Aerospace Science and Technology 58 (2016) 369–379

Contents lists available at ScienceDirect

Aerospace Science and Technology

www.elsevier.com/locate/aescte

Weapon systems accuracy evaluation using the error spectrum

Weishi Peng [a,b,*,1], YangWang Fang [b,2], Renjun Zhan [a,3], Weijia Wang [b,1]

[a] School of Materiel Engineering, People Armed Police Engineering University, Xi'an, 710086, China
[b] School of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an, 710038, China

ARTICLE INFO

Article history:
Received 3 June 2016
Received in revised form 27 August 2016
Accepted 31 August 2016
Available online 6 September 2016

Keywords:
Weapon systems accuracy evaluation
Error spectrum
Bootstrap method
Multiple attribute decision making

ABSTRACT

Weapon systems accuracy is a vitally important tactical and technical index, which has a significant impact on the process of the weapon systems design and test. Based on different accuracy criteria, we proposed a new evaluation method for the weapon systems accuracy. First, a new bootstrap method, using the correlation coefficient criterion, is presented to evaluate a weapon systems accuracy when the sample size n satisfied $n \geq 10$. Second, measures for weapon systems accuracy are introduced to represent the different performance of weapon systems. Furthermore, an attributes matrix is proposed to combine the above measures. Third, inspired by the Pitman's closeness measure, we designed a weapon systems accuracy attributes competition matrix to contain the entire pairwise competition results, and then its positive eigenvector is employed to the weapon systems accuracy evaluation. Finally, a numerical example is provided to illustrate the effectiveness and feasibility of the proposed method. It is shown that the new evaluation method can not only evaluate the weapon systems accuracy before and after improvement, but also rank the accuracy of the designed weapon systems to the evaluated weapon systems.

1. Introduction

Weapon systems accuracy represents the dispersion characteristic between the target and the warheads land, which plays a more and more important role in the design and the evaluation of a weapon system (see, e.g. [1–6]). In practical applications, several accuracy criteria were proposed for different weapon systems, such as accuracy, precision, system error, firing accuracy and CEP. The accuracy is the degree of closeness of the average impact point to the target, while the precision accounts for the dispersion of the impact points. System error is an exact value, which can be estimated based on several experiments in the same conditions. Firing accuracy is used to measure the firing error of the gun system, and it usually includes accuracy and precision. CEP is a hit accuracy criterion, which is often characterized with a CEP. CEP is the radius of a circle containing 50% of the hitting probability and centered at the target [7]. Therefore, CEP is used to describe the firing error of the weapon system [8].

However, there exist three challenges in the process of the weapon systems design and test. One of the challenges, in the process of the weapon systems test, is that we want to know the weapon systems accuracy before and after improvement. To solve the above problems, [9] presented the sequential probability ratio test to make statements concerning the unknown mean of a Gaussian process, which requires an expected number of observations considerably smaller than the fixed number of observations needed by the current most powerful test methods. Furthermore, the latter provides a control over the errors of the first and second kinds to exactly the same extent (has the same α and β) as the sequential test. The second challenge is that we also need to rank the accuracy of the designed weapon systems with respect to one another. Several methods have been proposed to deal with this problem, such as the AHP [10], fuzzy-AHP [1], fuzzy arithmetic operations [2], ranking fuzzy numbers [3], response surface method using grey relation analysis [11] and intuitionistic fuzzy sets [12]. Last but not least is the small sample problem. As mentioned above, the assessment of weapon systems accuracy is performed based on the test data of weapon systems. However, the classical statistic approaches which are based on the large sample are useless because a small number of experiments are usually performed. Therefore, the system simulation technology and the semi-physical simulation were proposed to increase the number of the test data (see, e.g. [4,18–24]).

