, but there is a large difference in the vertical coordinate values. At the same time, it was found that the difference between the estimated values calculated based on the corrected data is smaller than the difference between the coordinate estimates calculated based on the original data. That is, based on the corrected data, the two different methods produce better agreement in terms of the estimated center of gravity position. In addition, the variances of the sample point coordinates before and after correction were calculated separately, and the results are shown in Table 4:

Table 4 Variance of Sample Point Coordinates Before and After Data Correction

	Horizontal coordinate variance $\sigma^2(x)(\mathrm{mm}^2)$	Vertical coordinate variance $\sigma^2(y)(\mathrm{mm}^2)$
Before correction	969.047	32733.618
After correction	231.718	7547.947

As can be seen from Table 4, the variance of the x and y coordinates of the corrected sample points has been optimized compared to the uncorrected values. The improved agreement between the two methods and the reduced variance of the sample point coordinates indicate that the correction method proposed in this paper has improved data quality and computational accuracy.

5 CONCLUSION

This paper is based on the two-point suspension method. At different pitch angles of the UAV, multiple sets of data are measured and recorded, then corrected. After converting the coordinate system, the torque equation is derived and solved. The estimated values of the UAV's center of gravity axial position and height position can be calculated using the Fermat method or the least squares method. The results show that the error in the center of gravity estimates obtained using different methods can be as low as 0.1% MAC.
With the increasing civilian development of China's maritime areas and changes in the maritime situation in recent years, the prospects for surface UAVs are becoming increasingly broad, and production will also grow accordingly. Weight and center of gravity measurement is a necessary process before and after the design, manufacturing, use, and maintenance of surface UAVs. The weight and center of gravity measurement method proposed in this paper has a clear and straightforward principle, a simple calculation process, easy weighing operations, low experimental costs, and high calculation accuracy. It has significant practical value and guiding significance for the weight and center of gravity measurement work of water based UAVs.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Wang Gang, Hu Yu, Song Bifeng. Optimal design of tailless UAV considering uncertainty of center of gravity position. Acta Aeronautica Sinica, 2015, 36(07): 2214-2224.
[2] Nursultan I, Sumil S, Dogan K, et al. Design, development and experimental validation of a lightweight dual-arm aerial manipulator with a COG balancing mechanism. Mechatronics, 2022(82): 102719.
[3] Feng Dongyun. Application of aircraft weighing equipment in aviation maintenance. Measurement and Testing, 2011(Supplement): 250-251.
[4] Jiang Dianxiang, Liu Zhuo, Lin Ziping. Research on the measurement method of aircraft weight and center of gravity. Trainer Aircraft, 2011(3).
[5] Raymer D. Aircraft Design,A Conceptual Approach. AIAA Education, 2014, 9(1): 78.
[6] Darvishpoor S, Roshanian J, Raissi A, et al. Configurations, flight mechanisms, and applications of unmanned aerial systems: A review. Progress in Aerospace Sciences, 2020(121): 100694.
[7] Aakif M, Asif I, Ahmad Y. Application specific drone simulators: Recent advances and challenges. Simulation Modelling Practice and Theory, 2019(94): 100-117.
[8] Huang Wei. A brief discussion on the impact of aircraft weight balance on flight safety. China Science and Technology Expo, 2015(037): 343.
[9] Li Wenqiang, Zhang Zhengji. Current status and development trend of aircraft weight and center of gravity measurement technology. Weighing Instrument, 2016, 45(4): 8-11.
[10] Dai Shilong. Aircraft weight and center of gravity measurement. Heilongjiang Science and Technology Information, 2013(7): 95.
[11] Liu Bin. Application and development of aircraft weight and center of gravity measurement technology. Aviation Maintenance and Engineering, 2004(1): 53-57.
[12] Jing S, Guangtong X, Zhu W, et al. Safe flight corridor constrained sequential convex programming for efficient trajectory generation of fixed-wing UAVs. Chinese Journal of Aeronautics, 2025(38): 103174.
[13] Lingling C, Qi L, Feng G, et al. Design, modeling, and control of morphing aircraft: A review. Chinese Journal of Aeronautics, 2022(35): 220-246.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3033

ENGINEERING ESTIMATION METHOD FOR HELIUM VOLUME IN TETHERED AEROSTATS

YuFan Xie
AVIC Special Vehicle Research Institute, Jingmen 448001, Hubei, China.
Corresponding Email: 815199892@qq.com

Abstract: Addressing the engineering need for precise estimation of helium volume within the envelope of tethered aerostats during station-keeping, and to overcome the increased errors of traditional methods in complex environments, a high-precision and practical engineering estimation method is proposed through the construction of a multi-factor coupling model. Firstly, for single influencing factors, two helium volume estimation methods are presented: one based on ground temperature variations and another on atmospheric parameter changes. Further, considering multiple influencing factors during aerial tethering, a multi-factor coupling computational model is developed. This involves simplifying the force model of the tethered aerostat and establishing force balance equations to correlate helium volume with measurable parameters such as wind speed and temperature. Additionally, a wind speed-tension fitting relationship is proposed to correct wind disturbance errors. Validation against field test data demonstrates that the estimation error of this helium volume method is within 5%. The proposed engineering estimation method provides reliable technical support for managing envelope helium volume during the long-term station-keeping of tethered aerostats, exhibiting significant practical application value.
Keywords: Helium volume; Multi-factor coupling; Engineering estimation; Experimental verification

1 INTRODUCTION

Tethered aerostats are lighter-than-air vehicles constrained by a tether, typically comprising an envelope, a tether cable, and ground-based mooring facilities[1][2]. They achieve lift by being filled with a buoyant gas (lighter than ambient air). Lacking a propulsion system, they remain aloft via a tether connected to ground facilities[3]. With technological advancements, helium has largely replaced hydrogen as the buoyant gas, extending station-keeping duration, enhancing payload capacity and economic efficiency. Consequently, tethered aerostats are increasingly deployed for applications such as Earth observation, early warning and detection, communication relay, disaster prevention and mitigation, environmental monitoring, and network coverage in both civil and military domains, becoming a focus of development for many nations[4][6].

The envelope, a critical component, houses the helium whose volume determines the aerostat's buoyancy performance. During station-keeping, the helium volume continuously fluctuates due to the influence of surrounding atmospheric temperature, pressure, and other factors. Furthermore, as deployment scenarios expand[7][9]—spanning plains, plateaus, oceans, and diverse latitudes, longitudes, and altitudes—accurately estimating the in-envelope helium volume based on actual measurement parameters becomes crucial for optimal operational performance. This study investigates helium volume estimation methods, providing a reference for engineering applications.

2 HELIUM VOLUME ESTIMATION METHODS BASED ON SINGLE FACTOR ANALYSIS

2.1 Method Based on Ground Temperature Variation

Assuming constant ground atmospheric pressure, the ideal gas law dictates that helium volume (V_{he}) is proportional to ground temperature (T). Thus, volume increases with rising temperature and decreases with falling temperature. This yields Equation 1:

$$V_{he} = V_{he}^{*} \frac{T}{T^{*}} \tag{1}$$

Where:

T^{*} is the atmospheric temperature at the initial ground state;

T is the atmospheric temperature at the real-time or future ground state;

V_{he}^{*} is the helium volume at the initial ground state;

V_{he} is the helium volume at the real-time or future ground state.

Given initial ground temperature and helium volume, Equation 1 estimates volume under temperature changes. This method aids in predicting if the ballonet volume reaches its limit during cooling, informing helium replenishment operations.

Comparisons between helium volumes calculated using this ground-temperature-based method and the ground test data, along with error curves, are shown in Figure 1 and Figure 2. The results indicate minor discrepancies, validating the method's high precision for ground-level helium volume estimation.

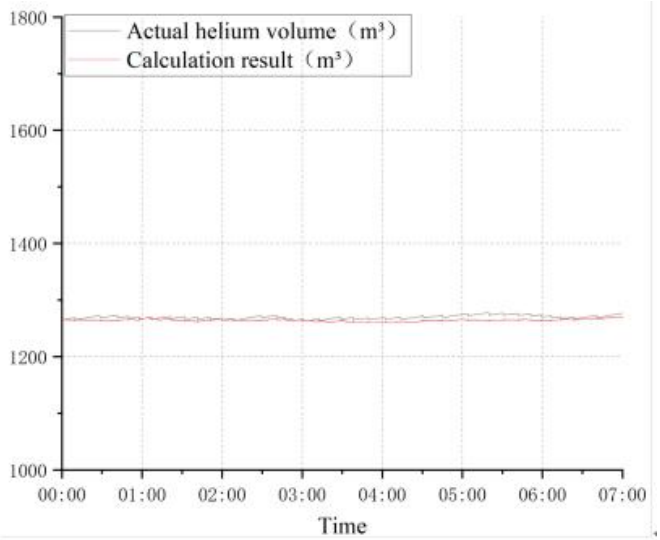

Figure 1 Comparison of Estimated Helium Volume Data and Test Data

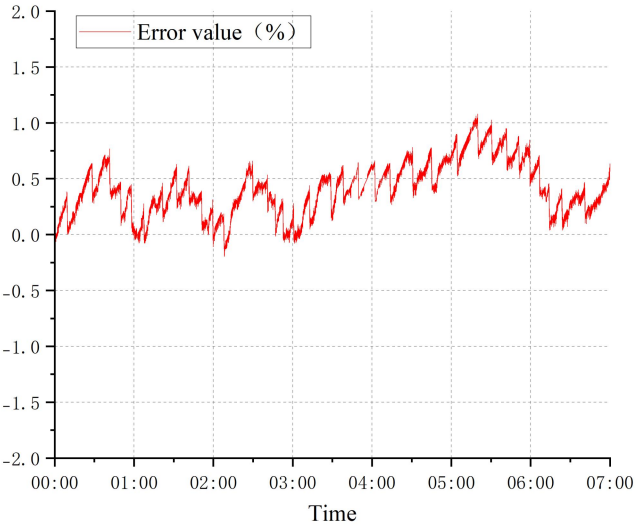

Figure 2 Helium Volume Estimation Error Curve

2.2 Method Based on Atmospheric Parameter Variation

During actual aerostat ascent and recovery, atmospheric temperature and pressure change significantly with altitude compared to the ground-moored state, substantially impacting helium volume. Based on the ideal gas law and standard atmospheric models, expressions for atmospheric temperature and pressure at different altitudes are given by Equation 2 and Equation 3. Considering the combined effect of pressure and temperature on volume, the expression for helium volume at different altitudes is derived as Equation 4:

$$T = T_0 + L \times H \tag{2}$$

$$P = P_0 \times (1 - 2.25577 \times 10^{-5} \times H)^{5.25588} \tag{3}$$

$$V_{he} = V_{he}^{*} \frac{P_0 T}{P T_0} \tag{4}$$

Where:
P is the atmospheric pressure at the current state;
T is the atmospheric temperature at the current state;
P_0 is the atmospheric pressure at the initial altitude state;
T_0 is the atmospheric temperature at the initial altitude state;
L is the temperature lapse rate, typically -0.0065 K/m;
H is the altitude.

2.3 Comparison and Analysis of the Two Estimation Methods

Both methods were used to estimate helium volume during actual ascent, and results were compared against test data. Method 1 used measured ground temperature data at different times with Equation 1. Method 2 used Equation 2 and Equation 3 to obtain atmospheric parameters at various ascent altitudes, then Equation 4 for volume estimation. Comparisons of estimated results against actual helium volumes derived from measured ballonet fullness, along with error curves, are shown in Figure 3 and Figure 4.

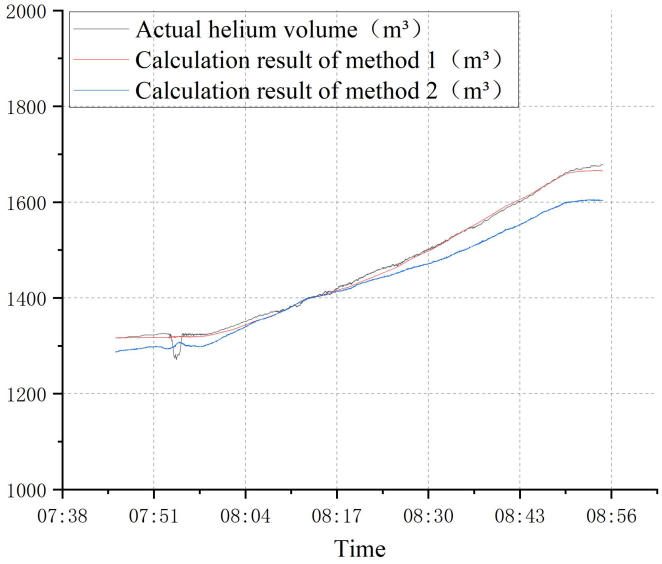

Figure 3 Comparison of Estimated Helium Volume Data and Test Data

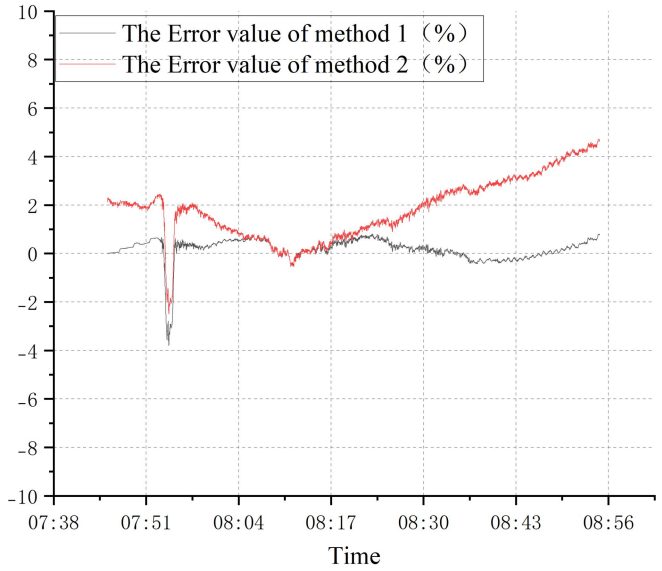

Figure 4 Comparison of Helium Volume Estimation Error Curves

Figure 3 and Figure 4 show that both methods capture the overall trend of helium volume change during ascent. However, estimation errors increase with altitude. Analysis suggests contributing factors include solar radiation heating the envelope interior and inaccuracies in atmospheric temperature data at different altitudes.

Furthermore,Figure 4 indicates Method 1 exhibits higher precision than Method 2 during ascent. Analysis of the atmospheric temperature comparison curve in Figure 5 reveals that Equation 2 predicts monotonically decreasing temperature with altitude. In reality, at a fixed location, ambient temperature increases diurnally while simultaneously decreasing with altitude gain during ascent. These opposing trends moderate actual envelope temperature changes. Method 1 uses measured ground temperature data, whose temporal trend aligns better with the actual conditions experienced during ascent, leading to smaller errors and higher precision.

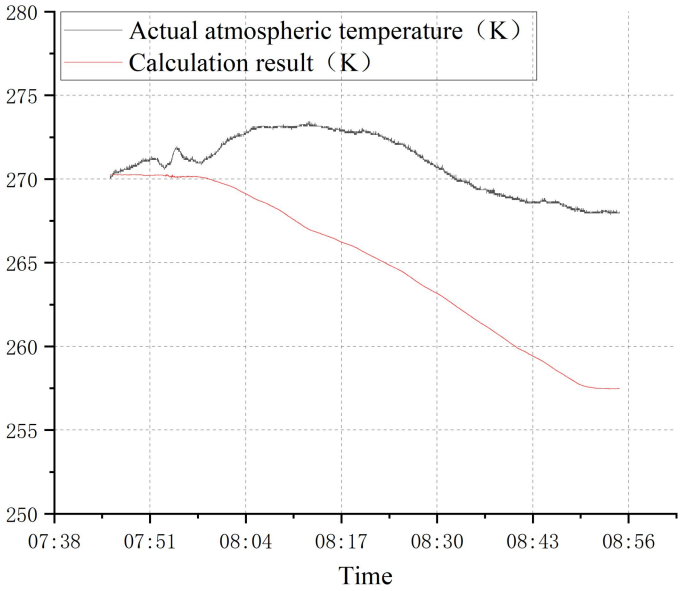

Figure 5 Comparison of Measured and Estimated Atmospheric Temperature Curves

The results and analysis above indicate that estimating helium volume during ascent requires considering both altitude-induced changes in atmospheric parameters and diurnal temperature variations. Estimating atmospheric state parameters under these multiple influences is challenging. Additionally, unknown helium leakage occurs during long-term station-keeping. Therefore, a helium volume estimation method based on multi-factor analysis is urgently needed.

