

Research on Chi-square Distribution-based Firing Accuracy Data Analysis

Zhiqiang Zhao^{1, a}, BoChuan Jiang^{1, b} and Jingang Zhang^{1, c}

¹ Science and Technology on Complex land Systems Simulation Laboratory, BeijingChina

^azzqsyp@163.com, ^b13901304418@139.com, ^ctank_zhang@sina.com

Keywords:Operational Test, firing accuracy, elements error, dispersion error, chi-square distribution, confidence degree.

Abstract. In the test and evaluation of operation, the test benefits may be greatly reduced because no useful conclusions may be obtained from test data due to missing or incompleteness of test information records. Based on the characteristics of gun firing and by separating firing errors, this paper proposes an idea using for analyzing firing accuracy in case of incomplete test data records, establishes a method for analyzing chi-square distribution-based firing accuracy data, and conducts test data analysis on test data of a certain type of weapon. The effectiveness of this method is proven by analysis results which have offered effective method for conducting test data analysis.

Introduction

In the equipment combat tests, firing accuracy test is one of the important test subjects which may supply the effectiveness evaluation of weapon and firing control with confident test data. Taking indirect-aiming gun firing as an example, the description of indicators of indirect aiming firing accuracy includes dispersion errors and median errors of elements. For analyzing the above-mentioned indicators based on test data, the coordinate of launching point of each round (accurate to each round, each gun), the bomb-fall coordinate and aiming point coordinate (accurate to each time's modification of gun sight and the results after shooting), as well as the detailed process of test organization. However, In the process of carrying out the actual test, due to defects in test design, incompleteness in test facilities, as well as the requirements on test based on actual combat, etc., the missing or incompleteness of bomb-fall test records may occur, hence, no useful conclusions may be drawn from test data, and accordingly no target set for test reached. Therefore, to implement effective analysis on test data, firstly, the attribution problem of bomb-fall shall be analyzed by adopting clustering method in the multivariate statistical analysis, followed by estimation of firing accuracy and calculation of confidence degree based on chi-square distribution, and at last, the number of tests added, determined.

Analysis Process Under the Conditions of Incomplete Test Data Records

Classification of firing errors.In gun firing, because each kind of method for determining elements shall go through a lot of measurements, calculations, and other preparatory works, each link may generate errors. Element error is a random vector with its starting point at aiming point, and end point at dispersion center. Its projections on shooting direction (x-axis) and on the direction perpendicular to shooting direction (z-axis) are respectively called distance errors (X_c) and direction errors (Z_c) which determine all elements. X_c and Z_c are usually independent to each other and obeying normal distribution, with its median errors denoted as E_d and E_f .

The dispersion may cause the points of burst dispersing from dispersion center. The projections of dispersion errors on x-axis and z-axis are called distance error (X_s) and direction error (Z_s) of dispersion. X_s and Z_s are usually independent to each other and obeying normal distribution, the dispersion's distance median error is B_d and the direction median deviation (error) is B_f . Any deviation of one bomb-fall point against aiming point is called firing error or projecting error of the point of burst. The distance error X of shooting is $X_c + X_s$, and direction error Z is $Z_c + Z_s$.

The distance median error E_x and direction median error E_z of shooting may be obtained from following equations: $E_x = \sqrt{E_d^2 + B_d^2}$ and $E_z = \sqrt{E_f^2 + B_f^2}$

Analysis process under the conditions with incomplete testing data records. The incompleteness and missing of part of test data records in test mainly contains the coordinate of projection points of each shell and the modified sights in shooting process and the coordinate of rear aiming points in shooting direction. When the coordinates of projection point of each shell is unknown, which shells are projected from the same gun can not be distinguished, and so are the grouping of shells. When the modified sights and the rear aiming points in shooting direction of each shell are unknown, the elements errors of the shooting may not be made sure. In case data is missing, the shells in short distances may be classified into one group, and the shells in the group may be regarded as coming from same gun by using clustering method in the multivariate statistical analysis. Based on the results of this classification, the dispersion median errors may be further estimated, and the median errors of elements, roughly estimated. The analysis process is shown in Fig. 1.