* Corresponding author.
E-mail addresses: peng_weishi@163.com (W. Peng), ywfang2008@sohu.com (Y. Fang), zhanrenjun@aliyun.com (R. Zhan), 519089154@qq.com (W. Wang).
1 Doctoral Student, Department of Aerospace Weapon Engineering.
2 Professor, PhD, Department of Aerospace Weapon Engineering.
3 Professor, PhD, Department of Non-lethal Engineering.

http://dx.doi.org/10.1016/j.ast.2016.08.032
1270-9638/

Nomenclature

Abbreviations

CEP	Circular Error Probability
AHP	Analytic Hierarchy Process
AEE	Average Euclidean Error
RMSE	Root Mean Square Error
HAE	Harmonic Average Error
GAE	Geometric Average Error
ME	Median Error
EM	Error Mode
IMRE	Iterative Mid-Range Error
ES	Error Spectrum
DES	Dynamic Error Spectrum
CDF	Cumulative Distribution Function
PDF	Probability Density Function
PMF	Probability Mass Function
AES	Area-Error Spectrum
RES	Range-Error Spectrum
WSAA	Weapon Systems Accuracy Attributes
WSAAC	Weapon Systems Accuracy Attributes Competition

Symbols

H_1	Alternative hypothesis
α	The probability of an error of the first kind (rejecting H_0 when H_0 is true)
H_0	Null hypothesis
β	The probability of an error of the second kind (accepting H_0 when H_1 is true)
$\boldsymbol{x}$	Original data
$\boldsymbol{x}^*$	Resampling data
$\boldsymbol{x}^s$	Simulation data
$\rho(\cdot)$	Correlation coefficient criterion
ε	The threshold of $\rho(\cdot)$ in the new bootstrap method
n	The number of the original data
m	The number of histogram
B	The number of the resampling data
n_s	The number of the performed simulation
$\rho(\boldsymbol{X},\boldsymbol{Z})$	The correlation between X and Y (where X and Y are downrange and cross-range misses, respectively)

In this paper, the weapon systems accuracy, based on the test data, is characterized by several statistics which reflect the different aspects of the weapon systems. More concretely, among the above accuracy criteria, the accuracy is the mean of the error data, and the precision is the standard deviation of the error data. However, the mean (i.e., the AEE) and standard deviation (the approximation of the RMSE) are easily dominated by the outliers [13]. To solve the above problem, several robust statistics were proposed (see, e.g. [13–17]), such as the HAE, the GAE, the ME, the EM and the IMRE. Among these, the HAE is suit to evaluate hit-or miss since it is dominated by the small error terms [16]. Furthermore, the GAE is a more typical value than AEE because it is less affected by extreme values. The ME is the middle error term or the arithmetic average of the middle two error terms; the error mode is the location of the highest peak of the histogram for the given error. Obviously, the ME and EM are decided by the "normal" or "typical" error, which are still not robust. Thus, the IMRE [17], a novel measure of central tendency, is presented to overcome the drawbacks of the median. In [16], the RMSE and AEE are pessimistic since they focus on the bad performance; the HAE is optimistic because it focused on the good performance; the GAE, ME, EM and IMRE are balanced since they are neither dominated by larger error nor affected by small error. Obviously, different statistics reflect different aspects of performance. All of the above-listed statistics can reflect only one aspect of the weapon systems accuracy. Thus, three comprehensive performance measures—the ES, desirability level, and relative concentration and deviation measures were proposed in [25–27]. Among these metrics, the ES can reveal more information about the estimation because it is an aggregation of several incomprehensive metrics. Since then, most existing researches have focused on the improvements of the ES (see, e.g. [30–33]).

The main contribution of this work is that a new evaluation method using the ES is proposed to evaluate the weapon systems accuracy according to the test data. First, a new bootstrap method is proposed to increase the number of the test data. Further, based on the correlation coefficient criterion, the new subsamples replicating by the new bootstrap method have faster convergence speed and lower square error than the bootstrap method. Second, several measures are introduced to the weapon systems accuracy, then, based on the above measures, a WSAA matrix is proposed to represent the performance of the weapon systems accuracy. Third, according to the WSAA matrix, let any two columns compared with each other, a WSAAC matrix is designed to store the competition results, and its positive eigenvector is employed to find whether the tested weapon satisfied design requirement or not.