3 HELIUM VOLUME ESTIMATION METHOD BASED ON MULTI-FACTOR ANALYSIS

3.1 Helium Volume Estimation Method Considering Multiple Factors

During station-keeping, environmental wind causes continuous changes in tether inclination angle, envelope pitch angle, and trajectory, resulting in a complex overall force model. Considering factors like tether angle, wind speed, pitch angle, and angle of attack, the force model is simplified.

The simplified force model is shown in Figure 6. In the tethered state, the envelope is in equilibrium under the action of: upward buoyant force from helium (F_b), aerodynamic lift from fins (F_L), structural weight (G_{struct}), and tension at the tether's upper end (T_1).

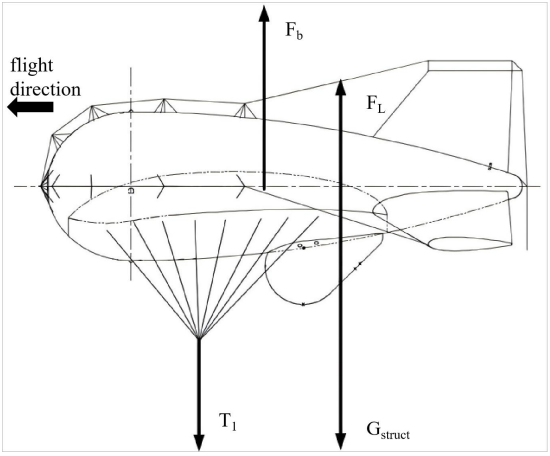

Figure 6 Simplified Force Model of Tethered Aerostat

Analyzing these forces: T_1 can be measured by onboard sensors; G_{struct} is known; the aerodynamic lift coefficient C_L is determined during design, enabling calculation of F_L. Therefore, the net buoyant force (F_{net}) generated primarily by the helium can be derived from the force balance Equation 5. As buoyancy is strongly related to helium volume, the volume during station-keeping can be solved.

The force balance equation is:

$$F_{net} = T_1 + G_{struct} - F_L \tag{5}$$

Where:
F_{net} represents the net buoyant force;
T_1 represents the tension at the upper end of the tether;
G_{struct} represents the structural weight;
F_L represents the aerodynamic lift force.
The net buoyant force F_net is derived through Equation 5 to Equation 9, resulting in Equation 10 and Equation 11:

$$F_{\text{net}} = F - G_{\text{air}} - G_{he} \tag{6}$$

$$F = \rho_{air} V_{\text{env}} \tag{7}$$

$$G_{\text{air}} = \rho_{air} V_{\text{he}} \tag{8}$$

$$G_{he} = \rho_{he} (V_{\text{气囊}} - V_{\text{side-env}}) \tag{9}$$

$$F_{\text{net}} = (\rho_{air} - \rho_{he})(V_{\text{env}} - V_{\text{side-env}}) \tag{10}$$

$$F_{\text{net}} = (\rho_{air} - \rho_{he}) V_{he} \tag{11}$$

T_1 is measurable. G_{struct} is known. F_L relates to wind speed (V_w) and the aerodynamic lift coefficient (C_L), expressed as:

$$F_L = \frac{1}{2} P v^2 \times V^{\frac{2}{3}} \times C_L \tag{12}$$

C_L depends on the angle of attack (α):

$$C_L = 0.029 \times (\alpha - 0.53) \tag{13}$$

Deriving from Equation 5 to Equation 13, the expression for helium volume V_{he} is:

$$\rho_{air}^{*} = 0.12492 \times \left(1 - 0.0000225577 \times H\right)^{4.25588} \times 9.80665 \tag{14}$$

$$\rho_{air} = \rho_{air}^{*} \times \frac{T^*}{T} \tag{15}$$

$$\rho_{he}^{*} = \frac{c}{n_{he}} \times 4 + \frac{(1-c)}{n_{air}} \times 29 \tag{16}$$

$$\rho_{he} = \rho_{he}^{*} \times \frac{T^*}{T} \tag{17}$$

$$V_{he} = \frac{T_1 + G_{\text{结构}} - \frac{1}{2} P v^2 \times V_{\text{env}}^{\frac{2}{3}} \times 0.029 \times (\alpha - 0.53)}{(\rho_{air}^{*} - \frac{c}{n_{he}} \times 4 + \frac{(1-c)}{n_{air}} \times 29) \times \frac{T^*}{T}} \tag{18}$$

Where:
V_{env} represents the envelope volume;
ρ_{air}^{*} represents the standard atmospheric density;
ρ_{air} represents the actual atmospheric density;
ρ_{he}^{*} represents the helium density at the ground-moored state;
ρ_{he} represents the helium density at the actual temperature;
T^* represents the ground temperature;
T represents the actual temperature;
c represents the measured helium purity on the ground;
n_{he}、n_{air} represent the molar masses of helium and air, respectively.

This estimation method depends on factors such as helium purity, wind speed, upper tether tension, and atmospheric pressure. During the aerial phase, upon measuring parameters like T_1, helium volume can be estimated using Equation 18.

3.2 Correction of the Estimation Method Based on Experimental Data

Figure 7 to Figure 11 show that during station-keeping, the maximum environmental wind speed reached 10.7 m/s. As per the derivation, wind speed affects aerodynamic lift (F_L), and the upper tether tension (T_1) fluctuates significantly with wind speed. In reality, the balloon and tether undergo coupled motion; the tether constrains the balloon, so T_1 increases with wind speed during gusts. Consequently, errors in T_1 measurement due to wind effects cause significant deviations in the estimated helium volume. Measuring tether state parameters during flight is difficult. Therefore, establishing a relationship between wind speed and T_1 is proposed to reduce errors induced by wind-affected tension measurements and improve estimation accuracy.

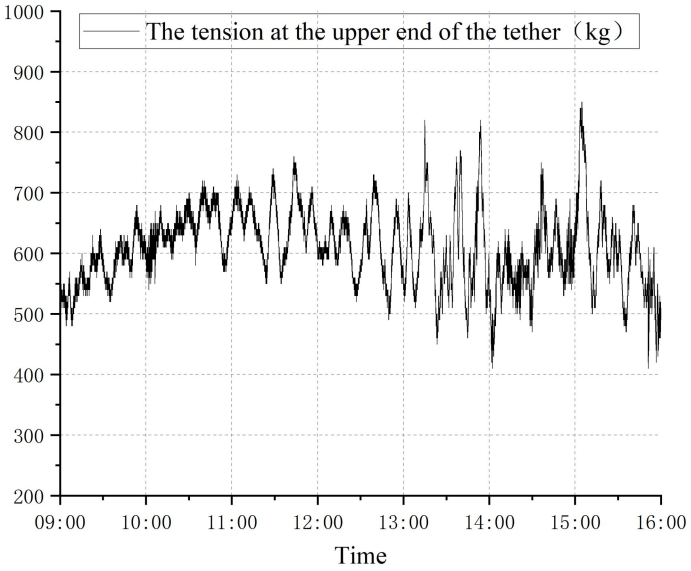

Figure 7 Measured Variation of Upper Tension during Station-Keeping

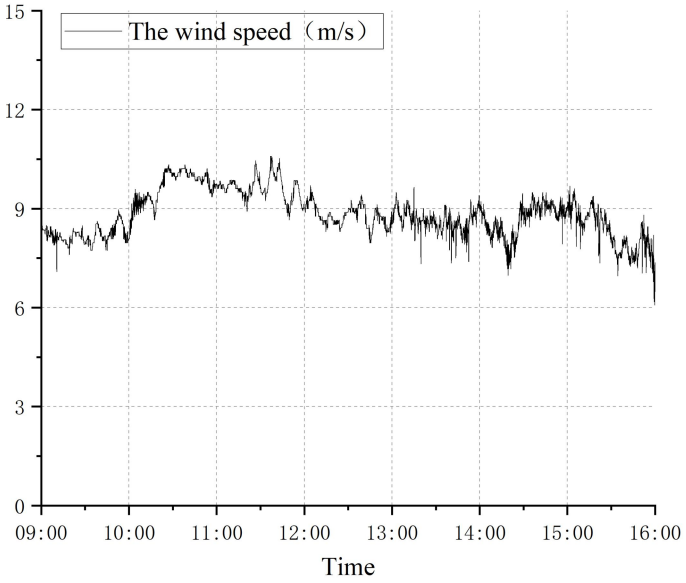

Figure 8 Measured Variation of Wind Speed during Station-Keeping

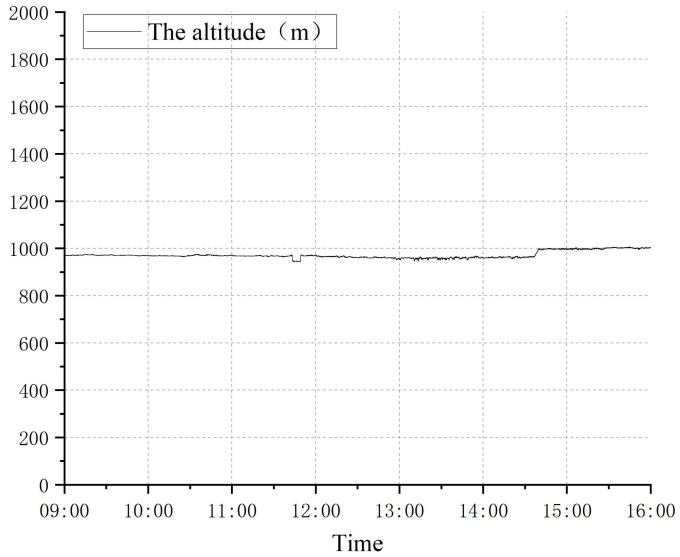

Figure 9 Measured Variation of Altitude during Station-Keeping

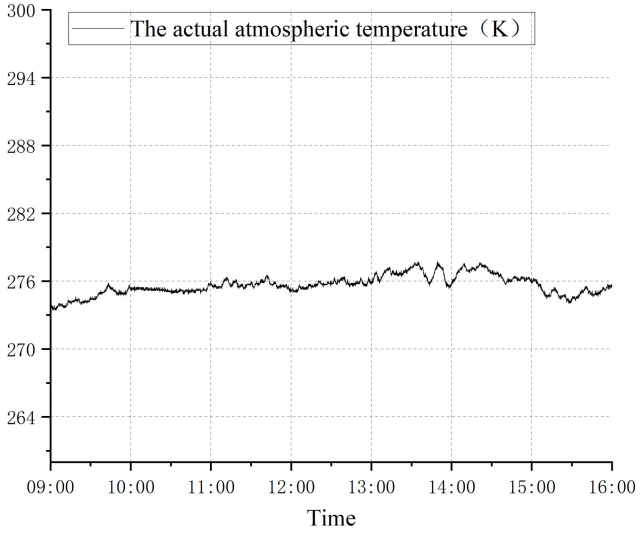

Figure 10 Measured Variation of Atmospheric Temperature during Station-Keeping

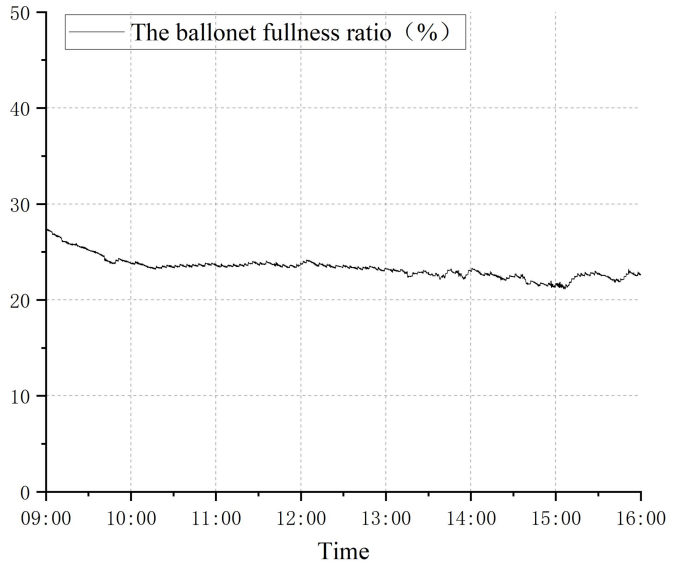

Figure 11 Measured Variation of Ballonet Fullness Ratio during Station-Keeping

A wind speed-upper tension relationship is proposed for rapid estimation:

$$T_1 = \eta v + T_1^* \tag{19}$$

Where:

η represents the rate of change of upper tension with wind speed;

T_1^* represents the upper tension under windless conditions.

Fitting the equation to measured data yielded:

η ≈ 8 (units: force/wind speed, e.g., N/(m/s)

T_1^* ≈ 2022.5 kg (force units implied)

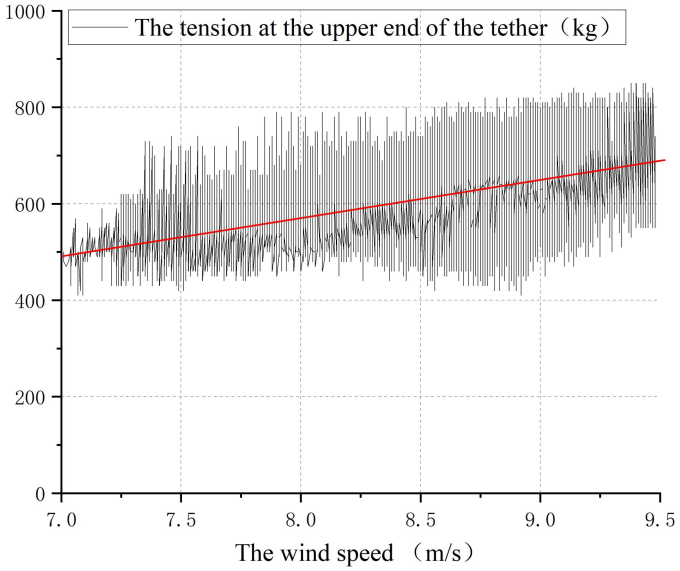

Figure 12 Wind Speed vs. Upper Tension Fitting Equation from Station-Keeping Data

Substituting Equation 19 into Equation 18 gives the corrected estimation formula:

$$V_{he}=\frac{(T_1-\eta v)+G_{struct}-\frac{1}{2}Pv^2\times V_{env}^{\frac{2}{3}}\times 0.029\times(\alpha-0.53)}{(\rho_{air}^{*}-\frac{c}{n_{he}}\times 4+\frac{(1-c)}{n_{air}}\times 29)\times\frac{T^{*}}{T}} \tag{20}$$

The estimated helium volumes using the corrected formula, compared against actual ascent test data, are shown in Figure 13 and Figure 14. The figures show that the estimated values fluctuate within the range of the measured helium volumes, consistent with test data variability. The estimation error remains within 5%, demonstrating that this multi-factor helium volume estimation method achieves high precision for tethered aerostats during station-keeping.