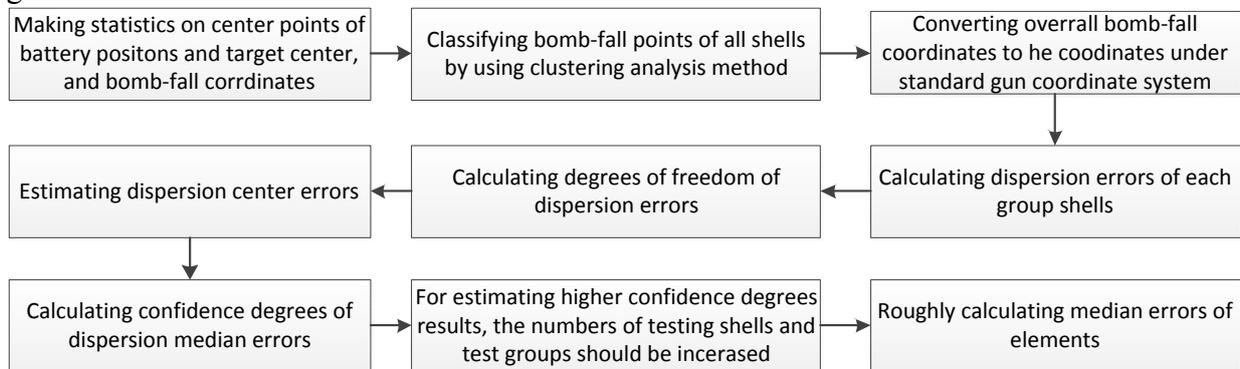


Figure 1 Analysis process under the conditions of incomplete test data records

Data Analysis Method on Firing Accuracy Based on Chi-square Distribution

Analyzing grouping attribution with Kmeans method. Set number of test groups as M , number of fired shells of each group, as N , the number of total guns, as K , as well as the bomb-fall coordinates of N shells in Group m testis $(x_{mi}, z_{mi}), i \in [1, N]$. Because no records about which shells among bomb-fall points of each group of N shells are projected from same gun kept in the Test Data Collection Table, therefore, first of all, the attribution issue of bomb-fall shall be analyzed with clustering analysis method in multivariate statistical analysis. In detail, Kmeans method is adopted, with N bomb-fall points each group being evenly divided into K collections, and each collection containing N/K bomb-fall points, which come from same gun. The one time classification process of Kmeans method is shown in Fig. 2.

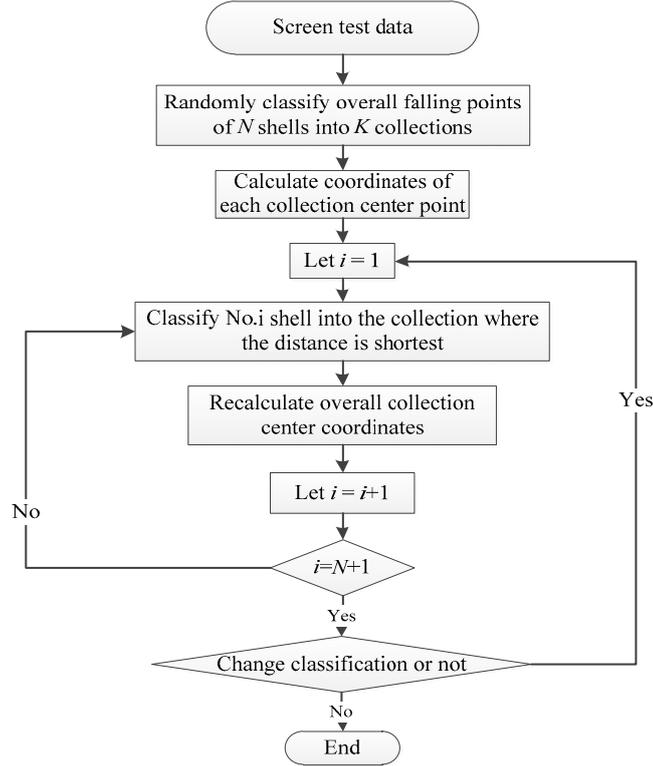


Figure2 Grouping analysis process using Kmeans method

Mark the coordinates of bomb-fall points of Type k as $(x_{ki}, z_{ki}), i \in [1, N/K]$, the center of gravity of Type k , (x_k, z_k) . Because the initial random classification may affect the final classification results, so in actual application, Kmeans method shall be adopted for multiple classifications, and the optimal classification shall be selected by following the principle of minimum distances within the class as bellow:

$$\min \sum_{k=1}^K \sum_{i=1}^{N/K} (x_{ki} - x_k)^2 + (z_{ki} - z_k)^2 \quad (1)$$