Considering the aforementioned introduction, in this paper, a new evaluation method for the weapon systems accuracy is proposed. The main novelty in this paper is summarized as follows. First, the proposed evaluation method can also be used to evaluate weapon systems accuracy under the small sample case based on the new bootstrap method. Second, more attributes of the weapon systems will be considered by using the error spectrum. Finally, the above evaluation method can not only evaluate the similar weapon systems but also to rank different types according to WSAAC matrix.

2. The new bootstrap method to resampling

In practical applications, the frequency of weapon systems test is getting down because of the high cost of the test, which leads to the limitation of the tested weapon systems observations. Therefore, before performing the new evaluation method, a new resampling method is proposed to increase the original sample. To the best of our knowledge, most existing research has focused on the small sample problem (see [18–24]). Among these, the most popular ones are bootstrap method introduced by [18] and Bayesian bootstrap method (random weighing method) proposed by [22] and [23]. In this paper, we utilized bootstrap method to resampling.

2.1. Bootstrap method

Bootstrap method is described in [18,19], and we quote it here. Based on the available limited samples $\boldsymbol{x} = \{x\}_{i=1}^{n}$, bootstrap method does provide an efficient way of estimating the distribution of statistical parameter by using the resampling technique, which has been applied in many applications [20,21]. The main idea is to replicate several sets of bootstrap samples $\{\bar{\boldsymbol{x}}^{*j}\}_{j=1}^{B}$ with replacement from the original data $\boldsymbol{x} = \{x\}_{i=1}^{n}$. A key issue in bootstrap method is how to perform the replication. Usually, Monte Carlo method is applied to replicate the new subsamples $\boldsymbol{x}^{*j} = (x_1^{*j}, x_2^{*j}, \cdots, x_n^{*j})$, which is carried out as follows:

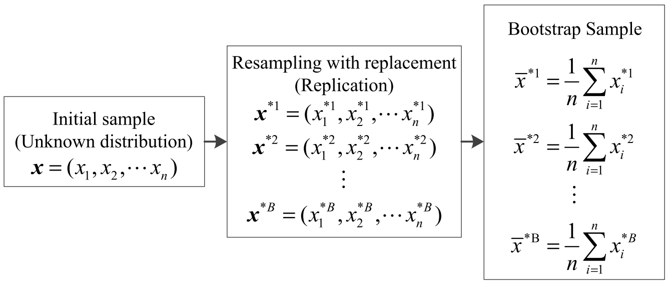


Fig. 1. Schematic representation of bootstrap method.

Step 1: Generate a random number R, which follows the uniform distribution $U(0, M)$, M is a positive integer satisfying $M \gg n$;

Step 2: Let $p = \mathrm{mod}(R, n)$, where $\mathrm{mod}(\cdot)$ is the mod function. If $\mathrm{mod}(R, n) = 0$, let $p = 1$, if $\mathrm{mod}(R, n) > n$, let $p = n$;

Step 3: Let $x_1^{*1} = x_p$, then repeat Steps 1 and 2 for n times to obtain a new sample $\boldsymbol{x}^{*1} = \{x_i^{*1}\}_{i=1}^n$.

Then repeat Steps 1–3 for B times, and denote $\bar{\boldsymbol{x}}^{*j}$ the average of each replication sample $\{x_i^{*j}\}_{i=1}^n$, we have the new bootstrap sample $\{\bar{\boldsymbol{x}}^{*j}\}_{j=1}^B$.

Obviously, the advantage of this method is that it only requires the initial set of samples, as shown in Fig. 1. Moreover, the number of the bootstrap resampling is chosen to be large enough so that it does not affect the quality of the results. As shown in [20], the number of resample B is typically taken from 500 to 5000. In order to ensure the accuracy of the bootstrap method and to save computation time, we let the number of resampling equals to 1000 in this paper.