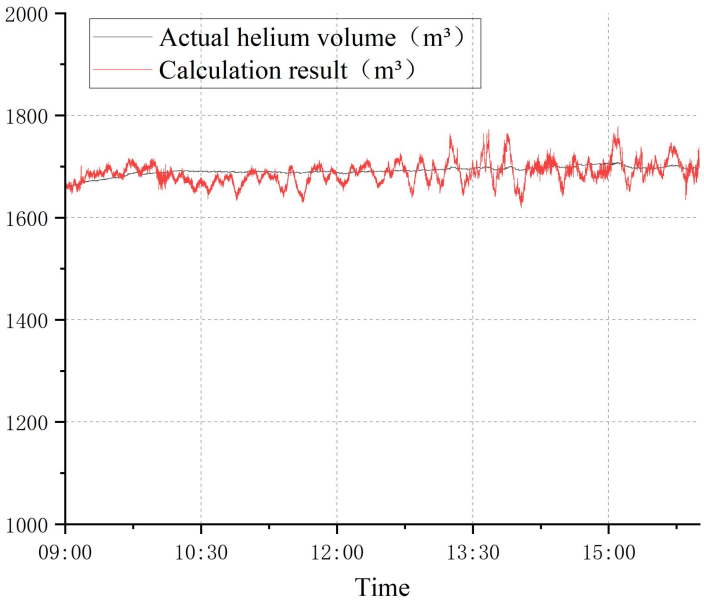

Figure 13 Comparison of Corrected Helium Volume Estimates and Test Values

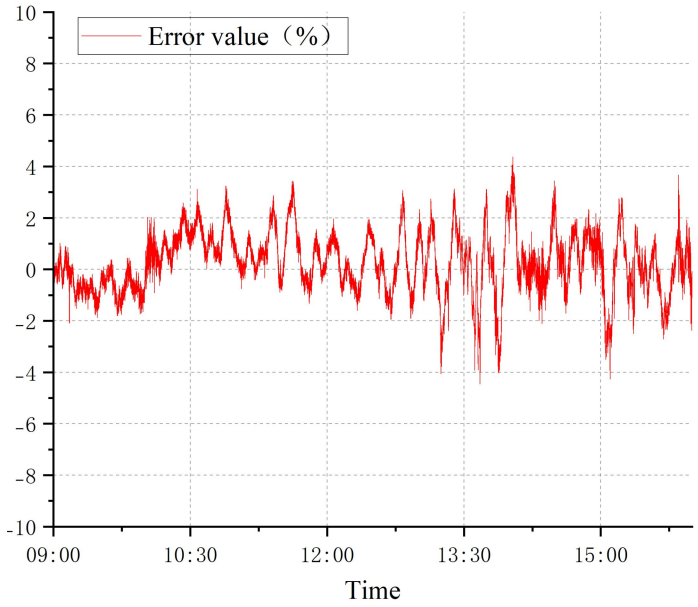

Figure 14 Magnitude of Error between Corrected Estimates and Test Values

4 CONCLUSION

This study systematically investigates engineering methods to meet the requirement for precise estimation of envelope helium volume in tethered aerostats during station-keeping. The main conclusions are as follows:

1.Single-factor methods (ground temperature and atmospheric parameters) were investigated. The ground-temperature-based method proved highly accurate under ground-moored conditions, suitable for predicting ballonet volume limits during cooling to inform helium replenishment. The atmospheric-parameter-based method reflected volume trends during ascent, but its error increased significantly with altitude, primarily limited by unaccounted factors like solar radiation heating and deviations in measured atmospheric parameters. Both methods exhibit limited applicability in the complex dynamic environment of aerial tethering.

2.A high-precision helium volume estimation method suitable for aerial tethering was proposed. This method correlates helium volume with real-time measurable key parameters (upper tether tension, wind speed, atmospheric temperature, etc.) by simplifying the aerostat's force model and establishing force balance equations. Specifically, a wind speed-tension fitting relationship was introduced to correct measurement errors in tension caused by wind disturbances. Model correction and validation using field test data demonstrated significantly improved accuracy. The corrected multi-factor coupling estimation method achieves errors within 5%, effectively overcoming the precision limitations of single-factor methods in complex environments.

3.The proposed high-precision, practical multi-factor coupling helium volume estimation method provides reliable technical support for real-time monitoring and management of envelope helium volume during long-term station-keeping missions of tethered aerostats. It significantly enhances estimation accuracy and engineering applicability in complex environments, holding substantial practical value. Future work could integrate intelligent sensor networks and adaptive algorithms to optimize robustness and real-time performance under extreme or transient meteorological conditions.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Jiang Y W. The Never-Vanishing "Sky Giant": The Operational Use of Russian Military Tethered Balloons. Ordnance Knowledge, 2017(12): 39-43.

[2] Jiang Y W. The Re-Emergence of Russia's Military Tethered Balloon System. Ordnance Knowledge, 2018(01): 38-42.

[3] Zhang D, Luo H, Cui Y, et al. Tandem, long-duration, ultra-high-altitude tethered balloon and its system characteristics. Advances in Space Research, 2020, 66(10): 2446-2465.

[4] Dasaradhan B, Das B R, Sinha M K, et al. A brief review of technology and materials for aerostat application. Asian Journal of Textile, 2018, 8: 1-12.

[5] Hembree B, Slegers N. Tethered aerostat modeling using an efficient recursive rigid-body dynamics approach. Journal of Aircraft, 2011, 48(2): 623-632.

[6] Deng X, L Ma Z Y, Luo X Y. Development and Application of Foreign Tethered Balloon Equipment and Its Enlightenment. Aerodynamic Missile Journal, 2020(06): 76–82.

[7] Kobayashi F. Direct sampling and bioanalyses of atmospheric bioaerosols using a tethered balloon over Syowa Station, Antarctica. Polar Science, 2022, 32: 100842.
[8] Cappelletti D, Petroselli C, Mateos D, et al. Vertical profiles of black carbon and nanoparticles pollutants measured by a tethered balloon in Longyearbyen (Svalbard islands). Atmospheric Environment, 2022, 290: 119373.
[9] Dai Q, Xing D, Fang X, et al. Numerical research on the thermal performance of high altitude scientific balloons. Applied Thermal Engineering, 2017, 114: 51-57.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3034

AE MONITORING OF REINFORCED CONCRETE STRUCTURES

JunTao Fan[1], Xu Li[2*]
[1]*CITIC Construction Co., LTD, Beijing 100027, China.*
[2]*Beijing Polytechnic University, Beijing 100176, China.*
Corresponding Author Xu Li, E-mail:lx3818262@163.com

Abstract: In this study, a novel damage evaluation method for concrete structures is introduced, which is based on the rising rate of acoustic emission. This method leverages the fact that when a concrete structure is subjected to stress, acoustic emissions increase, and the rate of this increase can be correlated with the extent of damage. To validate the effectiveness of this proposed method, a seismic damage test was conducted on a frame shear wall structure, which is a common and critical component in many building constructions. The results of the test clearly demonstrated that this method can accurately evaluate the damage status of the structure. More importantly, it has the capability to effectively locate areas where significant damage has occurred. This makes it a potentially valuable tool for post-disaster assessment and structural health monitoring, as it can help engineers and inspectors quickly identify and prioritize areas that require repair or further investigation. Overall, this study presents a promising approach to enhance the safety and reliability of concrete structures in the face of seismic events.
Keywords: Concrete; Acoustic emission; Accumulative time parameter; Seismic damage; Power spectral density

1 INTRODUCTION

Acoustic emission describes the occurrence in which materials emit elastic waves when undergoing fracture. The AE technique is capable of detecting the formation and development of microcracks within a structure. Structural damage is primarily attributed to fracturing processes[1]. Therefore, the AE technology is widely applied in the non-destructive testing and damage assessment of reinforced concrete structures. Currently, many scholars have proposed AE assessment methods specifically for reinforced concrete structures. Ohtsu[2], JI proposed the AE rate theory, which is used to evaluate the compressive strength of actual concrete structures[3]. Based on the Kaiser effect of concrete, the Japanese Society for Non-Destructive Testing (JSNDI) has proposed a NDIS-2421 standard[4], which is used to assess the damage condition of reinforced concrete structures[5]. In addition, there is another method derived from seismology called the b-value theory[6], which is used to predict the entire process from the formation of micro-cracks to the macroscopic expansion of cracks in concrete. Carpinteri extended the b-value theory and proposed a damage assessment method based on the cumulative rate of acoustic emission events[7], which assesses the damage of concrete structures according to the growth rate of acoustic emission events. Liu proposed a new b-value estimation procedure and applied it to the expansion and rupture test[8]. The results confirmed that the b-value depends on the material's in-homogeneity and stress.

Earthquake disasters are often one of the main factors causing damage to concrete structures. The aforementioned method has not yet been applied to the assessment of damage in concrete structures that have been damaged by earthquakes. This paper proposes a method for assessing the damage of concrete structures based on the cumulative growth rate of acoustic emission events, and validates it using an earthquake damage test of a concrete frame-shear structure.

2 THE PRINCIPLE OF CUMULATIVE TIME PARAMETER IN ACOUSTIC EMISSION

In concrete structures, the pulse wave released instantaneously upon the formation of a tiny crack is called an acoustic emission event. The commonly used acoustic emission characteristic parameters include the number of acoustic emission events, the acoustic emission rate, the cumulative acoustic emission energy, the number of acoustic emission oscillations, the duration, and the rise time.

Among the methods for assessing damage in concrete structures using acoustic emission event statistics, the b-value theory is the most representative. The *b*-value theory is proposed based on the following fact: for low-frequency acoustic emission events, their amplitudes are usually relatively high, while for high-frequency acoustic emission events, the amplitudes are generally low. Therefore, Richter calculated the slope of the amplitude distribution, that is, the *b*-value, to statistically analyze the amplitude distribution pattern of acoustic emission events. The *b*-value is shown as following

$$\log_{10} M = a - b(A_{dB} / 20) \tag{1}$$

In which, M represents the cumulative number of AE events where the peak amplitude of the signal (measured in dB) is greater than A_{dB}, a is a constant, and the value of b indicates the slope of the distribution of AE events with different amplitudes.

In the *b*-value theoretical analysis, both parameters *a* and *b* are time-dependent parameters. Therefore, the damage state of the structure is expressed as the rate of increase in the number of AE events at a certain level. Equation (1) is re-expressed in terms of the amplitude A_0 in volts, and a cumulative time parameter β_t for acoustic emission is introduced as follows[8].

$$N(\geq A_0, t) = 10^{a(t)} A_0^{-b(t)} = t^{\beta_t} \tag{2}$$

In the formula, $N(\geq A_0, t)$ represents the cumulative number of AE events with an amplitude greater than A_0. β_t reflects the rate of damage growth. As shown in Figure 1, $\beta_t = 1$ serves as a critical value for damage: when $\beta_t < 1$, the growth rate of N is slow; when $\beta_t > 1$, N increases sharply, the structural damage becomes severe, the bearing capacity is lost, and the structure is in a state close to collapse. According to reference[7], the value of *b* can be approximately calculated using the number of acoustic emission ringing events N^*, so the time parameter β_t of the amplitude distribution can be approximately calculated using the acoustic emission ringing number.

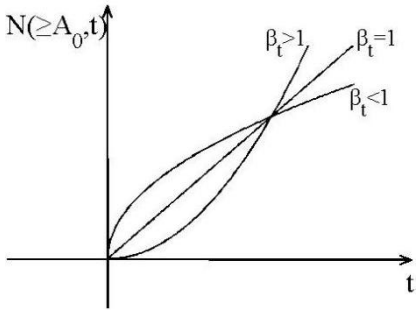

Figure 1 The Illustration of β_t with Three Conditions

3 ACOUSTIC EMISSION MONITORING EXPERIMENT OF CONCRETE STRUCTURES

This paper conducted an AE monitoring test on the seismic damage of a scaled-down reinforced concrete model to verify the feasibility of using the cumulative time parameter of acoustic emission for the assessment of seismic damage in concrete structures. The model was a 1/5-scale two-way double-span three-story reinforced concrete eccentric frame-shear wall structure, as shown in Figure 2. The tensile reinforcement in the model was made of 3mm galvanized iron wire, and 0.9mm galvanized iron wire was used as the stirrups. The reinforcing bars in the shear walls and slabs were double-layer galvanized iron wire mesh with a diameter of 2mm and a spacing of 20mm. The structure was poured with micro-particle concrete. The foundation base was cast with C30 concrete. The cross-sectional dimensions of the columns were 80mm×80mm, the cross-section of the beams was 50mm×100mm, and the thickness of the slabs was 30mm.

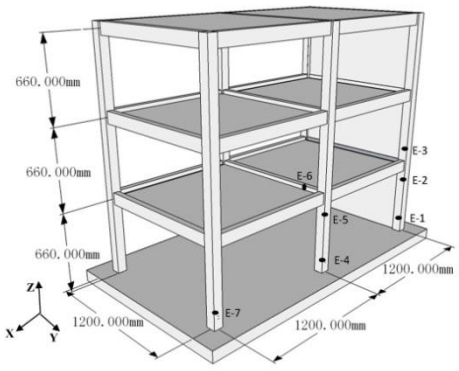

Figure 2 The Dimensions of Model and the Distribution of PZT Sensors

The sensor layout is shown in Figure 2, with numbers E-1 to E-7. The AE sensor signals were collected using the dSPACE system, with a sampling frequency of 10 kHz, and the acceleration in both the horizontal and vertical directions of each layer was simultaneously measured. Figure 3 is a photo of the test model.

Figure 3 The Experimental Photo

Simultaneous excitation in the horizontal direction (Y-axis) and the vertical direction (Z-axis) was carried out. The north-south component and the vertical component of the El-Centro wave in 1940 were used as the seismic input for this experiment. The input seismic wave conditions are shown in Table 1. A total of 13 sets of conditions were input, and each condition had a different amplitude of the seismic wave.

Table 1 The Earthquake Waves List

earthquake wave	EC-1	EC-2	EC-3	EC-4	EC-5	EC-6	EC-8	EC-9	EC-10	EC-11	EC-12	EC-13
peak value（g）	0.16	0.24	0.33	0.4	0.45	0.5	0.58	0.6	0.66	0.7	0.73	0.86

The data of acoustic emission and velocity sensors were simultaneously collected, and the cumulative ringing number of the acoustic emission signal was calculated. The specific method was as follows: set the threshold voltage to 5 mV. When the waveform of the acoustic emission signal passed the threshold voltage during the descending phase, the ringing number was recorded.

The loading time and the cumulative acoustic emission ringing number were normalized, and using equation (2) as the fitting target, the acoustic emission amplitude cumulative time parameter β_t was calculated using the least squares method. Figure 4 shows the normalized *t*-N* curve during the loading process of EC-11 and the β_t calculated by the least squares method. Among them, R-square is the determination coefficient of the fitting equation.

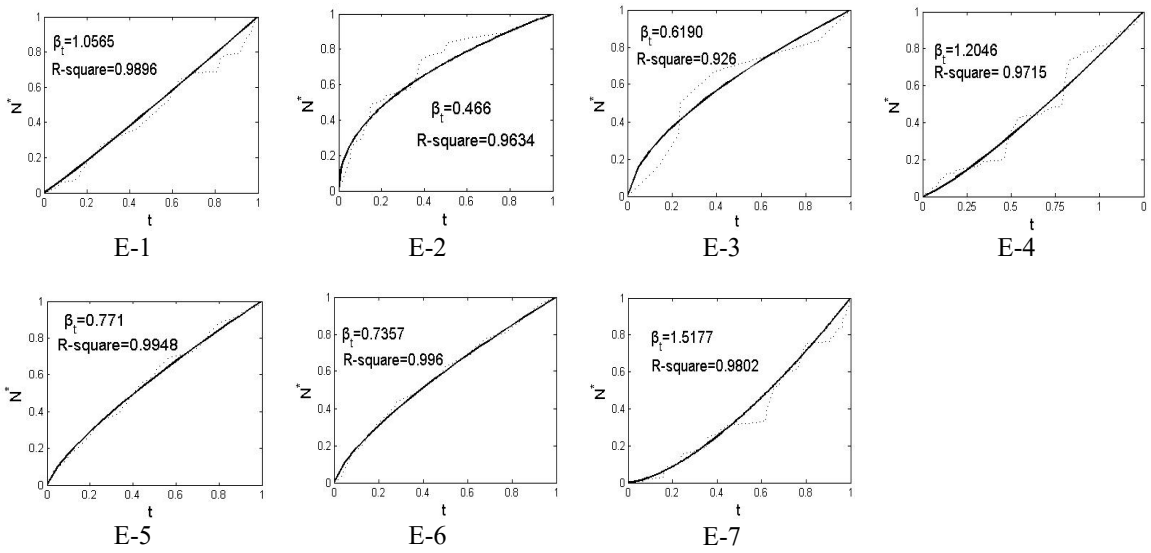

Figure 4 The *t-N** Curve and β_t during EC-11 Loading

It can be seen that the β_t calculated from the acoustic emission signals of E-7, E-4, and E-1 is greater than 1, indicating that during this loading process, the acoustic emission phenomena in the area near these three sensors have increased sharply, and the damage is relatively severe. Figure 5 shows the cumulative ringing numbers of the acoustic emission signals from E-1 to E-7 after all seismic wave loading is completed. Comparing the sensor signals at the bottom of the structure, the cumulative ringing number of E-7 is the highest, followed by E-4, and E-1 has a relatively smaller number of ringings. This is consistent with the results in Figure 4. It indicates that the bottom of the side columns has the most severe damage, followed by the middle columns. In contrast, the bottom of the column connected to the shear wall has

less damage. This is because the stiffness distribution along the X direction of the structure is uneven. The stiffness on the shear wall side is larger while that on the side of the side columns is smaller, resulting in the deformation of the side columns being greater than that of the shear wall. Therefore, the damage along the X direction becomes increasingly severe. Figure 6 shows the damage photos near E-1, E-4, and E-7. From the figure, it can also be seen that there are unpenetrating cracks near E-1, penetrating cracks at the bottom of E-4, and more than two penetrating cracks at E-7. The results judged from the photos are consistent with the analysis.