Coordinate conversion. The aiming points of each gun are unknown, because of part shortage of such data as shooting intervals, distance difference, etc. Compared with gun range, gun spacing may be ignored, therefore, for estimating dispersion errors in distance and direction, the coordinates of bomb-fall points shall be converted to the right angle coordinate system XOZ with target center as original point, the line between group center and target center as x -axis. In detail, set the coordinate of target center on original coordinate system $X'O'Z'$ as (x_t, z_t) , the gun group center coordinates, as (x_p, z_p) . The conversion formula for converting the coordinate (x, z) in original coordinate system $X'O'Z'$ into the coordinate (X, Z) in new coordinate system XOZ is shown as follows:

$$\begin{aligned} X &= (x - x_t) \cos \theta - (z - z_t) \sin \theta \\ Z &= (x - x_t) \sin \theta + (z - z_t) \cos \theta \end{aligned} \quad (2)$$

Where θ is the direction angle among x -axis in coordinate system $X'O'Z'$ and the connection line between launch point and target center.

Calculation of dispersion errors. Set the bomb-fall point coordinate of No. i shell of No. k group in group No. m of tests as $(x_{mki}, z_{mki}), i \in [1, N/K], k \in [1, K], m \in [1, M]$, then the computing formula for grouping center (x_{mk}, z_{mk}) of Group No. k in Group No. m Test is as follows:

$$x_{mk} = \frac{K}{N} \sum_{i=1}^{N/K} x_{mki}, \quad z_{mk} = \frac{K}{N} \sum_{i=1}^{N/K} z_{mki} \quad (3)$$

Then the dispersion errors of No. i shell of Group No. k is (x_{smki}, z_{smki}) .

$$x_{smki} = x_{mki} - x_{mk}, \quad z_{smki} = z_{mki} - z_{mk} \quad (4)$$

Calculating freedom degree of dispersion errors. In all collections $\{(x_{smki}, z_{smki}), i \in [1, N/K], k \in [1, K], m \in [1, M]\}$ constituted by dispersion errors, there are altogether MN coordinates. Because each bomb-fall point group satisfies equation $\sum_{i=1}^{N/K} x_{smki} = 0, \sum_{i=1}^{N/K} z_{smki} = 0$, therefore, overall dispersion errors have $m(N - K)$ degrees of freedom.

Estimation of dispersion median errors based on chi-square distribution. Because dispersion errors obey normal distribution, the following equation may be obtained

$$\frac{\rho^2 \sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} x_{mki}^2}{B_d^2} \sim \chi^2(M(N-K)) \quad (5)$$

$$\frac{\rho^2 \sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} z_{mki}^2}{B_f^2} \sim \chi^2(M(N-K)) \quad (6)$$

Where $\rho = 0.6745$ for artillery constant, $\chi^2(M(N-K))$ is chi-square distribution with its freedom degree being $M(N-K)$, then

$$E(\rho^2 \sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} x_{mki}^2 / M(N-K)) = B_d^2 \quad (7)$$

$$E(\rho^2 \sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} z_{mki}^2 / M(N-K)) = B_f^2 \quad (8)$$

Therefore, the estimated value of dispersion median errors of distance and direction are respectively as follows:

$$\widehat{B}_d = \sqrt{\rho^2 \sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} x_{mki}^2 / M(N-K)} \quad (9)$$

$$\widehat{B}_f = \sqrt{\rho^2 \sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} z_{mki}^2 / M(N-K)} \quad (10)$$

Calculation of confidence intervals of dispersion median errors. The following marks are introduced:

$$S_d = \sqrt{\sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} x_{mki}^2 / M(N-K)} \quad (11)$$

$$S_f = \sqrt{\sum_{m=1}^M \sum_{k=1}^K \sum_{i=1}^{N/K} z_{mki}^2 / M(n-K)} \quad (12)$$

