World Journal of Engineering Research
ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3021

LEAST SQUARES AND CLUSTER ANALYSIS BASED METHODOLOGY FOR SONIC BOOM LOCALIZATION OF ROCKET DEBRIS

XianYang Zhou[1*], WeiRong Zhang[2], SiQi Ji[1], Yang Song[1]
[1]*College of Arts and Sciences, Northeast Agricultural University, Harbin 150030, Heilongjiang, China.*
[2]*Guiyang No. 25 High School, Guiyang 550025, Guizhou, China.*
Corresponding Author: XianYang Zhou, Email: xianyangzhou1799@163.com

Abstract: Aiming at the transonic sonic boom localization problem generated by the falling rocket debris, this paper proposes a spatio-temporal localization method of debris based on acoustic monitoring. The sonic boom signal is received by multiple monitoring devices, and the effects of time error and equipment layout on positioning accuracy are analyzed to construct a single target and multi-target cooperative positioning model. For single wreckage, the four-sphere intersection principle and the least-squares method with error compensation are adopted to realize the accurate solution of three-dimensional coordinates and sonic boom moment through four monitoring devices, and the localization results are 110.57° longitude, 27.16° latitude, 957.96 m elevation, and 100.753 s. For multi-wreckage scenarios, the classification model of vibration wave signals is proposed, and the classification model of vibration wave signals is proposed by combining the Pearson correlation coefficient and the K-means clustering algorithm. A vibration wave signal classification model is proposed, and a time difference threshold constraint (≤5s) is introduced to optimize the wreckage matching, which effectively solves the problem of overlapping multi-source sonic boom signals. The results of this paper show that the proposed method significantly improves the robustness and accuracy of rocket debris localization in complex scenes, and can provide theoretical support for the rapid confirmation of the drop point.
Keywords: Rocket debris; Transonic sonic boom; Error compensation; Least squares; K-means

1 INTRODUCTION

Traditional methods for localizing rocket debris mainly rely on radar or visual monitoring, and these methods have problems such as limited accuracy and susceptibility to interference in complex environments. Especially in the crash area with complex terrain and variable environment, the traditional methods have limited localization effect. As a non-contact localization means[1], acoustic monitoring has gradually become a research hotspot in the field of rocket debris localization due to its advantages such as less influence by environmental factors and wide coverage.

In recent years, the localization technology based on acoustic monitoring has made significant progress in several fields[2]. For example, in the field of indoor localization, ultra-wideband (UWB) technology has realized the precise positioning of target objects by virtue of its high accuracy and strong anti-interference ability[3]. In addition, sound source localization technology based on microphone arrays has been widely used in the fields of robot auditory perception and speech recognition. The successful application of these technologies provides valuable reference and lessons for the sonic localization of rocket debris[4].

However, the sonic boom localization problem of rocket wreckage has its special characteristics [5]. The high intensity and short duration characteristics of the sonic boom signal, as well as the complex aerodynamic and thermodynamic effects during the falling process, all increase the difficulty of localization. To address the above problems, this paper proposes a spatio-temporal localization method for rocket debris based on sonic monitoring[6]. The method receives sonic boom signals through multiple monitoring devices, analyzes the influence of time error and equipment layout on positioning accuracy, and constructs a single target and multi-target cooperative positioning model. For single wreckage, the four-sphere intersection principle and the least squares method with error compensation are adopted to realize the accurate solution of 3D coordinates and sonic boom moments. For multi-wreckage scenarios, combining Pearson correlation coefficient and K-means clustering algorithm[7], the vibration wave signal classification model is proposed, and the time difference threshold constraint is introduced to optimize the wreckage matching, which effectively solves the problem of overlapping of multi-source sonic boom signals[8]. The experimental results show that the method significantly improves the robustness and accuracy of rocket debris localization in complex scenes, and provides theoretical support for the rapid confirmation of the drop point[9].

The research in this paper not only provides technical support for the safe recovery of rocket debris, but also opens up new ideas for the application of acoustic positioning technology in aerospace field.

2 SINGLE WRECKAGE LOCALIZATION MODEL BASED ON FOUR-SPHERE INTERSECTION WITH DYNAMIC ERROR COMPENSATION

2.1 Equipment coordinate normalization and four-sphere intersection principle

© By the Author(s) 2025, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

2.1.1 Normalization of device coordinates

Firstly, the geographic coordinates (longitude, latitude and elevation) of the monitoring equipment are converted into units in meters, and since the distance value per latitude is approximated to be 111.263 km, and the distance value per longitude is approximated to be 97.304 km, the following conversions can be performed, and the locations of the seven testing equipment in the Cartesian coordinate system are shown in Table 1:

$$\begin{cases} x_i = x \times 97304(meter) \\ y_i = y \times 111263(meter) \\ z_i = z(meter) \end{cases} \tag{1}$$

$(x_i, y_i z_i)$ $(i = 1,2...7)$denotes the positional coordinates of the seven monitoring devices A to G.