Furthermore, as shown in Fig. 1, the key step in bootstrap method is the replication. Although Monte Carlo method is regarded as a significant technical approach in solving the above replication problem, it has some limitations. To see this, here is an example: assume that the initial set of samples are (r_1, r_2, r_3, r_4), then bootstrap method randomly selects five samples are (r_1, r_1, r_3, r_4), (r_1, r_3, r_3, r_4), (r_1, r_2, r_2, r_4), (r_1, r_1, r_3, r_3), (r_2, r_2, r_4, r_4), then, the average of each replication sample is calculated to design a new sample $(\bar{r}_1^*, \bar{r}_2^*, \bar{r}_3^*, \bar{r}_4^*, \bar{r}_5^*)$. We can see from the replication sample (r_1, r_1, r_3, r_3) that only two points information (i.e., (r_1, r_3)) have been used. In other words, much information has been lost in the new replication sample. To solve the above problem, a new bootstrap method is proposed based on the correlation coefficient criterion.

2.2. The new bootstrap method

As discussed above, the correlation coefficient criterion is utilized to quantify the closeness between the initial set $\boldsymbol{x} = \{x\}_{i=1}^n$ and the new replication set $\boldsymbol{x}^{*j} = \{x_i^{*j}\}_{i=1}^n$, which is defined as

$$\rho\big(f(x); f\big(x^{*j}\big)\big) = \frac{\int f(x) f(x^{*j}) dx}{[\int f(x)^2 dx \int f(x^{*j})^2 dx]^{1/2}} \tag{1}$$

where $f(x)$ and $f(x^{*j})$ are the PDF of $\boldsymbol{x} = \{x\}_{i=1}^n$ and $\boldsymbol{x}^{*j} = \{x_i^{*j}\}_{i=1}^n$, respectively.

Since the PDF of $\boldsymbol{x} = \{x\}_{i=1}^n$ and $\boldsymbol{x}^{*j} = \{x_i^{*j}\}_{i=1}^n$ are rarely available, here, we use the histograms instead of the corresponding discretized PDF. Therefore, (1) is approximated to

$$\rho\big(f(x); f\big(x^{*j}\big)\big) \approx \frac{\sum_{k=1}^m h(x_k) h(x_k^{*j})}{[\sum_{k=1}^m h(x_k)^2 \sum_{k=1}^m h(x_k^{*j})^2]^{1/2}} \tag{2}$$

where $h(\cdot)$ is a histogram function with m bins $(m < n)$; x_k is the center of the k-th bin ($\{x_k\}_{k=1}^m$ is the set of x_k) and x_k^{*j} is still the center of the k-th bin ($k = 1, 2, \cdots m$). It was assumed that the above correlation coefficient criterion is positive definite: $\rho(f(x); f(x^{*j})) = 1$ if $f(x) = f(x^{*j})$ or $h(x) = h(x^{*j})$ and $\rho(f(x); f(x^{*j}))$ has the standard range: $0 \le \rho(f(x); f(x^{*j})) \le 1$.

Obviously, keeping on iterating through Steps 1–3 in bootstrap method, we will obtain the new replication sample unless the above correlation coefficient criterion is greater than ε. Here, ε a real number satisfying $\varepsilon \in [0, 1]$, typically determined by the user. Therefore, a detailed pseudocode description of the new bootstrap method is summarized as follows:

Step 1: Generate a random number R, which follows the uniform distribution $U(0, M)$, M is a positive integer satisfying $M \gg n$;

Step 2: Let $p = \mathrm{mod}(R, n)$, if $\mathrm{mod}(R, n) = 0$, let $p = 1$, if $\mathrm{mod}(R, n) > n$, let $p = n$;

Step 3: Let $x_1^{*1} = x_p$, then, repeat Steps 1 and 2 for n times to obtain a new sample $\boldsymbol{x}^{*1} = \{x_i^{*1}\}_{i=1}^n$.

Step 4: calculate the correlation coefficient criterion between new sample $\boldsymbol{x}^{*1} = \{x_i^{*1}\}_{i=1}^n$ and the initial set $\boldsymbol{x} = \{x\}_{i=1}^n$. Then, repeat Steps 1–3 until convergence: $\rho(f(x); f(x^{*j})) \ge \varepsilon$. If an appropriate stopping condition holds then output the new sample $\boldsymbol{x}^{*1} = \{x_i^{*1}\}_{i=1}^n$.