Then, the confidence interval of B_d with its confidence degree being $1-\alpha$ is as follows:

$$\left(\rho \sqrt{M(N-K)} S_d / \sqrt{\chi_{\alpha/2}^2(M(N-K))}, \rho \sqrt{M(N-K)} S_d / \sqrt{\chi_{1-\alpha/2}^2(M(N-K))} \right) \quad (13)$$

And the confidence interval of B_f with its confidence degree being $1-\alpha$ is as follows:

$$\left(\rho \sqrt{M(N-K)} S_f / \sqrt{\chi_{\alpha/2}^2(M(N-K))}, \rho \sqrt{M(N-K)} S_f / \sqrt{\chi_{1-\alpha/2}^2(M(N-K))} \right) \quad (14)$$

Where the $\chi_{\alpha/2}^2(M(N-K))$ and $\chi_{1-\alpha/2}^2(M(N-K))$ are the quantile $\alpha/2$ and $1-\alpha/2$ on chi-square distribution with its confidence degree being $M(N-K)$.

Calculation of confidence degrees of dispersion median errors in given tolerance conditions. In the analysis of test results, hoping that the estimated values of dispersion median errors B_d and B_f is wholly equal to true value is an impossible and unrealistic task, therefore, a certain deviation between estimated value and true value is usually admissible, which is called tolerance, marked as Δ . This value shall be determined according to actual conditions. In case of existence of tolerance, the differences between estimated value and true value shall not exceed the probability of tolerance. The probability is called confidence degree and calculated as follows:

$$f_1(\alpha) = \rho \sqrt{M(N-K)} S_d / \sqrt{\chi_{\alpha/2}^2(M(N-K))} \quad (15)$$

$$f_2(\alpha) = \rho \sqrt{M(N-K)} S_d / \sqrt{\chi_{1-\alpha/2}^2(M(N-K))} \quad (16)$$

Then, the confidence interval of distance dispersion median error with its confidence degree being $1-\alpha$ is $[f_1(\alpha), f_2(\alpha)]$. This formula denotes that according to the existing estimated data B_d , the possibility for falling in intervals of $[f_1(\alpha), f_2(\alpha)]$ is $1-\alpha$. In case the true value B_d truly exists in that interval, the midpoint of the interval may be taken as estimated value, while the differences between estimated value and true value of B_d will be half of that intervals at most. Therefore, set $\Delta_\alpha = 0.5 \times [f_2(\alpha) - f_1(\alpha)]$ as the tolerance with its confidence degree being $1-\alpha$. Based on existing data, if the midpoint of confidence interval is taken as the estimated value of B_d , then, the probability that the differences between true value and estimated value will not exceed Δ_α will be $1-\alpha$.

Based on the above analysis, when a tolerance Δ is given, the problem is transformed into how to solve α , so as to make $\Delta_\alpha = \Delta$, while the $1 - \alpha$ is the confidence degree corresponding to tolerance $\Delta_\alpha = \Delta$.

Calculate number of tests needing to be added for making the dispersion median errors to fall in the given admmissive scope. Section 3.6 and 3.7 mainly deal with the calculation of confidence degrees to which the parameters may reach under the given tolerance conditions, or the calculation of tolerance and confidence intervals estimated by parameters under the given confidence degree conditions. But in reality, the estimation of parameters are needed to be put below the given tolerance with a higher confidence degree, while the data obtained from existing test results are not sufficient to supply information volume. In this case, tests need to be added. The test numbers need to add shall be estimated according to existing test results so as to supply theoretical direction for the following added tests.

Following is a description based on the example of distance dispersion median error B_d . In Section 3.6, the confidence interval of B_d under the confidence degree of $1 - \alpha$ may be denoted as follows:

$$\left(\rho\sqrt{n}S_d / \sqrt{\chi_{\alpha/2}^2(n)}, \rho\sqrt{n}S_d / \sqrt{\chi_{1-\alpha/2}^2(n)} \right) \quad (17)$$

Where, $n = M(N - K)$ is the freedom degree of dispersion error, S_d is the sample average value in the distance direction of dispersion error, and both of them are only related to test data.