Table 1 Three-dimensional Coordinates after Conversion

Equipment	Longitude (m)	Latitude (m)	Elevation (m)
A	9370485	2992440	824
B	9416300	3020160	727
C	9410520	3056350	742
D	9371335	3060750	850
E	9394540	3037870	786
F	9389695	3071310	678
G	9353995	2983310	575

2.1.2 Four-sphere intersection localization (math.)

In this paper, the sonic boom arrival time is used to construct a time difference for multistation passive localization[10], and since acoustic waves propagate at a constant velocity in the air, the relationship between the arrival time and the speed of sound can be utilized to calculate the distance from the detection equipment to the rocket wreckage.

Assuming that the start time of the sonic boom in the wreckage is t_0, and the time for the sonic boom to reach each piece of equipment is t_i $(i = 1,2,3...7)$, then the time required for the sonic boom to be transmitted is $(t_i - t_0)$. It is known that the velocity of propagation of the vibration wave is v = 340 m/s, and the distance between the wreckage and each piece of equipment is:

$$d_i = v \cdot (t_i - t_0) \tag{2}$$

Assuming that the 3D coordinates of the wreckage when the sonic boom occurs are(x, y, z) and the 3D coordinates of the monitoring equipment are (x_i, y_i, z_i), the true distance between the monitoring equipment and the wreckage can be expressed as:

$$D_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} \tag{3}$$

Neglecting a number of error factors such as ground curvature, time errors, etc., the following equation should hold:

$$\sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} = v \cdot (t_i - t_0) \tag{4}$$

Three balls in three-dimensional space should intersect at two points, but due to the wreckage of the beginning of the sonic boom time is also unknown, it is still necessary to solve an equation, then at least four monitoring devices need to be arranged to accurately locate the wreckage to each detection device as the center of the sphere, the four spheres of the intersection as shown in Figure 1.

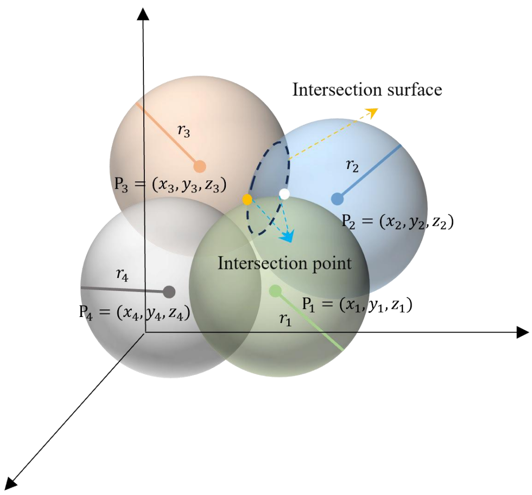

Figure 1 Four-sphere Intersection

That is, the program consists of one sonic boom signal source and four ground monitoring devices, as shown in Figure 2. The principle of localization is to record the moment when the sonic boom signal source and each monitoring station receives the radiation source signal[11], and calculate the time difference between the received localization signals, so as to solve the position of the target.

Figure 2 Sonic Boom Signal Propagation

The equations for passive localization can be listed as Equation 5, and by solving the joint equations by association the positional coordinates of the bomb and arrow wreckage segments can be calculated.

$$\begin{cases} \sqrt{(x_1-x)^2+(y_1-y)^2+(z_1-z)^2} = v\cdot(t_1-t_0) \\ \sqrt{(x_2-x)^2+(y_2-y)^2+(z_2-z)^2} = v\cdot(t_2-t_0) \\ \sqrt{(x_3-x)^2+(y_3-y)^2+(z_3-z)^2} = v\cdot(t_3-t_0) \\ \sqrt{(x_4-x)^2+(y_4-y)^2+(z_4-z)^2} = v\cdot(t_4-t_0) \end{cases} \tag{5}$$