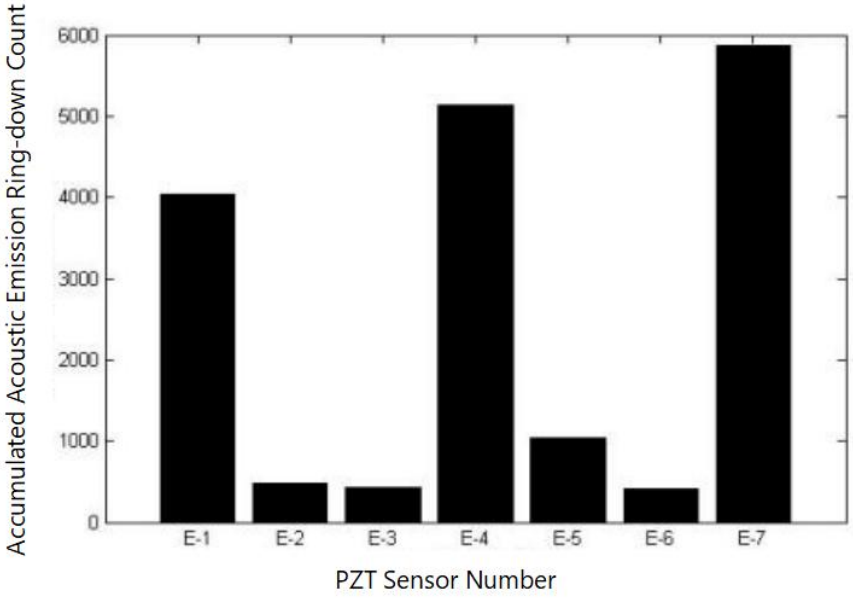

Figure 5 The Total Acoustic Emission Count from E-1to E-7

Near E-1 Near E-4 Near E-7

Figure 6 The Damage Photo Near E-1, E-4 and E-7

4 CONCLUSION

This paper presents a damage assessment method for concrete structures under seismic loads based on the cumulative time parameter of acoustic emission, and the feasibility of the method is verified through an earthquake damage experiment of a concrete frame-shear structure. The results show that this method can effectively evaluate the damage degree of concrete structures, and by using an acoustic emission sensor array, it can effectively detect the damaged areas with larger damage.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Liu Jiahao, Chen Jiang, Xue Yang, et al. Review of the application of acoustic emission technology in green manufacturing. International Journal of Precision Engineering and Manufacturing-Green Technology, 2024, 11(3): 995-1016.

[2] Masayasu Ohtsu, Hiroshi Watanabe. Quantitative damage estimation of concrete by acoustic emission. Construction and Building Materials, 2001, 15(5): 217-224.

[3] JI Hongguang, ZHANG Tiansen, CAI Meisen, et al. Experimental study on concrete damage by dynamic measurement of acoustic emission. Chinese Journal of Rock Mechanics and Engineering, 2000, 19(2):165-168.

[4] Masayasu Ohtsu, Masakatsu Uchida, Takahisa Okamoto. Damage Assessment of reinforcement concrete beams qualified by acoustic emission. ACI Structrual Journel, 2002, 99(4): 411-417.
[5] LI Dongsheng, CAO Hai. Acoustic emission monitoring and evaluation of damage evolution for polyvinyl alcohol fiber concrete. Journal of vibration and shock, 2012, 31(9): 29-32.
[6] Richter CF. Elementary seismology. Freeman, NewYork. 1958.
[7] Carpinteri A, Lacidogna G, Niccolini G, et al. Critical defect size distributions in concrete structures detected by the acoustic emission technique. Meccanica, 2008, 43(3): 349-363.
[8] Liu X, Han M, He W, et al. A new b value estimation method in rock acoustic emission testing. Journal of Geophysical Research: Solid Earth, 2020, 125(12): e2020JB019658.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI:https://doi.org/10.61784/wjer3035

APPLICATION OF REINFORCEMENT LEARNING IN AUTONOMOUS DRIVING SCENARIOS: PATH PLANNING USING POLICY GRADIENT METHODS

JunRan Wu
School of Mathematical, Chengdu University of Technology, Yibin 644000, Sichuan, China.
Corresponding Email: Wjr13031955011@foxmail.com

Abstract: This study explores reinforcement learning-based autonomous path planning in China's mixed traffic flow, employing a policy gradient approach to optimize decision-making through environmental interaction. China's traffic environment, characterized by a diverse mix of vehicles, pedestrians, and non-motorized traffic, presents unique challenges. To address these, a state space is meticulously constructed, incorporating real-time traffic light status, the precise positions and orientations of surrounding vehicles, providing a comprehensive representation of the traffic scenario. The action space, covering operations such as going straight, turning left, turning right, and stopping, enables the autonomous system to make practical driving decisions. A "time-saving and violation-avoidance" reward function is designed, effectively transforming the path planning task into a sequential decision optimization problem. The policy network, parameterized by a three-layer neural network, learns from the environment through repeated trials. Initial reward fluctuations gradually stabilize, achieving a 100% success rate with an average of 18.2 steps, successfully realizing shortest-path selection. However, the negative average reward (-1.22) indicates that the current reward function may have an excessive penalty bias. While validating the effectiveness of the policy gradient method for intelligent transportation, the results underscore the necessity of refining the reward function to strike a balance between negative penalties and positive incentives. This framework provides valuable methodological guidance for autonomous driving in complex scenarios, yet further optimization of the reward mechanisms remains essential for practical implementation.
Keywords: Autonomous driving; Policy gradient methods; Reward function; Policy network

1 INTRODUCTION

We have now entered an era surrounded by intelligent agents. Intelligent agents can be found in various fields, such as the AI field, the autonomous driving field, the geographical simulation field, and the economic calculation field. The arrival of intelligent agents has brought many new experiences to our lives. Piero even once said, "Machines are the key to the happiness of our future lives,... Machine intelligence is likely to be indispensable for solving the most serious problems of our time[1]. ” So, how can we use machines as a carrier to simulate a certain aspect of human intelligence? In Wu Fei's work, it was introduced that "To simulate a certain aspect of human intelligence using machines as a carrier, it can be achieved through methods such as logical reasoning centered on symbolism, search exploration centered on problem-solving, and decision-making intelligence centered on data-driven approaches." Inspired by the laws of biological learning, reinforcement learning interacts with the environment through a trial-and-error mechanism. It learns and optimizes by maximizing the cumulative rewards, and ultimately achieves the optimal strategy[2]. Artificial intelligence aims to enable machines to learn, think, and understand just like humans, that is, to simulate human intelligence using computers. Since the external environment provides very little information directly, the decision-making agent in reinforcement learning must interact with the environment through trial and error, and continuously optimize the control strategy from the action-evaluation process to improve the control performance of the system. Therefore, essentially, reinforcement learning approximates the mapping relationship of "state-action" through parameterized functions to find the optimal strategy for solving decision-making problems[3]. We hope that machines can learn autonomously just like humans. What we need to do is to enable the intelligent agent to maximize the cumulative rewards obtained from its responses during the interaction with the environment. This reward is abstract. It can be numerical addition or subtraction, time addition or subtraction, or other forms. Since the intelligent agent may face multiple states, we can simplify this problem to the issue of an autonomous vehicle passing through a traffic light intersection. Then, the final cumulative return can be regarded as the time required for the autonomous vehicle to complete this section of the journey.

Typical reinforcement learning has the following three characteristics[4]:

1) Different actions generate different rewards.
2) Rewards are delayed in time.
3) The return of a certain action is related to the specific environment in which it is performed.

Wahba proposed two methods related to machine learning models: supervised learning and unsupervised learning. Unsupervised learning does not require a decision-making process, and its learning objective is merely to obtain the distribution of data within the same category. Supervised learning provides data information with labels and can make

single-step decisions based on the label information. However, these two learning methods completely fail to meet the data processing and decision-making needs of intelligent agents in practice. In the actual environment, the state is changing. Since supervised learning requires the construction of a large number of labeled samples, it is difficult to be applied in practice. Currently, most studies use the deep reinforcement learning method to train the model . Reinforcement learning does not require a supervision signal to directly guide the learning. It only depends on a feedback reward signal to evaluate its "trial-and-error" process, and indirectly guides the intelligent agent to learn in the direction of maximizing the feedback reward value, thus reducing the dependence on an accurate system model[5].
Reinforcement learning refers to the learning of the mapping from environmental states to actions, with the aim of maximizing the cumulative reward value obtained by the actions from the environment. The basic components of reinforcement learning include a set of states S that represent the environment, a set of actions A that represent the actions of the intelligent agent, and the reward r for the intelligent agent. In the actual process, the interaction between these two roles includes the state, action, and reward (the reward can be either positive or negative), as shown in Figure 1 Environmen-Agent . The large-scale application of reinforcement learning requires the use of generalization function approximation[6], such as neural networks, decision trees, or instance-based methods. In the past decade, the dominant method has been the value function method. The goal of reinforcement learning is to enable the intelligent agent to learn the optimal control strategy in an unknown environment, thereby maximizing the expected future returns[7]. The core idea of reinforcement learning is to continuously interact between the intelligent agent and the environment, and select reasonable actions with the goal of maximizing the cumulative return. This coincides with the process of acquiring experiential knowledge and making decisions in human intelligence.

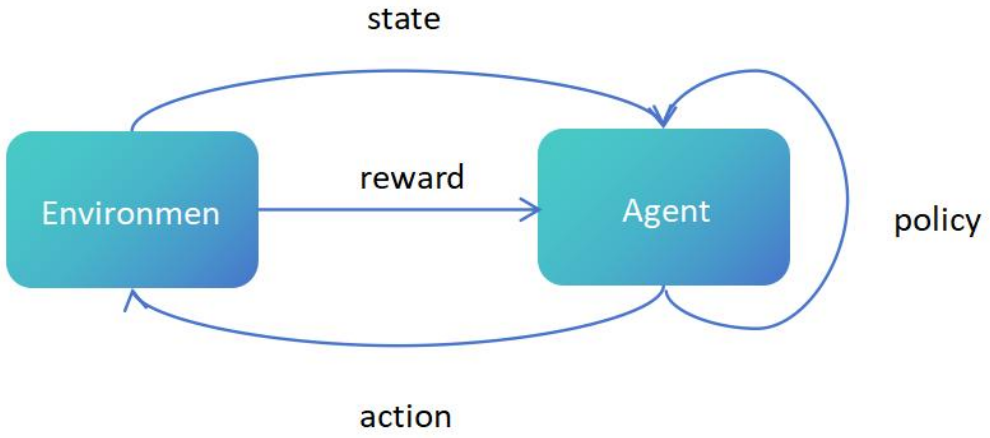

Figure 1 Environmen-Agent

The behavior policy function is the code of conduct for an intelligent agent. In essence, it is a mapping, which maps the set of environmental states to the probability distribution function or probability density function of the set of behaviors, guiding the intelligent agent to select the best behavior. The reward function is the evaluation criterion for the intelligent agent. Usually, positive rewards are given for good behaviors, and negative rewards are given for the opposite. The value function is used to describe the quality of the current state, including the state value function and the action value function [8]. Generally, the interaction process between the intelligent agent and the environment in reinforcement learning can be modeled as a Markov Decision Process (MDP)[9]. The core idea of model-based reinforcement learning originates from the theory of dynamic programming. Its core algorithm realizes the optimization process of policy iteration and value iteration by applying the Bellman equation and the Bellman optimality equation. Model-free reinforcement learning, on the other hand, adopts a learning method of directly interacting with the environment for sampling. The current research mainly focuses on three technical directions: optimization algorithms centered around the value function, optimization methods based on the search in the policy space, and optimization methods combined with the construction of the environment.

2 BACKGROUND

Under the traffic conditions of mixed traffic flow in China, the traditional solution of presetting situations in advance is obviously no longer suitable for today's autonomous driving. Cars need to make autonomous decisions in situations that have not been preset by humans in advance. The autonomous decision-making technology dynamically adjusts the path through reinforcement learning to find the optimal path. In recent years, many scholars have made progress in the field of path planning by using reinforcement learning methods[10-12]. As a key technology of autonomous driving [13], path planning serves as a bridge connecting environmental perception and motion control. The quality of the planned path directly affects the driving trajectory of the vehicle. It is the foundation for achieving autonomous driving and has broad scientific research value and commercial application prospects[14].

3 PROBLEM FORMULATION

Suppose an autonomous vehicle needs to reach its destination through a rectangular grid route, where each grid intersection is equipped with a traffic light displaying the remaining time for each signal phase. When the traffic light is green, the vehicle is allowed to go straight or turn left; when it is red, the vehicle can only stop or turn right. Based on literature review, in real-world scenarios, traffic light cycles at different intersections vary, and even within the same intersection, the durations of straight-going and left-turning green phases may differ. For simplicity, we assume that all traffic lights along the route share identical settings, with red and green lights in different directions switching simultaneously. The research objective is to determine how the autonomous vehicle can reach its destination in the shortest possible time under these conditions.

The problem exhibits the characteristics of reinforcement learning algorithms: the agent conducts trial and exploration in the environment, and optimal policy selection is made based on rewards from the environment. We model this problem as follows:

3.1 Problem Modeling

Environment Setup:

An *N*N* grid of intersections, with the starting point at coordinates (1, 1) and the destination at *(N, N)*. The vehicle's initial orientation is north. The path matrix is denoted as W, the average speed of the vehicle is v, and the total traffic light cycle duration is 160 seconds—80 seconds for the green phase and 80 seconds for the red phase. The north-south direction starts with a green light followed by a red light, while the east-west direction starts with a red light followed by a green light.

State Space Definition:

$S = \left(x_{1i}, x_2, x_3 \right)$,where $x_{1i} = \left(x_{11}, x_{12}, x_{13} \right)$ respectively represent the distances traveled by the vehicle when going straight, turning left, and turning right to the next intersection (unit: meters).

x_2 represents the remaining time for the transition from red - light to green - light (if it is green - light, then it is directly 0). $x_3 \in \left\{ N, E, S, W \right\}$ represents the current driving direction of the vehicle.

Action Space Definition:

The action vector $a_t \in R^4$ is one of the four unit vectors:

$$a_t \in \left\{ e_1, e_2, e_3, e_4 \right\} \tag{1}$$

where:

$e_1 = (1,0,0,0)$: go straight ; $e_2 = (0,1,0,0)$: turn left ;

$e_3 = (0,0,1,0)$: turn right ; $e_4 = (0,0,0,1)$: stop.