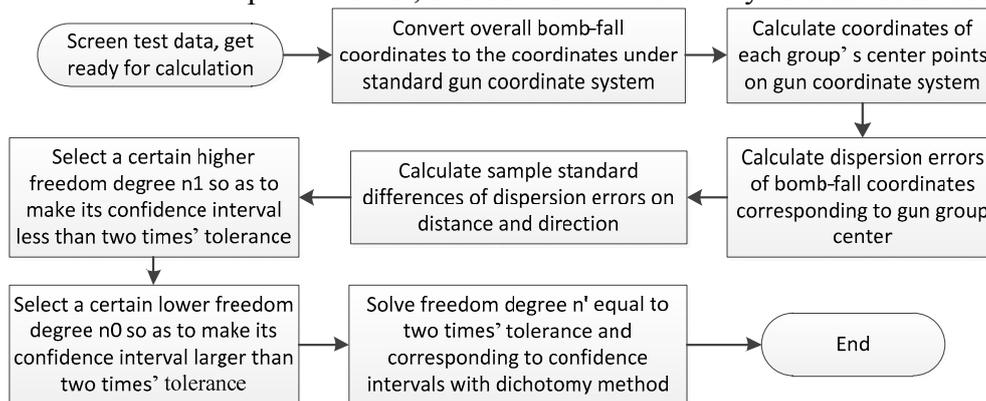


Figure 3 Solution Flow Chart for Added Test Numbers

When confidence degree was determined, α is a constant, then, half of the width of confidence interval is a function of freedom degree, marked as $\Delta(n)$. Known from strong law of large numbers, S_d is the standard deviation tending to dispersion error, therefore, when the new S_d obtained from added tests has no big difference from the original S_d , the width of confidence interval then is mainly determined by the number of added tests, namely the number of newly-increased freedom degree. Based on the above discussion, it is suggested to estimate the number of tests needed so as to ensure the estimated value of B_d falling in the tolerance Δ with its confidence degree being $1 - \alpha$, namely, solving the n' which makes $\Delta(n') = \Delta$ established. While $n' - n$ is the number of tests needing to add. The solving flow chart is shown in Fig. 3.

Estimation of elements median errors. The process of analysis method of elements' distance and direction median errors is similar to that of dispersion error analysis, and the main differences lie in the identification of error sources, here it will not go into detail.

Examples of Test Data Analysis

Taking the indirect aiming test of mortar platoon (equipped with four mortars of a certain type) as an example, the tests of three subjects are conducted covering 4km, 4.8km and 6.2km. The following is an analysis conducted on the subject with 4.8km shooting distance. In this subject, three groups of tests are done with 64 shells shooting per group. The bomb-fall points of three groups are shown in Fig. 4. Fig. 4 has described that optimum width fire shooting may be practiced in mortar platoon, especially seen from the third group of data, the bomb-fall points obviously

present an evenly stacking. Known from confirmation, three rounds of shooting are conducted in third group test, four shells launched in each round of shooting per gun with no changes exerted on sight and shooting directions in the whole process of tests. While in the process of the first two groups of tests, after shooting four shells each gun, the No.2, No.3 and No.4 guns have respectively modified five mils rightward. No detailed information about aiming points were recorded in the tests. Therefore, although the bomb-fall points of the first two groups of shooting tests appeared stacking, the stacking is not evenly spread. It is difficult to judge which shells came from the same gun, therefore, the data obtained from the first two groups of tests cannot be put into use.