2.2 Least Squares Optimization Algorithm with Error Factors

However, in reality, the signal propagation environment is not ideal, and various influencing factors interfere with the propagated signal, resulting in a positive error in the measured value of the sonic boom's arrival time or arrival time difference, thus making the equation unsolved. Because the measurement error is usually linear with the measurement distance increases the pattern of growth, so here to introduce an error factor β ($0 \leqslant \beta \leqslant 1$). Then the actual distance between the monitoring equipment and the wreckage can be expressed as.

$$D_i = \beta\cdot d_i \tag{6}$$

Then Equation 5 can be transformed into:

$$\begin{cases} \beta^2 {d_1}^2 = (x_1-x)^2+(y_1-y)^2+(z_1-z)^2 \\ \beta^2 {d_2}^2 = (x_2-x)^2+(y_2-y)^2+(z_2-z)^2 \\ \vdots \\ \beta^2 {d_i}^2 = (x_i-x)^2+(y_i-y)^2+(z_i-z)^2 \end{cases} \tag{7}$$

Let the distance from the monitoring equipment to the wreckage be: ki=xi2+yi2+zi2 and, at the same time, transform equation (7):

$$\begin{cases} \beta^2({d_1}^2-{d_2}^2)+2(x_1-x_2)+2(y_1-y_2)y+2(z_1-z_2)z = K_1-K_2 \\ \beta^2({d_1}^2-{d_3}^2)+2(x_1-x_3)+2(y_1-y_3)y+2(z_1-z_3)z = K_1-K_3 \\ \vdots \\ \beta^2({d_1}^2-{d_2}^2)+2(x_1-x_2)+2(y_1-y_2)y+2(z_1-z_2)z = K_1-K_2 \end{cases} \tag{8}$$

It is further rewritten in least squares matrix form:

$$AX = b \tag{9}$$

$$where\ b = \begin{bmatrix} K_1-K_2 \\ K_1-K_3 \\ \vdots \\ K_1-K_i \end{bmatrix}, A = \begin{bmatrix} ({d_1}^2-{d_2}^2) & 2(x_1-x_2) & 2(y_1-y_2) & 2(z_1-z_2) \\ ({d_1}^2-{d_3}^2) & 2(x_1-x_3) & 2(y_1-y_3) & 2(z_1-z_3) \\ \vdots & \vdots & \vdots & \vdots \\ ({d_1}^2-{d_i}^2) & 2(x_1-x_i) & 2(y_1-y_i) & 2(z_1-z_i) \end{bmatrix}$$

where β^2 is the squared value of the measurement error factor, $X=(\beta^2,x,y,z)$ is the vector representation of all unknown quantities, and where (x,y,z) is the 3D coordinates of the rocket wreckage. Therefore the final solution value can be obtained using the least squares algorithm as:

$$X = (A^TA)^{-1}A^Tb \tag{10}$$

2.3 Single Wreck Localization Results and Error Analysis

In this paper, the least squares algorithm with error factor is used for the solution[12]. The error analysis of the proposed method is performed using data, the test data has a total of 1 wreck and 7 monitoring devices, and the arrival time of the signal received corresponding to each device, the results are shown in Table 2:

Table 2 Positioning Results

Parameters	X(latitude)	Y(longitude)	Z(elevation)	T(seconds)
numerical value	110.57	27.16	957.96	100.753

This paper finally obtained the sonic boom point longitude 110.57 °, latitude 27.16 °, altitude 957.96m, sonic boom moment for 100.753 s. In order to more intuitively show the best monitoring equipment and the wreckage of the location of the relationship between the data, plotting the detection of equipment coordinates as shown in Figure 3.