Above are the state space and action space of this problem. Next, the state transition function (deterministic) is defined as follows: Let the current time be t, the accumulated passage time be T_t , the state be $S = \left(x_{1i}, x_2, x_3 \right)$ and the action be $a_t = e_i$.

Then the next state is $S = \left(T_{1i,k+1}, t_{2,k+1}, x_{3,k+1} \right)$, and the calculation logic is as follows:

(1) $x_{1i,k+1}$ Path Update

According to the current position and vehicle orientation $x_{3,t}$, using the path matrix W and action $a_t = e_i$, calculate the distances in each direction after the next position reaches the new node.

The following Table 1 Vehicle Orientation Changes is used to illustrate the changes in the direction of the car.

Table 1 Vehicle Orientation Changes

Current Orientation	Go Straight	Turn Left	Turn Right
N	N	W	E
E	E	N	S
W	W	S	N
S	S	E	W

(2) Calculation of the remaining time for red - light to turn green for $x_{2,k+1}$:Calculate the current cumulative passage

$$\begin{cases} x_{3,t} \in \left\{ N, S \right\} \rightarrow green\ first, then\ red \\ x_{3,t} \in \left\{ \mathrm{E}, W \right\} \rightarrow red\ first, then\ green \end{cases} \tag{2}$$

Calculate the current cumulative time: $T_{k+1} = T_t + \dfrac{v distance\ of\ action - taken\ direction}{v} + x_{2,t}$

Determine the state of the traffic light: $p = T_{t+1} \bmod 160$

When $x_{3,t} \in \{N, S\}$, $x_2 = \begin{cases} 0 & , p < 80 \quad green \\ 160 - p, & p > 80 \quad red \end{cases}$;When $x_{3,t} \in \{E, W\}$, $x_2 = \begin{cases} 0 & , p < 80 \quad r \\ 160 - p, & p > 80 \quad red \end{cases}$

Design objective of the reward function: save time and avoid violations.

(1) Time Cost

$$r_1 = -\frac{d(a_t)}{v}, d(a_t) = \begin{cases} x_{11}, & if\ a_t = e_1 \\ x_{11}, & if\ a_t = e_1 \\ x_{13}, & if\ a_t = e_3 \\ 0\ , & if\ a_t = e_4 \end{cases} ; \tag{3}$$

(2) Violation Penalty : $r_2 = \sigma_\delta(red(x_3, \phi))I_{a_t = e_1 \vee a_t = e_2} + \sigma_\delta(red(x_3, \phi))I_{a_t = e_4}$;

(3) Right - turn Reward when x_2 is Large $r_3 = \sigma(\frac{x_2}{\tau})\sigma_\delta(red(x_3, \phi))I_{a_t = e_3}$;

Finally, the total reward is set as $r(s_t, a_t) = w_1 r_1 - w_2 r_2 + w_3 r_3$

Table 2 Symbol Explanation is for explaining the symbols.

Table 2 Symbol Explanation

Symnol	Explanation
a_k	Current Action
$I_{a_t = e_i}$	Indicator function, which equals 1 when $a_t = e_i$
$red(x_3, \phi)$	Judge whether the current is a red light
$\sigma(z)$	Used for smoothing processing
x_2	The time left for the red light
$h_j(s_t; \theta)$	The j -th action approximated and output by the neural network.
θ	Parameter set of the neural network
γ	Discount rate
w	Weight

In the reinforcement learning framework, the learning objective is to maximize the expected cumulative reward. This objective guides the agent to make optimal decisions in successive actions, aiming to achieve the highest long - term reward and optimize its environmental behavior.

$$J(\theta) = \mathrm{E}\left[\sum_{t=0}^{T} \gamma^t \cdot r(s_t, a_t)\right] \tag{4}$$

Decision function $\pi(s, a)$:

$$\pi(a_t \mid s_t; \theta) = \frac{\exp(h_i(s_t; \theta))}{\sum_{j=1}^{4} \exp(h_j(s_t; \theta))}, a_t = e_i \tag{5}$$

With the specific form of the decision function, we can further analyze the impact of policy parameters on performance. Let θ represent the vector of policy parameters, μ represent the performance of the corresponding policy. In the policy gradient method, the policy parameters are updated approximately proportionally to the policy gradient: $\Delta\theta \approx \alpha \frac{\partial \mu}{\partial \theta}$.

Where α represents the learning rate. Through the above formula, it can be ensured that θ converges to a locally optimal strategy.

Policy Gradient Theorem:

$$\frac{\partial \mu}{\partial \theta} = \sum_S d^{\pi}(S) \sum_a \frac{\partial \pi(a \mid S;\theta)}{\partial \theta} Q^{\pi}(S,a) \tag{6}$$

The policy gradient formula provides a method for the gradient of the parameter θ. Substituting the decision function into this theorem enables further exploration of the optimized policy function. Assume $g_w : S \times a \to \mathbb{R}$ approximates $Q^{\pi}(S,a)$, substitute into the formula to find the extreme point:

$$\sum_S d^{\pi}(S) \sum_a \frac{\partial \pi(S,a)}{\partial \theta} \left[Q^{\pi}(S,a) - g_w(S,a) \right] = 0 \tag{7}$$

3.2 Strategy Optimization Implementation

In practical engineering, parameterize the decision function using a neural network, use the *PolicyNet* function, and adopt a three - layer fully - connected network. Input the state vector, and the output is transformed into a decision function (action probability distribution) $\pi(a_t \mid S;\theta)$ through the *Soft* max function. As shown in Figure 2 Parameter Update Flowchart.

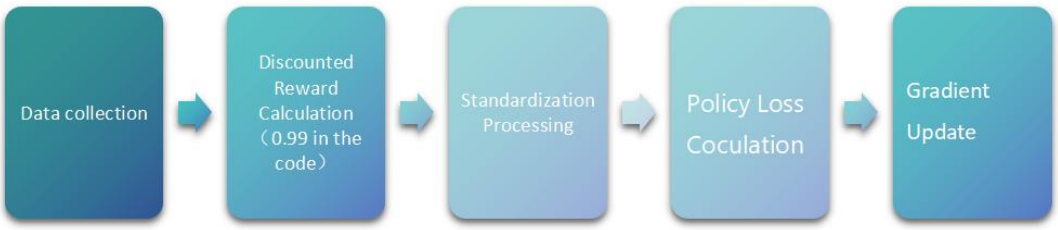

Figure 2 Parameter Update Flowchart

During the training process, the change in the number of cumulative - reward episodes of the agent is shown in the figure:

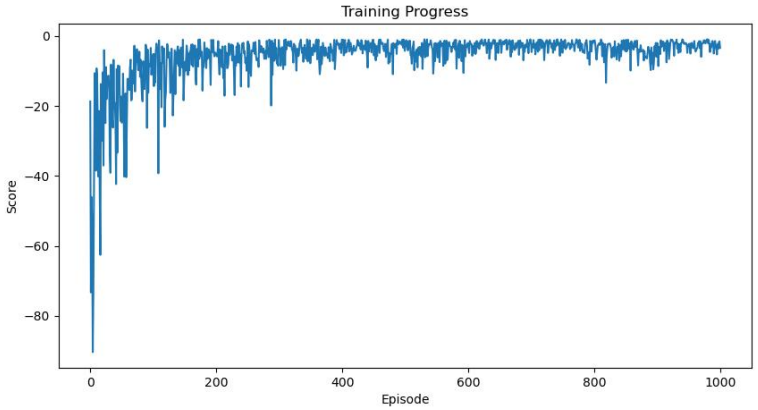

Figure 3 Reward Variation

As can be seen from the Figure 3 Reward Variation, in the early stage, the agent explores the environment and the reward fluctuates greatly. In the later stage, with the progress of training, the reward gradually stabilizes at a relatively high level, indicating that the policy network has effectively learned and chosen the shortest path.

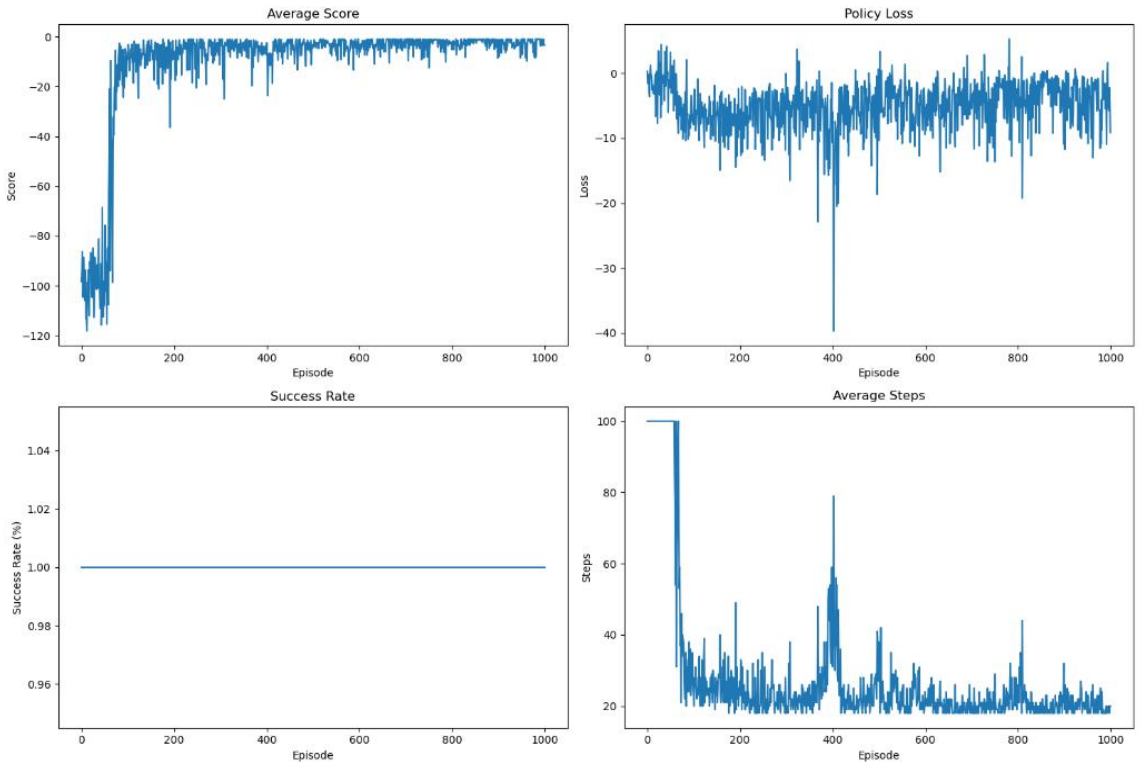

Figure 4 Data Experiment Visualization Results

As can be seen from Figure 4 Data Experiment Visualization Results, the average score fluctuates greatly and has a low value in the early stage of training. This is because the agent hardly obtains positive rewards in the early exploration stage. In the later stage, the average score rises and tends to stabilize, indicating that the agent has learned to implement effective strategies, which shows that the policy gradient algorithm successfully enables the driverless car to learn high - reward strategies. The policy loss fluctuates during training but generally tends to decrease, reflecting the effectiveness of the algorithm. The success rate is nearly 100% with almost no fluctuation, indicating that the model strategy is correct. The agent has good adaptability, can stably achieve the goal, and has a high degree of reliability and generalization ability. The average number of steps fluctuates greatly with peak values in the early stage, and the number of steps to complete the task is unstable. However, it decreases and stabilizes at a lower level in the later stage, indicating that the agent has found a better action sequence and improves efficiency.
Here are the key metrics tracked from Table 3 Data Presentation: success rate, average steps per episode, average reward, and individual episode rewards/steps:

Table 3 Data Presentation

Success Rate	Average Steps:	Average Reward	Rewards per epi -sode	Steps per episode
100.0%	18.2	-1.22	[-0.92,-0.89,-0.94,-1.05,-2.30']	[18, 18, 18, 18, 19]

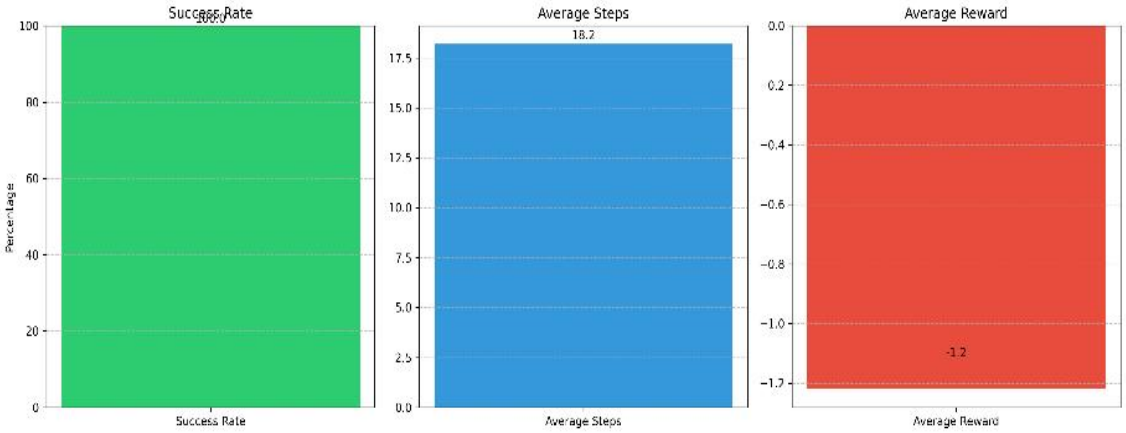

Figure 5 Data Presentation

We can see from Figure 5 Data Presentation table that Average episode steps stabilized at 18.2, demonstrating consistent task - completion patterns and reliable behavioral repeatability.
Notably, the average reward remained negative at -1.22. Despite high success rates and stable episode lengths, this negative reward profile indicates the agent's decision - making under the current reward schema fails to generate net positive returns. Key contributors may include:
Reward Function Formulation: Over - penalization of minor deviations or insufficient goal - achievement incentives.
Policy Optimization Bias: Frequent suboptimal choices prioritizing task completion (high success) over reward maximization, likely due to exploration - exploitation imbalances.