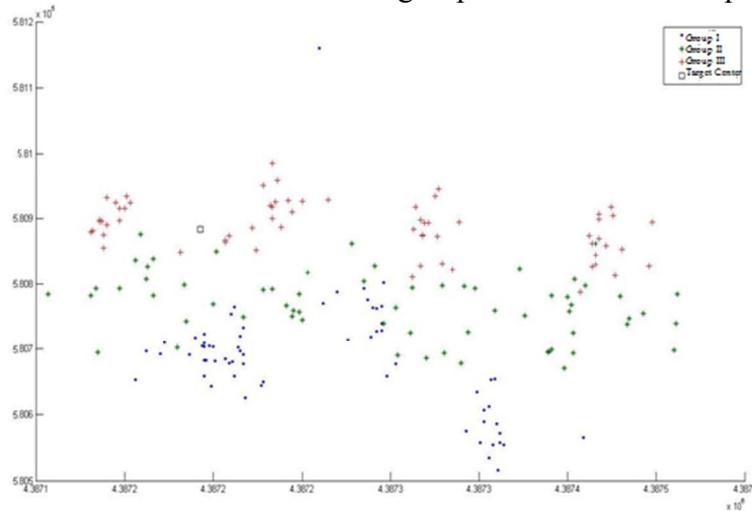


Figure4Dispersion Map for ThreeGroups's Bomb-fall Points

As shown by Figure 4, the bomb-falls in third group's test are classified according to the "Point Suppression Data Acquisition Table" and in the sequence of projecting four shells per gun. In Fig. 5, each heap of bomb-fall points is not distributed to the same gun, therefore, it may be concluded that items are not recorded according to the sequence of shooting. As shown in Fig. 6, each heap of bomb-fall points is distributed to the same gun. By comparing Fig. 5 and Fig. 6, the latter is much closer to reality. The bomb-fall distribution model shown in Fig. 6 will be adopted in the following data analysis.

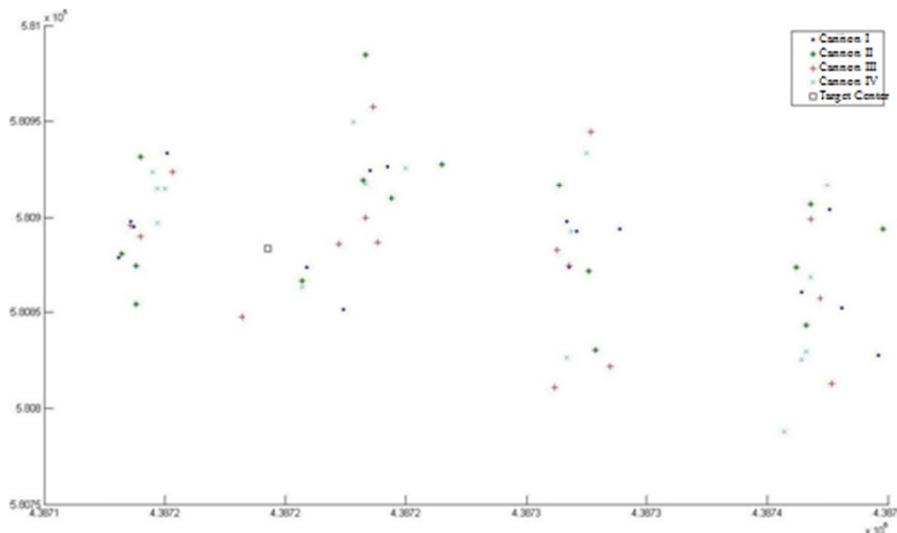


Figure5 Dispersion Map for ThirdGroup's Bomb-fall Points ClassifiedAccording to Guns

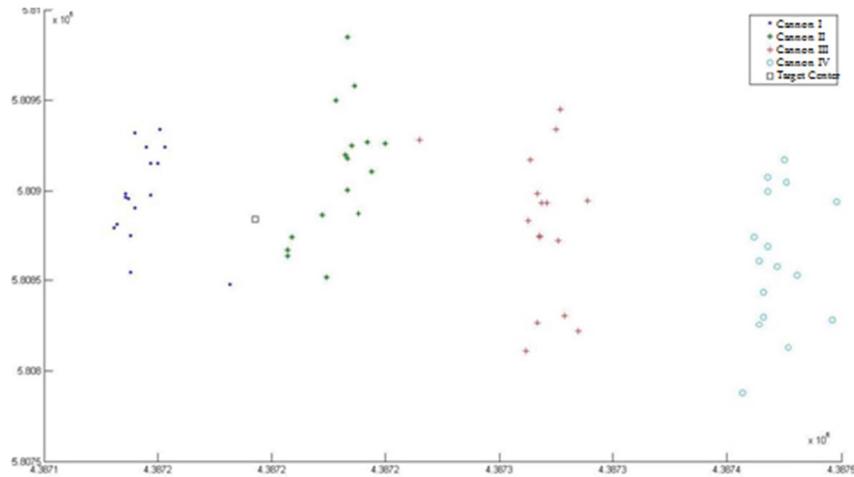


Figure6 Dispersion Map for ThirdGroup's Bomb-fall Points Classified after Rearrangement According to Guns