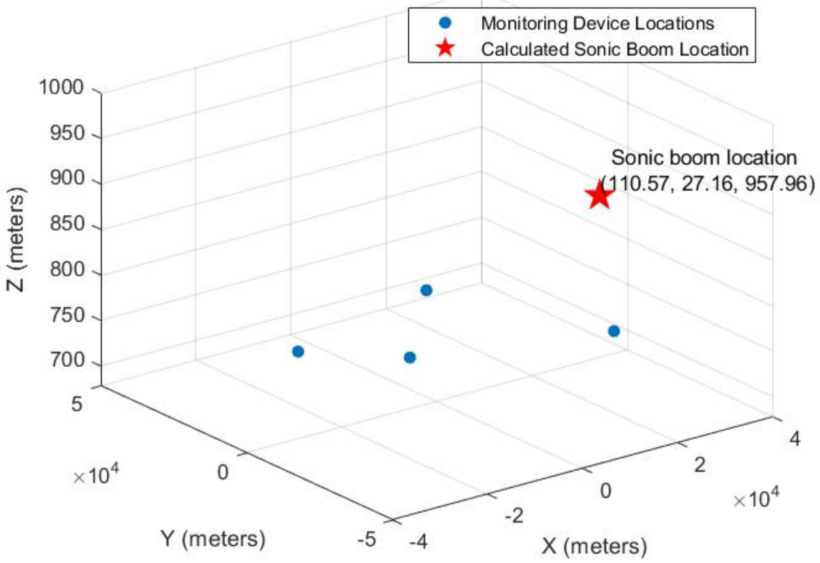

Figure 3 Monitoring Equipment and Wreckage Coordinates

3 CLUSTER CLASSIFICATION AND CO-LOCALIZATION OF MULTI-WRECK SONIC BOOM SIGNALS

3.1 Vibration Wave Correlation Analysis Based on Pearson's Coefficient

For multiple wreckage sonic boom localization analysis, each monitoring device received four vibration waves in chronological order, but could not identify the corresponding relationship between vibration waves and wreckage.

Due to the differences in the sonic boom time of different wreckage, so the arrival of different detection equipment also has a time difference, but belonging to the same wreckage sonic boom time has a certain correlation, so this paper adopts the correlation analysis[13], through the calculation of the distance between any two monitoring equipment and the sonic boom received the correlation coefficient between the difference in time to make a judgment, if the sonic boom signal comes from the same wreckage, the correlation coefficient is close to 1, and for the different wreckage of the sonic boom signal, the correlation coefficient is very small. If the sonic boom signal comes from the same wreckage, the correlation coefficient will be close to 1, while for different wreckage, the correlation coefficients will be small and not the same, which can be k-mean clustering of vibration waves to identify a group of signals from the same wreckage.[14]

In three-dimensional space, let (x_i, y_i, z_i) be the coordinates of the monitoring device i, (x_j, y_j, z_j) be the coordinates of the monitoring device j, d_{ij} be the distance between the device i and j, Δt_{ij} be the time difference between the reception of the sonic boom of the device i and j, $(i, j = 1,2,\cdots,n,\ i \neq j)$, then the distance and time difference between any two monitoring devices are as follows.

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{11}$$

$$\Delta t_{ij} = t_i - t_j \tag{12}$$

The arrival time of a sonic boom signal is a continuous value in seconds, and the Pearson correlation coefficient applies to continuous data, reflecting a linear relationship between the variables.

From a physical point of view, the propagation time of sound waves and distance usually follow a linear relationship. As the distance from the source increases, the time of propagation of the sound wave also increases. Since the data is a continuous arrival time of the sonic boom signal, and assuming that there is a linear relationship between time and distance, the Pearson's correlation coefficient can be chosen, which provides a quantitative measure of the strength of the relationship between the variables, and is calculated using the following formula.

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}} \tag{13}$$

Where *-1≤ r ≤1*, the closer r is to 0, the weaker the correlation, the closer it is to 1, the stronger the positive correlation, and the closer it is to -1, the stronger the negative correlation.

3.2 K-means Clustering Optimization Model with Time Constraints

The K-means algorithm is a commonly used iterative, unsupervised clustering algorithm[15], which achieves the division of the samples in the dataset into a predetermined k clusters by minimizing the distance between the samples to make the samples within the same cluster as similar as possible, and as different as possible between the different clusters, and the algorithm is shown in Table 3.

Table 3 Pseudo-code of K-means Algorithm

```
import      X = {x_j} //data set;  k//number of clusters
    for  k  in max_epochs:
        m_1^0 , m_2^0 ...... m_k^0 //Initialize the center of mass
        for  i  in k:
            d_ji =  d(x_j, m_i) //Calculate distance
        assign x_j to m_i  if j = arg min d_ji//Assigned to the nearest center of mass
                                        label_x_j = j//Get Tagged
        end for
        for  i  in k:
                m_i = (1/|S_j|) Σ_{x_i∈S_j} x_i//Update the center of gravity
        end for
        if |Δm_i| < ε
            end for
        else:
            break
    end for
export   m_1, m_2...... m_k//Best Quality Heart; Y = {label_x_j}//Clustered Tags
```

After calculating the correlation coefficients between the distances of different monitoring devices and the time difference between the reception of sonic booms, this paper uses the K-means cluster analysis algorithm in machine learning to cluster the correlation coefficients and construct a K-means cluster analysis model to obtain the debris corresponding to the vibration waves received by different monitoring devices[16][17].