4 CONCLUSION

This paper applies policy gradient methods to address the path planning problem in autonomous driving under mixed traffic flow scenarios. By constructing a state-action space and designing a reward function focused on time efficiency and safety, a three-layer neural network is used to parameterize the decision function, realizing iterative optimization of the policy network. Experimental results show that the agent can stably achieve the shortest path with a 100% success rate, verifying the effectiveness of the policy gradient approach. However, the negative average reward indicates the need for refining the reward function to balance penalty and incentive mechanisms. This study provides a valuable framework for integrating reinforcement learning into intelligent transportation systems, highlighting both the potential of policy gradient methods and the importance of reward function design for practical autonomous driving applications. In the future, we plan to expand the application scope of this approach by incorporating more complex traffic scenarios, such as adverse weather conditions and emergency situations. Additionally, we aim to improve the generalization ability of the model through the integration of transfer learning techniques, enabling it to adapt more effectively to different road networks and traffic regulations.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Piero Scaruffi. The Nature of Intelligence: 64 Big Questions in the Fields of Artificial Intelligence and Robotics. Beijing: People's Posts and Telecommunications Press, 2018.
[2] Sun Changyin, Mu Chaoxu. Several Key Scientific Issues of Multi-Agent Deep Reinforcement Learning. Acta Automatica Sinica, 2020, 46(7): 1301-1312.
[3] Dong Hao, Yang Jing, Li Shaobo, et al. Research Progress of Robot Motion Control Based on Deep Reinforcement Learning. Control and Decision Making, 2022, 37(2): 278-292.
[4] Li Kaiwen, Zhang Tao, Wang Rui, et al. Research Progress of Combinatorial Optimization Based on Deep Reinforcement Learning. Acta Automatica Sinica, 2021, 47(11): 2521-2537.
[5] Xiong Luolin, Mao Shuai, Tang Yang, et al. A Survey of Integrated Energy System Management Based on Reinforcement Learning. Acta Automatica Sinica, 2021, 47(10): 2321-2340.
[6] Yang Ting, Zhao Liyuan, Liu Yachuang, et al. Dynamic Economic Dispatch of Integrated Energy System Based on Deep Reinforcement Learning. Automation of Electric Power Systems, 2021, 45(5): 39-47.
[7] Sutton R S, McAllester D, Singh S, et al. Policy gradient methods for reinforcement learning with function approximation. Advances in neural information processing systems, 1999, 12.
[8] Chen Jiapan, Zheng Minhua. A Survey of Robot Manipulation Behavior Research Based on Deep Reinforcement Learning. Robot, 2022, 44(2): 236-256.
[9] Wang Han, Yu Yang, Jiang Yuan. A Survey of Advances in Multi-Agent Reinforcement Learning Based on Communication. Scientia Sinica (Informationis), 2022, 52(5): 742-764.
[10] Liu Hongqing, Wang Shimin. Research on Vehicle Routing Problem Based on Reinforcement Learning. Computer Applications and Software, 2021, 38(8): 303-308.
[11] Zhang Rongxia, Wu Changxu, Sun Tongchao, et al. Research Progress of Deep Reinforcement Learning and Its Application in Path Planning. Journal of Computer Engineering & Applications, 2021, 57(19).
[12] Huang Dongjin, Jiang Chenfeng, Han Kaili. Three-Dimensional Path Planning Algorithm Based on Deep Reinforcement Learning. Journal of Computer Engineering & Applications, 2020, 56(15).
[13] Xu Hongxin, Wu Zhizhou, Liang Yunyi. A Review of Research on Path Planning Methods for Autonomous Driving Vehicles Based on Reinforcement Learning. Application Research of Computers/Jisuanji Yingyong Yanjiu, 2023, 40(11).
[14] Zhu Maofei, Hu Fangya, Li Nake, Zhu Shouli, Wu Qiong. A Survey of Path Planning Algorithms for Driverless Vehicles. Agricultural Equipment & Vehicle Engineering, 2023, 61(11): 18-22.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3036

CURRENT STATUS AND APPLICATION TRENDS OF EQUIPMENT FAILURE PREDICTION AND HEALTH MANAGEMENT TECHNOLOGY

Jie Liu
China Special Vehicle Research Institute, Jingmen 448035, Hubei, China.
Corresponding Email: 1499703502@qq.com

Abstract: With the rapid advancement of science and technology, the integration, complexity, and intelligence of weapons and equipment have significantly increased. Traditional fault diagnosis and maintenance support technologies are becoming increasingly inadequate in meeting emerging requirements. To address the demands of information warfare for rapid, reliable, and accurate operation of military systems, Prognostics and Health Management (PHM) technology has progressively evolved. PHM represents an innovative approach to health condition management, developed by leveraging the latest advancements in modern information and artificial intelligence technologies. It constitutes a systems engineering discipline derived from practical engineering applications, continuously refined into systematic and structured methodologies focused on monitoring, predicting, and managing the health status of complex engineering systems. Today, PHM technology is receiving widespread attention globally and is considered a core enabler in the development of next-generation weapons systems and the realization of autonomous logistics support. This article provides an overview of the current applications of PHM technology, aiming to serve as a reference for its implementation in equipment management.
Keywords: PHM; Failure prediction; Health management; Current status

1 INTRODUCTION

With the ongoing advancements in the complexity, integration, and intelligence of modern weaponry and equipment, along with evolving maintenance concepts and reforms in maintenance methodologies, Prognostics and Health Management (PHM) technology has progressively matured. PHM involves the real-time monitoring of various operational status parameters and characteristic signals through advanced sensor technologies. It leverages intelligent reasoning algorithms and models to assess equipment health conditions, predict potential faults prior to their occurrence, and generate maintenance and support decisions by integrating available resource information, thereby enabling condition-based maintenance [1-3]. A comprehensive PHM system typically encompasses functionalities such as fault prediction, fault isolation, fault diagnosis, health management, and life tracking. For complex systems and equipment, PHM should be capable of performing multi-level and hierarchical diagnostics, prognostics, and health management. The workflow of a PHM system generally consists of seven core processes [4,5], as illustrated in Figure 1.

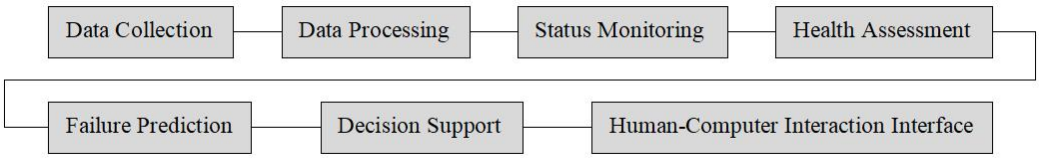

Figure 1 The Framework of the PHM System

2 THE CURRENT STITUATION OF TECHNOLOGICAL DEVELOPMENT AT HOME AND ABROAD

2.1 The Current Situation of Technological Development Abroad

The development of PHM technology abroad has progressed through five distinct stages: external programming, built-in testing (BIT), intelligent BIT, comprehensive diagnosis, and PHM integration. Concurrently, the evolution of maintenance decision-making technology has encompassed post-failure maintenance, periodic preventive maintenance, condition-based maintenance (CBM), and intelligent maintenance[6-9]. At the application level, PHM implementation has advanced from component-level and subsystem-level solutions to system-level integration that encompasses all major subsystems across the entire platform. Currently, PHM technology is being extensively researched and widely adopted by leading military powers such as the United States and the United Kingdom. It has become a critical component in the development and deployment of next-generation weapons systems, including aircraft, ships, and armored vehicles. A notable example of its application can be found in the F-35 fighter jet and related equipment systems.

2.1.1 Application of the PHM system in the F-35 aircraft

The F-35 aircraft represents one of the most prominent implementations of PHM technology. Its PHM system comprises both airborne and ground-based components. The airborne system employs a hierarchical intelligent reasoning architecture, integrating various types of diagnostic and predictive software at multiple design levels. This structure enables the comprehensive application of fault diagnosis and prediction technologies, spanning from individual components to the entire system level. The ground-based system integrates status data, performance metrics, trend analysis, fault diagnostics, prognostics, and remaining useful life predictions for both individual aircraft and fleets. Based on an expert system framework, this integration supports large-scale maintenance forecasting for entire combat units and facilitates autonomous logistics support.

According to statistics from the U.S. military, after the F-35 fighter jet incorporated PHM (Prognostics and Health Management) technology, the non-repeatable fault rate decreased by 82%, maintenance manpower requirements were reduced by 20% to 40%, logistics support scale was cut by 50%, the number of spare parts types increased by 25%, aircraft operation and maintenance costs dropped by 50% compared to previous models, and the service life reached 8,000 flight hours. These figures clearly illustrate the significant effectiveness of PHM in reducing maintenance and support expenditures, enhancing the safety, availability, and reliability of military equipment, ensuring mission success, and improving overall combat capability.

2.1.2 Engine PHM system design abroad

Rolls-Royce's T900 series Engine Health Management System (EHMS) comprises both onboard and ground station components. The onboard system primarily employs sensor technology to monitor engine conditions and collect operational data, performing preliminary data processing, storage, and triggering alarms when necessary. The ground station is responsible for advanced signal processing and data analysis based on the collected information, enabling comprehensive fault diagnosis and predictive maintenance.

Pratt & Whitney has developed a universal architecture for aero-engine PHM systems based on data fusion techniques. This architecture utilizes appropriate algorithms to analyze high-frequency signals obtained from structural health and vibration monitoring sensors, which are then calibrated using low-frequency signals from airway and fuel/lubrication oil sensors. The calibrated signals from gas path and lubrication systems first pass through an analytical module that identifies performance variations across engine modules, integrates oil system parameters, and detects anomalies in the gas path. Subsequently, all characteristic information is transmitted to the advanced fault diagnosis and gas information fusion module. By integrating this data with the fault code information from the engine's full-authority digital electronic controller, the module generates a comprehensive health assessment of the engine and identifies potential faults. Finally, by incorporating input from maintenance personnel and pilots along with historical engine maintenance records, the fault isolation inference engine module formulates a recommended maintenance plan.

2.1.3 The common application situations of PHM systems abroad

The application situations of common PHM systems in foreign countries such as the United States and the United Kingdom are shown in Table 1.

Table 1 The Application Situation of PHM Technology Abroad

Category	PHM system	Application situation
Spacecraft	IVHM、ISHM	X34; X37
Fixed-wing aircraft	PHM、HUMS	F-35; F-22; R-2; C-130; C-17; RQ-7A/B; P-8A; EF-2000
Helicopter	HUMS、JAHUMS	AH-64; UH-60; CH-47; RAH-66; EH-101; NH-90

2.2 Current Situation of Domestic Technological Development

China started relatively late in the field of PHM technology and has a weak research foundation. In the late 1970s, the maintenance of aviation equipment had been following the traditional experience-based maintenance methods of the 1930s and 1940s abroad. The inspection time and content were directly determined by experience, and the maintenance was carried out by senses in accordance with the prescribed time and content. Since the 1980s, China has vigorously developed research on PHM-related technologies such as condition monitoring, fault prediction and reliability maintenance. And it was included in the national "863" development program.

2.2.1 Theoretical exploration and in-depth research on PHM technology in China

In terms of the design of engine PHM systems, Sun Jianzhong from Nanjing University of Aeronautics and Astronautics has conducted research on the health management issues of the new generation of domestic military and civil aviation engines in three aspects: condition monitoring, health status assessment, and remaining life prediction modeling. This provides reliable methods and technologies for improving the health management level of domestic civil aviation engines. To solve the problem of fault prediction for liquid rocket engines, Shen Jiao from National University of Defense Technology proposed a fault prediction method based on error prediction correction. This method has significantly improved the prediction accuracy and adaptability compared with the single process neural network prediction model.

In terms of the design of the PM system for weapons and equipment, Yang Hongjun from the 76327 Unit of the People's Liberation Army of China, based on the analysis of the characteristics of level 3 faults of faulty equipment in the armored command information system, and starting from the actual needs of the maintenance and support of armored command information system equipment, designed the architecture and functional composition of the PHM system for armored command information system equipment. And effective theoretical exploration was carried out on the construction of the PHM system. He Xianwu and others from the Naval Aeronautical Engineering College have studied the effectiveness and feasibility of applying PHM technology in the maintenance and support of anti-ship missiles, and designed the PHM system structure of the anti-ship missile weapon system and the sensor network structure in the maintenance and support of anti-ship missiles. Zhang Liang and others from the Air Force Engineering University have focused on the technical characteristics, difficulties and maintenance and support requirements of the new generation of combat aircraft of our military Various schemes of the airborne PHM system architecture were compared and analyzed, and a hierarchical architecture integrating four layers, namely module/unit level PHM, subsystem level PHM, regional level PHM and platform level PHM, was proposed.

In the design of the PHM system for unmanned aerial vehicles (UAVs), the research group led by Feng Guoqiang from the Air Force Engineering University utilized the self-equilibrium theory in biology to design the fault-tolerant control and fault diagnosis system for UAVs. By designing the "equilibrium point", they determined the safe state of the UAVs and applied fault-tolerant control and fault diagnosis to fault prediction and health management. Su Xujun and others from Army Engineering University proposed a method for predicting faults in unmanned aerial vehicle (UAV) systems using BP neural networks, and presented the modeling and calculation methods of BP neural networks. The calculations show that this method can be used for UAV fault prediction. Peng Lelin and others from the Liberation Vehicle Artillery Academy established the topological structure of the system equipment based on the fault characteristics of the unmanned aerial vehicle (UAV) system, and on this basis, constructed the logical architecture of the UAV PHM system. Lei Yaolin and others from the 54th Research Institute of China Electronics Technology Group Corporation designed an unmanned aerial vehicle (UAV) health prediction and health management system based on telemetry data and intelligent interpretation. They also conducted simulation experiments on the fault prediction algorithm using simulated fault samples, indicating that the UAV health prediction system has a good fault detection capability.

2.2.2 PHM technology by the Air Force Radar Academy

Wang Hanzhong from the Air Force Radar Academy, in view of the system structure and fault characteristics of modern radar equipment, adopted a hierarchical and modular fault diagnosis method, organically combined intelligent diagnosis technologies such as neural networks, expert systems and fuzzy reasoning, and designed a comprehensive intelligent fault diagnosis system for modern radar equipment. He also focused on the knowledge acquisition, knowledge base and reasoning construction of the system. In-depth research was conducted on the main functional modules and key issues such as the extraction of characteristic parameters of diagnostic information. The simulation results show that the system can conduct efficient and accurate automatic diagnosis of complex faults, overcome the limitations of traditional diagnosis methods, and effectively improve the fault diagnosis ability of modern radar equipment.

2.2.3 Technical applications of PHM technology in fields such as aviation, aerospace, shipping, and weaponry

In the aviation field of our country, focusing on model technology research and development, we have carried out test design and verification for flight control actuator systems, rotary actuator drive equipment, hydraulic power systems, component casings, power supply systems, avionics processing machines, and metal composite material airframe structures. Research and related verification on diagnosis and performance table regression prediction technologies: In the aerospace field. At present, the satellite power system mainly conducts in-orbit status monitoring, performance degradation prediction, operation management and life extension of solar cell arrays, batteries and controllers. Manned spaceflight has also carried out status monitoring and fault-tolerant control for some key systems. In the field of shipping, technical applications such as condition monitoring, fault diagnosis, operation and auxiliary maintenance decision-making have been carried out for key equipment in the main and auxiliary systems. In the field of weaponry, engineering applications such as on-board condition monitoring and auxiliary maintenance guidance in a network environment, enhanced diagnosis in the monitoring center, and auxiliary decision-making for tasks and maintenance have been implemented for launch vehicles.

3 DEVELOPMENT TRENDS

The future development of PHM technology will gradually move towards intelligence, networking and specialization. The development of PHM technology in China should fully draw on foreign research achievements, focus on independent innovation, and at the same time, do a good job in theoretical research on PHM technology, technological innovation, infrastructure construction, and the cultivation of soft power of talents. Form a PHM technology system with Chinese characteristics and oriented towards the new generation of equipment for the future.

3.1 Intelligence

In recent years, with the continuous deepening of theoretical research and development of condition monitoring and diagnosis technologies, and the continuous breakthroughs and industrial applications of modern sensor technologies with high precision, high performance and high information volume have been achieved. New methods for condition monitoring and fault diagnosis are constantly emerging, such as expert diagnosis, fuzzy diagnosis, neural network

diagnosis, and the combination of the above-mentioned various diagnoses. Intelligence has become the future development trend.

3.2 Networking

Through networked online equipment monitoring, the transmission of equipment status and data storage and analysis are achieved, realizing full coverage, dynamic and continuous monitoring of equipment. Moreover, diagnostic analysis of monitoring data can be conducted to predict and determine the current damage degree and risk level of the equipment, ensuring the safe and stable operation of the equipment. Through intelligent diagnosis and analysis, the networked online monitoring system not only provides monitoring for the operational status of the equipment but also offers a scientific basis for regular and irregular safety inspections and maintenance of the equipment.

3.3 Specialization

In the future, the technology of equipment condition monitoring and fault diagnosis in our country will develop in a specialized direction, with increasingly detailed division of labor. The implementation of condition monitoring and fault diagnosis will mainly be integrated. Overall solution providers that can offer professional equipment condition monitoring and fault diagnosis software and hardware systems and have rich diagnostic technology talents will become the mainstream. China will gradually form industry-wide, regional and even national-level equipment condition monitoring and fault cloud diagnosis centers.