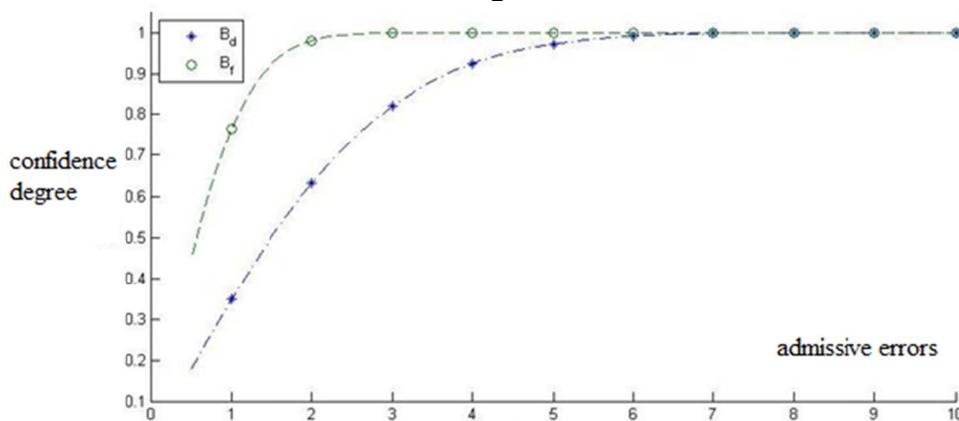


Figure7 Change Curve of Confidence Degree of Dispersion Median Errors

Estimation of dispersion median errors of distance and direction. The dispersion median errors of distance and direction obtained from third group's tests data are respectively $B_d = 24.05\text{m}$, $B_f = 9.07\text{m}$, the curve of confidence degree changing along with the admissible errors is shown in Fig. 7.

Take tolerance $\Delta = 2, 3, 4, 6, 8, 10$, respectively calculate the confidence degrees when $B_d = 24.05\text{m}$ and $B_f = 9.07\text{m}$, obtained results as shown in Table 1.

Table 1 Confidence Degree of Dispersion Errors

$\Delta (m)$	2	3	4	6	8	10
B_d Confidence Degree	0.63	0.82	0.92	0.99	1.00	1.00
B_f Confidence Degree	0.98	1.00	1.00	1.00	1.00	1.00

As shown in Fig. 7 and Table 1, if the existing test data are available, the amount of information will be determined. The higher the acceptable tolerance, the higher the confidence degree of estimated values. Taken B_d as an example, if the tolerance of B_d not exceed 4m may be accepted, the confidence degree of $B_d = 24.05\text{m}$ will be 92%, while if the tolerance of B_d does not exceed 2m, then the confidence degree of $B_d = 24.05\text{m}$ is only 63%.

The confidence degrees of dispersion median errors of distance and direction which may be reached under the conditions of different tolerances are given in Table 1. The following Table 2 has given the calculation results of the confidence intervals of B_d and B_f under the given confidence degrees. Actually, the width of confidence interval is two times that of tolerance.

Table 2. Confidence Interval of Dispersion Errors

Confidence Degree	0.95	0.9	0.85	0.8	0.75	0.7
B_d Confidence Interval	(20.4, 29.3)	(20.9, 28.3)	(21.3, 27.8)	(21.6, 27.3)	(21.8, 27.0)	(22.1, 26.7)
B_f Confidence Interval	(7.7, 11.0)	(7.9, 10.7)	(8.0, 10.5)	(8.1, 10.3)	(8.2, 10.2)	(8.3, 10.1)

Numbers of tests needed to add.The above analysis focuses either upon calculation of estimated value's confidence degree under the conditions of admmissive errors, or upon the calculation of confidence interval (equivalent to calculation of admmissive errors) under the conditions of given confidence degree. In the data analysis of actual tests, the estimated value is at least hoped to fall in a certain tolerance scope under the given confidence degree. But under the circumstance that data are given, the information amount is determined. In case information amount failed to support the estimated values' falling in the scope of admmissive errors with its confidence degree, the information amount must be added, accordingly tests must be increased. The following deals with how to add number of tests according to test data. The number of tests may be obtained by calculating total number of tests minus the existing number of tests. As detailed down to the dispersion median errors B_d and B_f of distance and direction, the total number of shells need to be calculated. The calculation results are shown in Table 3.