The first need for data preprocessing, unified data outline. Assuming that there are m debris in the air, the ground has n monitoring equipment, remember the position coordinates of the ith debris is (x_i, y_i, z_i), the position coordinates of the jth monitoring equipment is (X_j, Y_j, Z_j), the moment of receiving the sonic boom signal of the ith debris is $T_{ij}(i = 1,2\cdots m, j = 1,2\ldots n)$, where c denotes the speed of propagation of the vibration wave, then there are:

$$\sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2 + (z_i - Z_j)^2} = c(T_{ij} - t_i) \tag{14}$$

Considering that the four wrecks may produce sonic booms at different times, but differ from each other by no more than 5 s, a constraint exists:

$$|t_i - t_j| \le 5 \tag{15}$$

After adding the constraints, at least 4m independent equations are required to determine the location and sonic boom moments of m wrecks. Similarly, similar to the single-wreck sonic boom localization, a system of overdetermined equations can be obtained by listing the corresponding four equations separately for each wreck[18], when the number of monitoring devices *n > 4*. In order to fully utilize the redundant information and improve the localization accuracy, a least squares problem can be constructed:

$$\min_{x_i, y_i, z_i, t_i} \sum_{i=1}^{m} \sum_{j=1}^{n} (\sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2 + (z_i - Z_j)^2 - c(T_{ij} - t_i)})^2 \tag{16}$$

Therefore, for determining the location and time of the sonic booms of the four wrecks in the air, it is necessary to arrange at least four monitoring devices based on the use of the cluster analysis model and the sonic boom localization model.

3.3 Multi-wreck Localization Equation System Construction and Solution

Taking a given four wrecks as an example, assuming that five monitoring devices are installed on the ground, their locations and the arrival times of the received vibration waves are shown in Table 4:

Table 4 Coordinates of Detection Equipment and Arrival Time of Sonic Booms

Equipment	X(km)	Y(km)	Z(km)	D1(s)	D2(s)	D3(s)	D4(s)
device M	1.0	2.4	0.0	2.0	3.3	4.4	6.6

device N	0.5	0.3	0.0	3.2	4.0	4.9	5.5
device U	3.9	0.4	0.0	2.0	2.6	4.2	8.0
device V	2.0	2.8	0.0	2.5	4.0	5.2	6.9
device W	1.5	2.0	0.0	4.3	5.3	7.0	7.1

Using the least squares model with an error factor, the optimal estimates of the positions and sonic boom moments of the four wrecks were obtained as shown in Table 5 below:

Table 5 Wreckage Coordinates and Sonic Boom Time Estimates

Wreckage	X(km)	Y(km)	Z(km)	sonic boom time(s)
Wreckage 1	-3.36	4.46	5.76	1.44
Wreckage 2	5.45	-3.08	7.83	5.54
Wreckage 3	1.64	3.75	2.56	0.85
Wreckage 4	-2.27	-4.63	7.02	3.38

Comparing and analyzing the localization results with the monitoring data, the error between the theoretical reception time of the vibration wave and the actual time of each monitoring device can be calculated, and the root mean square error is 0.43s . The error is within the allowable range, indicating that the localization results are more accurate and the clustering algorithm and localization model are effective.

For the multi-rocket debris sonic boom localization problem, this paper proposes a method based on least squares estimation and K-means clustering. By establishing a mathematical model for multi-wreckage localization, combined with vibration wave signal identification and clustering analysis, the sonic boom signals of different wreckages are effectively separated and matched. The least squares method is used to optimize and solve the wreckage location and sonic boom moment, and the example validation shows that the method has high accuracy and feasibility. The results show that the method can efficiently deal with the problem of the overlapping of the sonic boom signals from multiple sources, and provides a reliable solution for the localization of the rocket wreckage in complex scenarios. The algorithm can be further optimized in the future to adapt to the dynamic environment requirements.

4 CONCLUSION

Aiming at the complexity of transonic sonic boom localization of rocket wreckage and the problem of multi-source signal aliasing, this paper proposes a spatio-temporal localization method based on acoustic wave monitoring, which significantly improves the localization accuracy and robustness in complex scenes through the single-target and multi-target synergistic localization model.

In single-wreckage localization, the four-sphere intersection principle and the least-squares method with error compensation are used to solve the three-dimensional coordinates of the wreckage and the moment of sonic boom by combining the time difference of the sonic boom signals from the four monitoring devices. By introducing the error factor optimization equation, the measurement error caused by environmental interference is effectively compensated, and high accuracy positioning is finally achieved, and the error analysis shows that the method can significantly reduce the influence of time error and equipment layout on the positioning results.