4 CONLCUSION

As a cutting-edge technology, Prognostics and Health Management (PHM) plays a crucial role in improving the reliability, safety, maintainability, and supportability of equipment. This paper reviews the current research status of PHM technology both domestically and internationally, and discusses its future development trends, with the objective of laying a solid foundation for the further advancement and practical application of equipment PHM systems.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Jinzhen Kong, Jie Liu, Jingzhe Zhu, et al. Review on Lithium-ion Battery PHM from the Perspective of Key PHM Steps. Chinese Journal of Mechanical Engineering, 2024, 37(04): 14-35.
[2] Wang Hong, Kulevome Delanyo Kwame Bensah, Zhaozi an. An integrated PHM framework for radar systems through system structural decomposition. Journal of Systems Engineering and Electronics, 2025, 36(01): 95-107.
[3] Yao Zhao, Li Hejia, Tu Chao, et al. Application of Fault Prediction and Health Management in Vehicle Control System. Journal of Physics: Conference Series, 2021, 1852(4). DOI: 10.1088/1742-6596/1852/4/042069.
[4] Hu Nannan, Su Xiaohao, Baolong Liu. A Review on Prognostics and Health Management. International Journal of Advanced Network, Monitoring and Controls, 2018, 3(1): 115-121. DOI: 10.21307/ijanmc-2018-022.
[5] Cao Yang, Liu Hongtian, Song Chao, et al. Research on a fault test system based on fault injection. Journal of Physics: Conference Series, 2021, 1894(1). DOI: 10.1088/1742-6596/1894/1/012023.
[6] Shen Shilei, Ma Linlong. Research on the Built-in Test Design of Civil Aircraft Flight Control System. Journal of Physics: Conference Series, 2022, 2252(1). DOI: 10.1088/1742-6596/2252/1/012010.
[7] Huang Jian Xin, Bian Ya Qin. Research on BIT Intelligence Maintenance Decision of Weapon Materiel. Applied Mechanics and Materials, 2011, 128-129(128-129): 224-228. DOI: 10.4028/WWW.SCIENTIFIC.NET/AMM.128-129.224
[8] Drees R, Young N. Role of BIT in support system maintenance and availability. IEEE Aerospace and Electronic S ystems Magazine, 2004, 19(8): 3-7.
[9] Straka Ondřej, Punčochář Ivo. Distributed design for active fault diagnosis. International Journal of Systems Scien ce, 2022, 53(3): 562-574. DOI: 10.1080/00207721.2021.1963501.

World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3037

HETEROGENEOUS FENTON CATALYTIC OXIDATION TECHNOLOGY FOR LEACHATE TREATMENT IN GENERAL SOLID WASTE LANDFILLS

Yan Zhou
Environmental Protection and Energy Conservation Center, Sinopec Fifth Construction Co.Ltd, Guangzhou 510145, Guangdong, China.
Corresponding Email: zhouyan.swuj@sinopec.com

Abstract: To address the challenges posed by the complex water quality, high pollutant concentration, and susceptibility to environmental influences in general solid waste landfill leachate, this study developed a combined treatment process based on heterogeneous Fenton catalytic oxidation. Taking a landfill leachate treatment project as an example, the process route "Denitrification Pretreatment - Heterogeneous Fenton Catalysis - Post-treatment System" was employed. Through the synergistic action of units including mechanical bar screening, coagulation sedimentation, air stripping deammonification, activated carbon adsorption, heterogeneous Fenton oxidation, A/O biological treatment, and advanced filtration/disinfection, efficient pollutant removal was achieved. Practical engineering application demonstrated stable system operation. The effluent quality met the discharge standards specified in Table 2 of the "Standard for Pollution Control on Municipal Solid Waste Landfill" (GB16889-2024), with removal rates for COD, BOD_5, ammonia nitrogen, total nitrogen, and suspended solids reaching 98%, 98.8%, 99.3%, 99.04%, and 98.6% respectively. The TOC degradation rate exceeded 60%. This process offers both environmental and economic benefits, providing a replicable engineering model for the harmless treatment of landfill leachate.
Keywords: Leachate; Heterogeneous Fenton catalysis; Solid waste landfill; Treatment process; Advanced oxidation technology

1 INTRODUCTION

Within China's current solid waste disposal system, incineration and landfilling remain the dominant technological pathways, yet both face significant environmental challenges[1]. Leachate generated during waste landfilling, classified as high-concentration organic wastewater, exhibits complex composition (containing humic substances, heavy metals, inorganic salts, and emerging contaminants like antibiotics and microplastics) and significant fluctuations in quality influenced by factors such as landfill age and climatic conditions[2]. Untreated discharge of this wastewater poses potential threats to ecosystems and human health through groundwater infiltration and soil accumulation[3].

Among typical landfill leachate treatment technologies, recirculation was widely used due to its simplicity and low initial cost. However, long-term operation can lead to salt and recalcitrant organic compound accumulation, causing scaling and fouling in reverse osmosis systems and reduced treatment efficiency[4]. While evaporation enables total volume reduction, the challenge of treating concentrated residues and high energy costs limit its widespread engineering application. Research has examined the effects of temperature and free ammonia on nitrification and nitrite accumulation in landfill leachate[5]. A granular sludge reactor (GSR) achieved efficient nitrogen removal via nitritation/denitritation for treating mature leachate with high ammonia nitrogen content[6], but its effectiveness for other organics was limited. Dynamic membrane bioreactors (DMBR) have also been compared for landfill leachate treatment[7]. In contrast, heterogeneous Fenton catalytic oxidation technology, generating hydroxyl radicals (·OH, redox potential 2.80 V) via H_2O_2 decomposition mediated on the catalyst surface, demonstrates unique advantages in the broad-spectrum degradation of organic pollutants[8]. Compared to traditional homogeneous Fenton systems, this technology overcomes drawbacks such as Fe^{3+} hydrolysis/precipitation, narrow operational pH range, and high sludge production. The recyclability of the supported catalyst further enhances process economics.

Therefore, based on the principles of "reduction, resource recovery, and harmless treatment", and targeting the characteristics of poor biodegradability and high heavy metal content in mature solid waste landfill leachate[9], this study developed a coupled system "Denitrification Pretreatment-Heterogeneous Fenton Catalysis-Advanced Post-treatment". Through the synergistic action of multiple units, efficient pollutant removal was achieved, providing an engineered solution for compliant leachate disposal.

2 WATER QUALITY CHARACTERISTICS OF LANDFILL LEACHATE

Leachate produced from municipal solid waste sanitary landfills is compositionally complex, with water quality and quantity varying greatly due to external factors. It typically contains substantial amounts of organic pollutants (including synthetic compounds from consumer products like cosmetics), inorganic heavy metals, inorganic salts, and emerging contaminants. Inappropriate discharge causes severe environmental pollution and threatens human health.

Leachate pollutants can be categorized into four groups: inorganic macro-pollutants, non-biodegradable organics, dissolved organic matter, and heavy metals. Key discharge indicators include Chemical Oxygen Demand (COD), color (CN), Ammonia Nitrogen (NH_3-N), Total Nitrogen (TN), and Total Phosphorus (TP)[10]. Organic pollutants often consist largely of hard-to-degrade humic substances like humic and fulvic acids.

Landfill age is a major factor influencing leachate composition. Researchers typically classify leachate based on landfill age: less than 5 years (young), 5-10 years (intermediate), and over 10 years (mature/old). Notably, as landfill age increases, COD levels gradually decrease, accompanied by increasing humification, leading to progressively worse biodegradability[11]. Changes in other water quality characteristics are shown in Table 1.

Table 1 Variation of Landfill Leachate Water Quality Over Time

Characteristic	Young	Intermediate	Mature
Filling time（year）	＜5	5-10	>10
pH	＜6.5	6.5-7.5	>7.5
COD（mg/L）	>10000	4000-10000	<4000
BOD5/COD	0.5-1.0	0.1-0.5	<0.1
Ammonia Nitrogen	<400	-	>400
Heavy Metals	Low - Medium	Low	Low
Biodegradability	Good	Moderate	Poor

3 LEACHATE WATER QUALITY INDICATORS

The target water body for this study is mature leachate from a general solid waste landfill site. Determination of its water quality parameters is based on 36 consecutive months of monitoring data from this landfill and the stringent requirements of the "Standard for Pollution Control on Municipal Solid Waste Landfill" (GB16889-2024). Due to a landfill age exceeding 10 years, the mature leachate exhibits characteristics of high ammonia nitrogen (3000-3500 mg/L), low biodegradability ($BOD_5/COD < 0.1$), and humic substance enrichment. Compared to young leachate, although its COD concentration is reduced below 4000 mg/L, the molecular structure of organic matter is more complex, and the ionic strength of heavy metals and inorganic salts is higher.

During water quality detection, a spatio-temporal dynamic analysis method was employed: samples were collected from different landfill areas (e.g., center, periphery) and during typical seasons (rainy season, dry season), covering 23 conventional indicators including COD, BOD_5, ammonia nitrogen, total nitrogen, suspended solids, and trace heavy metals. The data showed COD concentration fluctuations up to 33.3% (4500-6000 mg/L). Ammonia nitrogen concentration exhibited seasonal fluctuations influenced by the waste degradation stage, consistent with the literature conclusion that "mature leachate quality is significantly regulated by landfill layer microbial activity". Specific design water quality indicators are shown in the table 2 below:

Table 2 Influent and Effluent Water Quality of Leachate

Parameter	Influent Quality	Effluent Standard
COD/(mg·L^{-1})	4500-6000	100
BOD_5/(mg·L^{-1})	2000-3000	30
Ammonia Nitrogen /(mg·L^{-1})	3000-3500	25
Total Nitrogen /(mg·L^{-1})	3200-4000	40
Suspended Solids /(mg·L^{-1})	1000-2000	30
Manganese /（μg·L^{-1}）	2160-3000	1500
Total Bacteria /（CFU·mL^{-1}）	1000-2000	1000

4 TREATMENT PROCESS FLOW DESIGN

The landfill leachate treatment process is illustrated in Figure 1: Leachate passes through a 5 mm mechanical bar screen to intercept debris before entering an equalization tank (HRT=24 h, mixing/homogenization). It is then pumped to a coagulation sedimentation tank where Polymeric Aluminum Ferric Chloride (PAFC) (2.0 g/L) and Polyacrylamide (PAM) (12 mg/L) are dosed to remove suspended solids. The pretreated water phase enters an air stripping deammonification tower (pH=11, gas-to-liquid ratio 5000:1) for ammonia removal, followed by deep impurity removal in an activated carbon fiber adsorption column (contact time 2 h). The purified water phase is pumped into the heterogeneous Fenton reactor (loaded iron-based catalyst, pH 3-5, H_2O_2 dosage 1.2 times theoretical value, HRT=2-3 h) for catalytic oxidation degradation of organics. The effluent is dosed with lime milk for neutralization to pH 7-8 to precipitate metal ions. The neutralized water phase enters an A/O biological tank (anaerobic 4 h / aerobic 8 h) for nitrogen and phosphorus removal, followed by chemical precipitation dosing Ferric Sulfate (30 mg/L) for enhanced phosphorus removal. The settled effluent sequentially passes through a quartz sand filter (particle size 0.5-1.2 mm) → activated carbon filter to adsorb residual pollutants, and finally undergoes chlorine dioxide disinfection (5-10 mg/L, contact time 30-60 min). The Fenton catalyst can be recycled. If effluent testing exceeds standards (COD >100 mg/L or NH_3-N >25 mg/L), the flow is diverted back to the biological tank or Fenton unit for enhanced treatment based on the parameter.

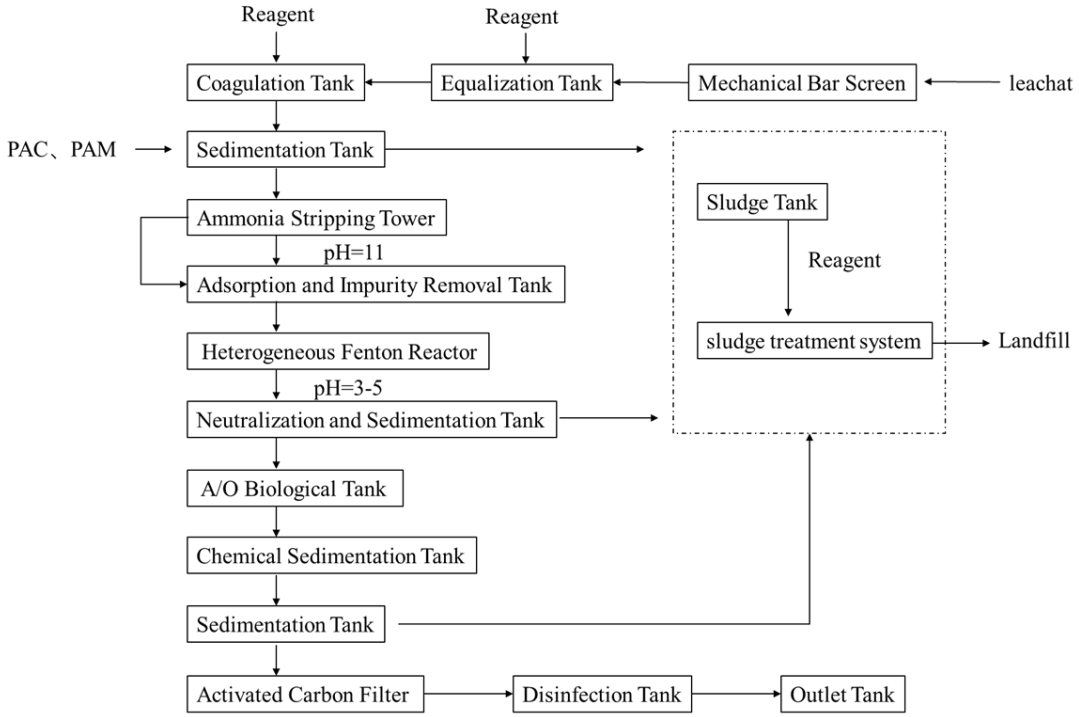

Figure 1 Process Flow Diagram

5 TREATMENT TECHNOLOGY SCHEME DESIGN

5.1 Pretreatment System

5.1.1 Bar screen

Leachate first passes through a bar screen. A mechanical bar screen is used to intercept larger suspended and floating solids (e.g., plastics, wood blocks, fibers) in the leachate, preventing these impurities from entering subsequent units and causing clogging or damage. The bar spacing is set at 5 mm, effectively intercepting larger particles and ensuring smooth operation of subsequent processes.

5.1.2 Equalization tank

Screened leachate enters the equalization tank. Its primary function is to balance water quality and quantity, as leachate generation and composition fluctuate significantly due to rainfall, waste filling rates, etc. The effective tank volume is 1000 m^3, with a hydraulic retention time (HRT) of 24 hours. Mixing equipment ensures thorough homogenization, preventing stratification and providing stable influent conditions for downstream treatment units.

5.1.3 Coagulation sedimentation

Leachate from the equalization tank enters the coagulation sedimentation tank. During coagulation, Polymeric Aluminum Ferric Chloride (PAFC) and Polyacrylamide (PAM) are dosed. PAFC acts as a coagulant; its hydrolyzed polynuclear hydroxyl complexes destabilize colloidal particles through adsorption, bridging, and charge neutralization. PAM acts as a flocculant, promoting aggregation of destabilized particles into larger flocs for easier sedimentation and separation.

Coagulant dosing experiments investigated the removal efficiency of COD and TN. At a constant PAM dose of 20 mg/L, COD and TN removal rates increased with PAFC dosage, reaching maxima at 2.5 g/L PAFC (67.4% COD removal, 54.2% TN removal). At a constant PAFC dose of 2.00 g/L, COD and TN removal showed no clear relationship with PAM dose, though a PAM dose of 24 mg/L yielded good results (62.2% COD, 50.0% TN removal). Results are shown in Table 3. Considering cost-effectiveness, optimal doses were determined as PAFC 2.50 g/L and PAM 20 mg/L. At these doses, COD and NH_3-N removal rates reached 61.5% and 47.9%, respectively, indicating effective treatment.

Table 3 Optimal Coagulant Dosage

Exp. No.	PAFC/（$g \cdot L^{-1}$）	PAM/（$mg \cdot L^{-1}$）	COD Removal /%	TN Removal /%
1	1.5	20	47.9	46.7
2	1.75	20	53.9	47.9
3	2.00	20	59.8	49.2
4	2.25	20	65.0	50.0
5	2.50	20	67.4	54.2
6	2.00	12	61.5	47.9
7	2.00	16	62.2	48.3
8	2.00	20	61.1	46.7
9	2.00	24	62.2	50.0

10	2.00	28	62.8	47.9

A tube settler is used, increasing the settling area and efficiency, effectively removing suspended solids, some heavy metals, and organics. However, the COD concentration in the coagulation effluent remains relatively high, requiring further treatment.