Table3 Estimated Number of Shells Need to be Added to B_d and B_f

Tolerance Confidence Degree	2m		3m		4m		6m		8m		10m	
	B_d	B_f										
95%	*286	47	*131	25	*77	18	39	12	25	10	19	9
90%	*203	35	*94	20	56	14	29	10	19	8	15	8
85%	*157	28	*74	16	44	12	24	9	16	8	13	7
80%	*125	24	59	14	36	11	20	8	14	7	11	7
75%	*102	20	49	12	30	10	17	8	12	7	10	6

Totally, 64 shells were used in the Group III Test. In Table 3 and Table 4, the number (marked with "*") larger than 64 denotes the number of tests to be added for making estimated values falling in the scope of tolerance under the given confidence degree. For instance, in Table 3, if trying to make the errors of estimated value $B_d = 24.05m$ less than 2m with the confidence degree being 70%, 20 shells should be increased. As shown in Table 4, the errors of the estimated value $B_f = 9.07m$ obtained from the existing test data will not exceed 95%. This result is consistent with that of Table 1.

Result analysis.By summarizing analysis data of test subjects with 4km, 4.8km and 6.2km, the relationship between the dispersion median errors of mortar of a certain type and the shooting distance is obtained by adopting regression analysis, and shown in Table 4.

Table 4 Relationship of $B_d(m)$ and $B_f(m)$ with the Changes of Gun Range D

Gun Range D	4km		4.5km		5km		5.5km		6km		6.2km	
	B_d	B_f										
Testl analysis data	20.5	6.3	22.65	8.1	24.8	9.9	26.95	11.7	29.1	13.5	29.96	14.22
Tactical and technical index data	26.7	13.3	30.0	15.0	33.3	16.7	36.7	18.3	40.0	20.0	41.3	20.7
Firing table data	14.9	3.7	16.9	4.1	18.9	4.4	20.9	4.9	23	5.5	23.8	5.9

Under such conditions that tolerance Δ is equivalent to 10% of tactical and technical index, the confidence degrees of test analysis obtained from teating data samples are all larger than 80%. The correspondence of test analysis data with the tactical and technical index data and firing table data has proved that the analysis results may be used in the data input of operational effectiveness evaluation and model verification.

Summary

In the test of equipment operation, firing accuracy is one of the important test subjects, which may supply effectiveness evaluation of weapon and firing control with reliable test data. This paper has put forward ideas for analyzing firing accuracy under the conditions of incomplete test data records by combining the characteristics of gun firing and by separating firing errors. On the basis of determining the attribution of bomb-fall by using method of clustering, the analyzing method for firing accuracy based on chi-square distribution was established and the test data analysis was

conducted on the test data of one type of weapon. The results of analysis show that the method was effective and may be used for conducting test data analysis as effective means.

Acknowledgement

In this paper, the research was sponsored by the Defense Pre-research Project (Project No.513300504) and the Pre-research Project Fund of Key Laboratory of Science and Technology for National Defence (Project No.9140C900401130C90001).

References

- [1] Firing Accuracy Analysis of Vehicle-mounted Howitzer, Cao Ning, Nanjing University of Science & Technology, 2013.
- [2] Research on Testing Method for Firing Accuracy of MLRS, Yao Zhi Jun, Nanjing University of Science & Technology, 2008.
- [3] Principle for Evaluation of Firing Efficiency (first edition). Cheng Yun Men, Beijing: People's Liberation Army Publishing House, 1986.
- [4] Research on Simulation of Gun Firing Accuracy, Yang Bo Zhong, Chen Yun Sheng and Cheng Xian Jin, Journal of Nanjing University of Science & Technology, 2002, 26(2): 127-130.
- [5] Research on Evaluation Method for System Accuracy of Modern SP Gun Howitzer, Tang Xue Mei and Yi Qun Zhi, Journal of Gun Launch & Control, 2001, 1: 28-33.