Aiming at the multi-wreckage scenario, a signal classification model based on Pearson correlation coefficient and K-means clustering is proposed. By calculating the correlation coefficients of time difference and distance of signals received by different monitoring devices, combined with the time difference threshold constraint, the multi-source sonic boom signals are effectively separated. The least squares method is utilized to construct a system of super-definite equations to collaboratively solve the positions and sonic boom moments of multiple wrecks. The validation of the algorithm shows that the root-mean-square error of multi-wreckage localization is 0.43 s. The results are in good agreement with the actual data, which verifies the feasibility and accuracy of the model.

The innovation of this paper is to introduce the error compensation mechanism and cluster analysis into the field of sonic localization, which solves the limitations of traditional methods in complex environments. The research results not only provide reliable theoretical support for the rapid recovery of rocket debris, but also open up a new path for the application of acoustic positioning technology in aerospace engineering. Future research can further optimize the error model and explore the real-time positioning algorithms in dynamic environments to adapt to a wider range of engineering scenarios.

COMPETING INTERESTS

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

REFERENCES

[1] Wang H. Research on mobile robot sound source localization method based on sound arrival time difference. Qingdao University of Science and Technology, 2020.

[2] Jin-Hong Xu. Three-dimensional localization method based on optimization principle in non-line-of-sight propagation environment. Zhejiang Gongshang University, 2018.
[3] Wang J T, Huang F Y, Li Z Z, et al. Research on indoor localization based on UWB ranging technology. Communication Power Technology, 2018, 35(05): 50-53+56.
[4] Wang Pei, Xu Shan, Xiong Wei, et al. Research on trajectory localization and monitoring technology of bullet and arrow wreckage based on TDOA. Fire and Command Control, 2023, 48(11): 139-144+151.
[5] An Jinxia, Qin Zhigang, Li Xuan, et al. Comprehensive situational analysis method and application of rocket debris landing site. Network Security and Data Governance, 2024, 43(11): 37-42.
[6] Chen Zhangrui. Reflection and analysis on launch vehicle recovery technology. Science and Technology Communication, 2019, 11(06): 140-142.
[7] Dai Yufan, Jin Hongping. Indoor localization method based on K-means and KNCN algorithm. Journal of Hubei Institute of Automobile Industry, 2024, 38(04): 59-63+68.
[8] Albert A, Alves S, André M, et al. Acoustic positioning for deep sea neutrino telescopes with a system of piezo sensors integrated into glass spheres. Experimental Astronomy, 2025, 59(1): 6.
[9] Fumiaki T, Motoyuki K. An approximate travel time calculation and a robust GNSS-acoustic positioning method using an MCMC technique. Earth, Planets and Space, 2022, 74(1).
[10] Geng Aoting. Research on key technology of multi-station passive localization based on AOA. Nanchang University, 2023.
[11] Wu P. Research on key technology of multi-station passive localization based on TDOA. National University of Defense Technology, 2019.
[12] Guo Jian, Zeng Zhihao, Huang Xihang. Secondary localization of mobile robot based on improved least squares method. Combined Machine Tools and Automated Machining Technology, 2024, (12): 62-66+71.
[13] Petrushka K, Petrushka I, Yukhman Y. Assessment of the Impact of Military Actions on the Soil Cover at the Explosion Site by the Nemerov Method and the Pearson Coefficient Case Study of the City of Lviv. Journal of Ecological Engineering, 2023.
[14] Xi Z L, Lu W. Application of Kmeans Algorithm with Pearson Coefficient in College Student Sports Quality and Sports Resources. Advanced Materials Research, 2014, 3078(905): 725-730.
[15] Zhang Xiaodong, Li Bo, Zhang Tiange, et al. Calculation and analysis of wind deviation of transmission lines based on K-mean clustering algorithm. Electrical Technology, 2023, 24 (12): 20-26+34.
[16] hemiakina A, Khrisanfov M, Chikurova N, et al. Getting Insights Into Chromatographic Properties of HILIC and Mixed-Mode Homemade Stationary Phases Using Principal Component and Cluster Analyses. Journal of Chemometrics, 2025, 39(3): e70019.
[17] Díaz T M, Ruiz M A, Henríquez L P. Functional PCA and cluster analysis for determining temperature patterns in Chile. Theoretical and Applied Climatology, 2025, 156(3): 186.
[18] Mao Chunxian. Research on the method of solving system of super definite linear equations based on vector DSP. Hunan University of Science and Technology, 2023.