5.2 Heterogeneous Fenton Catalytic Oxidation

5.2.1 Reaction principle

Pretreated leachate enters the heterogeneous Fenton reactor. Within the reactor, the heterogeneous Fenton catalyst reacts with hydrogen peroxide (H_2O_2). The catalyst is an iron-based material supported on activated carbon. The reaction principle involves: On the catalyst surface, H_2O_2 decomposes under the catalysis of iron oxides to generate highly oxidizing hydroxyl radicals (·OH, E^0 = 2.80 V). These radicals non-selectively oxidize and decompose organic pollutants in the leachate, breaking down macromolecules into smaller compounds or even mineralizing them to CO_2 and H_2O. Concurrently, iron ions on the catalyst surface undergo redox cycling (Fe^{3+} reduced to Fe^{2+}, Fe^{2+} reacting with H_2O_2 to regenerate ·OH), sustaining the reaction. Being supported, the catalyst minimizes loss and allows reuse, enhancing stability and economics.

5.2.2 Reaction condition control

Strict control ensures optimal performance. Reaction pH is maintained between 3-5, where catalyst activity and ·OH generation efficiency are high. Online pH monitors provide real-time feedback for automatic sulfuric acid or sodium hydroxide dosing. Temperature is controlled at 30-40°C using a heating/cooling system. H_2O_2 dosage is set at 1.2 times the theoretical requirement based on influent COD, ensuring sufficient oxidant. Hydraulic Retention Time (HRT) is 2-3 hours, allowing sufficient contact time between pollutants and ·OH.

5.2.3 Reaction equipment

The heterogeneous Fenton reactor is a Continuous Stirred Tank Reactor (CSTR) made of corrosion-resistant stainless steel. An internal agitator (100-150 rpm) ensures thorough mixing of leachate, catalyst, and H_2O_2, enhancing reaction efficiency. A bottom distribution system ensures uniform influent entry, preventing short-circuiting. A vent at the top collects and treats minor amounts of harmful gases generated during the reaction.

5.3 Post-treatment

5.3.1 Neutralization and precipitation

Effluent from the Fenton reactor enters the neutralization/precipitation tank. Due to its acidic nature, lime milk is dosed first to adjust pH to 7-8, precipitating metal ions (e.g., iron, heavy metals) as hydroxides. Lime dosage is automatically controlled based on real-time pH and metal ion concentration monitoring. Precipitation and separation further reduce metal ion content in the effluent.

5.3.2 Filtration

The neutralized effluent enters the filtration unit, combining sand filtration and activated carbon filtration. The sand filter uses quartz sand (0.5-1.2 mm) to remove residual fine suspended solids and colloidal matter, reducing turbidity. The granular activated carbon (GAC) filter utilizes its highly porous structure and large surface area to adsorb residual organics, color, odor, and trace heavy metals, polishing the effluent quality.

5.3.3 Disinfection

The filtered effluent undergoes disinfection using chlorine dioxide (ClO_2). ClO_2 is a strong oxidant effective at killing bacteria, viruses, and pathogens[12]. Dosage is 5-10 mg/L, generated on-site. A contact time of 30-60 minutes ensures effective disinfection, preventing microbial hazards upon discharge.

6 OPERATION AND PERFORMANCE ANALYSIS

(1) The physical pretreatment technology in this project, addressing the suspended solids issue in leachate, developed a physical process. By screening oil-absorbing media and optimizing filtration conditions, the levels of SS, oil content, and turbidity in the water phase were reduced. This configuration ensured high filtration efficiency and excellent water quality.

(2) The advanced oxidation technology for the pretreated water phase developed heterogeneous Fenton catalysis. Experimental conditions including H_2O_2 ratio, reaction pH, flow rate, and temperature were optimized. Utilizing the highly oxidizing hydroxyl radicals generated in-situ as the primary oxidant, recalcitrant organic pollutants in the wastewater were oxidized and degraded, even mineralized to CO_2, H_2O, and small carboxylic acids, while simultaneously reducing ammonia and total nitrogen, achieving the goal of harmless treatment.

(3) Post-treatment technology was developed to complement the advanced oxidation unit, targeting residual TOC, ammonia nitrogen, and total nitrogen in the effluent. Adsorption was employed to separate harmful substances, utilizing the well-developed pore structure, vast surface area, and various functional groups of the adsorbent to remove diverse organic pollutants and metal ions from the leachate. In this case study, effluent concentrations of COD, BOD, NH_3-N, and SS met discharge standard limits. Figure 2 shows the leachate treatment performance.

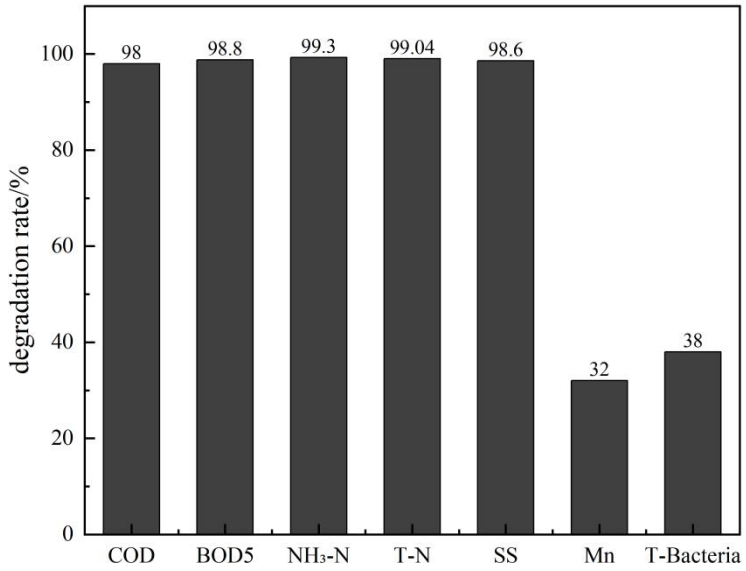

Figure 2 Leachate Treatment Performance

(4) Data collected during stable operation of the equipment showed TOC removal exceeding 60%, reaching a maximum of 75%. All indicators met design requirements, and treatment costs were reduced by 17.5%. Figures 3, 4, and 5 illustrate H_2O_2 decomposition and TOC removal performance.

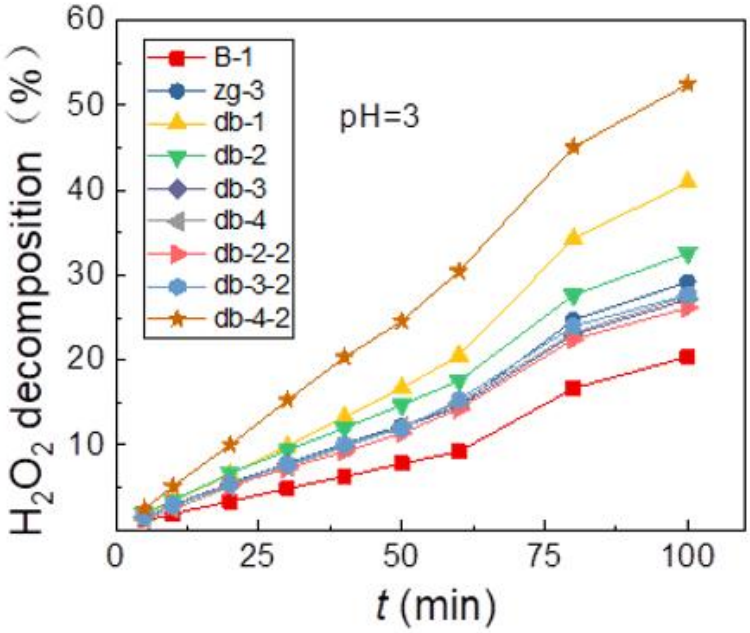

Figure 3 H_2O_2 Decomposition Rate

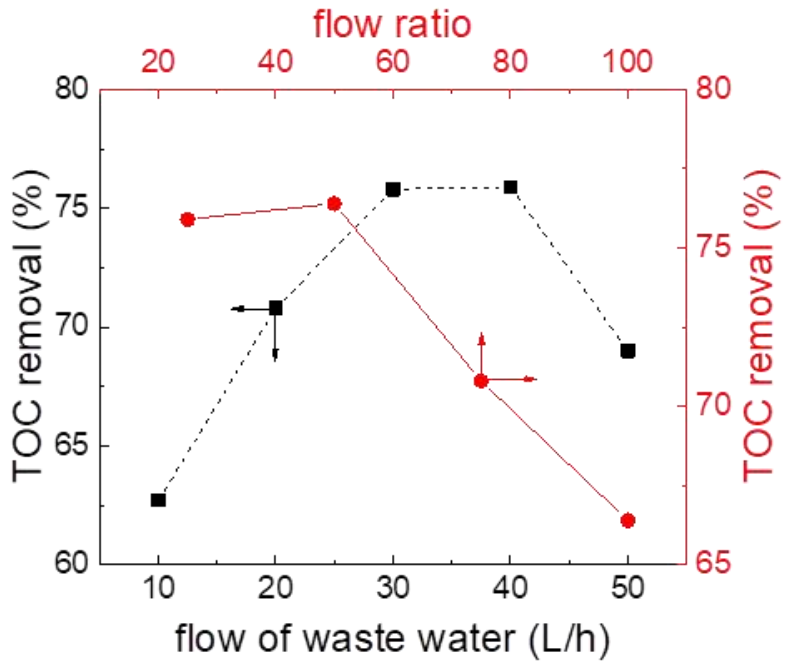

Figure 4 TOC Removal Performance

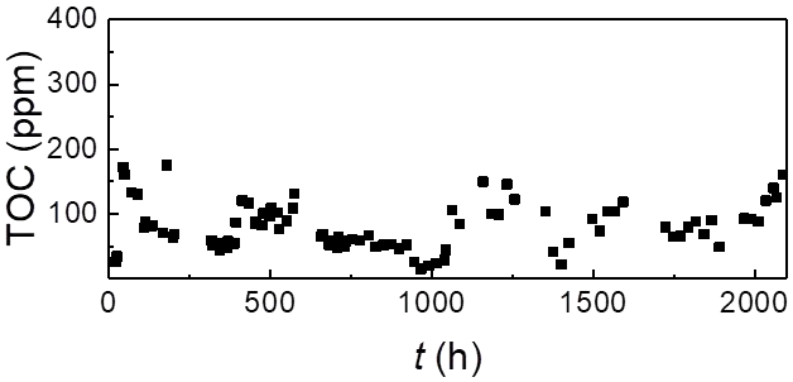

Figure 5 TOC Removal Performance

7 CONCLUSION

This study focused on the characteristics of mature leachate from general solid waste landfills and constructed a combined process of "Denitrification Pretreatment - Heterogeneous Fenton Catalytic Oxidation-Advanced Post-treatment". Engineering practice validation yielded the following key results:

(1) Significant Process Synergy: Pretreatment units (screening, coagulation sedimentation) effectively removed suspended solids, laying the foundation for deep treatment. Synergy between air stripping deammonification and activated carbon adsorption enhanced pre-removal of nitrogen and recalcitrant organics. The core heterogeneous Fenton unit, utilizing a supported iron-based catalyst under optimized conditions (pH 3-5, H_2O_2 1.2x theoretical, HRT 2-3 h), efficiently generated hydroxyl radicals via Fe^{3+}/Fe^{2+} cycling on the catalyst surface, achieving broad-spectrum degradation and mineralization of organic pollutants, overcoming the limitations of homogeneous Fenton (narrow pH range, high sludge yield). Post-treatment units (A/O biological treatment, chemical precipitation, filtration/disinfection) ensured stable, compliant effluent quality. The synergistic linkage of all units formed a complete technical chain from pretreatment to advanced purification.

(2) Stable Compliant Effluent: Continuous operational monitoring data demonstrated removal rates of 98% for COD, 98.8% for BOD_5, 99.3% for ammonia nitrogen, 99.04% for total nitrogen, and 98.6% for suspended solids. TOC degradation exceeded 60%. Effluent quality consistently met all discharge standards specified in Table 2 of GB16889-2024. This validates the combined process's high efficiency in removing concentrated, recalcitrant pollutants from mature landfill leachate, solving a long-standing technical challenge in the industry.

(3) Techno-economic and Environmental Benefits: The recyclability of the supported catalyst in the heterogeneous Fenton unit significantly reduced catalyst consumption and sludge disposal costs compared to homogeneous Fenton, contributing to an overall cost reduction of 17.5%. The entire process adheres to the principles of "reduction, resource recovery, and harmless treatment," minimizing pollutant discharge at the source and preventing secondary pollution of soil and groundwater ecosystems. Aligned with green and low-carbon development, the process holds significant potential for replication and promotion.

In summary, the combined process developed in this study breaks through the technical barriers for treating mature solid waste landfill leachate, achieving a synergistic balance of high purification efficiency, economic viability, and environmental protection. It provides a replicable and promotable technical paradigm for similar projects, holding significant practical value and exemplary significance for advancing landfill leachate treatment technology in China and improving ecological environmental quality. Future research could focus on catalyst long-term stability mechanisms and intelligent process control to further enhance application efficacy.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Gan Tao. Analysis of Industrial Solid Waste Disposal and Comprehensive Utilization Measures. Low Carbon World, 2025, 15: 28-30.
[2] Ma Jiamin, Ren Xueyu, Zeng Mengyuan, et al. Research Status on Characteristics and Treatment Technologies of Non-sanitary Landfills. Environmental Engineering, 2023, 1-18.
[3] Qin Jianyou, Shu Lin, Tan Qiuyan, et al. Current Status and Management Strategies for Leachate Monitoring in Municipal Solid Waste Landfills. Low Carbon World, 2025, 15: 25-27.
[4] Peng Chuanbin. Application of Landfill Leachate Treatment Technology. Energy Conservation and Environmental Protection, 2025: 64-70.
[5] Kim D J, Lee D I, Keller J. Effect of temperature and free ammonia on nitrification and nitrite accumulation in landfill leachate and analysis of its nitrifying bacterial community by FISH. Bioresource Technology, 2006, 3(97), 459–468. https://doi.org/10.1016/j. biortech.2005.03.032.
[6] Zou X, Mohammed A, Gao M, et al. Mature landfill leachate treatment using granular sludge-based reactor (GSR) via nitritation/denitritation: Process startup and optimization. Science of the Total Environment, 2022, 844, 157078. https://doi.org/10.1016/j.scitotenv.2022.157078.
[7] Mora M, Fernández M, Gómez J M, et al. Kinetic and stoichiometric characterization of anoxic sulfide oxidation by SONR mixed cultures from anoxic biotrickling filters. Applied Microbiology and Biotechnology, 2015, 99, 77–87. https://doi.org/10.1007/s00253-014-5688-5.
[8] Wang Shuangyu, Yin Zishan, Gu Shanshan, et al. Research Status of Heterogeneous Fenton Catalysts. Liaoning Chemical Industry, 2025, 54: 324-327.
[9] Zeng Xiaolan, Zhang Yuxi, Ding Wenchuan, et al. Study on Combined Process for Compliance Treatment of Mature Landfill Leachate. National Drainage Committee Annual Conference, 2015: 7.
[10] Renou S, Givaudan J G, Poulain S, et al. Landfill leachate treatment: Review and opportunity. Journal of Hazardous Materials, 2008, 150(3), 468-493.
[11] Lin S H, Chang C C. Treatment of landfill leachate by combined electro-Fenton oxidation and sequencing batch reactor method. Environmental Science, 2000, 34(17), 4243-4249.
[12] Mei Yi, Ma Lanting, Qu Chengtun. Landfill Leachate Treatment and Clean Incineration Technology. Petrochemical Industry Application, 2022, 41(12): 7-10.