Step 5: Repeat Steps 1–4 for B times, the new bootstrap sample $\{\bar{\boldsymbol{x}}^{*j}\}_{j=1}^B$ are obtained.

In this paper, for the weapon systems accuracy evaluation, let (x_0, z_0) be the target position, assume that we fire a series of weapons as the target and mark where the warheads land $(\boldsymbol{X}, \boldsymbol{Z}) = (x_1, z_1), (x_2, z_2), \cdots, (x_n, z_n)$, which are independent identically distributed; then the downrange and cross-range misses are $\tilde{x} = \{x_i - x_0\}_{i=1}^n = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ and $\tilde{z} = \{z_i - z_0\}_{i=1}^n = (\tilde{z}_1, \tilde{z}_2, \cdots, \tilde{z}_n)$, respectively. Then, denote $(\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{Z}}) = (\tilde{x}_1, \tilde{z}_1), (\tilde{x}_2, \tilde{z}_2), \cdots, (\tilde{x}_n, \tilde{z}_n)$ as the corresponding fall point error.

According to (2), it is inconvenience to calculate the correlation coefficient criterion between the test data $(\boldsymbol{X}, \boldsymbol{Z})$ and the replication data $(\boldsymbol{X}^*, \boldsymbol{Z}^*)$. Therefore, the relative error is proposed to transform the above two data.

Usually, we use the following relative distance between the target and warhead.

$$c_i = \sqrt{x_i^2 + z_i^2} \quad i = 1, 2, \cdots n \tag{3}$$

Similarly, we can define a relative distance error between the target and warhead

$$\tilde{c}_i = \sqrt{\tilde{x}_i^2 + \tilde{z}_i^2} \quad i = 1, 2, \cdots n \tag{4}$$

So, $(\tilde{x}_1, \tilde{z}_1), (\tilde{x}_2, \tilde{z}_2), \cdots, (\tilde{x}_n, \tilde{z}_n)$ can be equivalently written as the relative distance error $(\tilde{c}_1, \tilde{c}_2, \cdots, \tilde{c}_n)$. Then, we obtain the new observations $(\tilde{c}_1^*, \tilde{c}_2^*, \cdots, \tilde{c}_B^*)$ based on the above new bootstrap method. Similarly, according to (4), a relative distance error $\tilde{c}_j^*$ correspond to a fall point error $\{(\tilde{x}_i, \tilde{z}_i)\}_{i=1}^n$. Thus, the replication data are given by

$$\big\{\big(\tilde{x}_j^*, \tilde{z}_j^*\big)\big\}_{j=1}^B = \left\{\left(\frac{1}{n}\sum_{i=1}^n \tilde{x}_i, \frac{1}{n}\sum_{i=1}^n \tilde{z}_i\right)\right\}_{j=1}^B \tag{5}$$

Essentially, we defined the relative distance or the relative distance error because of the following reasons. First, it is not convenient to calculate the correlation coefficient criterion for the two-dimensional test data. Second, as shown in Fig. 1, the new bootstrap method is very easy to be implemented after the two-dimensional test data combined to one-dimensional data. Third, we can see that the relative distance or the relative distance represents the whole error of the weapon system fall point. Furthermore, the relative distance error can be applied in the error spectrum directly. Hereafter, we utilized the relative distance error data ($\tilde{\boldsymbol{c}}^* = (\tilde{c}_1^*, \tilde{c}_2^*, \cdots, \tilde{c}_B^*)$) to represent the weapon systems accuracy.

3. Measures for the weapon systems accuracy evaluation

3.1. Average Euclidean error and root mean square error

According to the new observations, there are several competitive classes of statistics in the weapon systems accuracy evaluation (see, e.g. [13,14,16]). Among these, the two most popular ones are the mean and the variance. In fact, the mean is the AEE, which is defined as

$$\mathrm{AEE}(\tilde{\boldsymbol{c}}^*) = \bar{c}^* = \frac{1}{B}\sum_{i=1}^{B}\tilde{c}_i^* \tag{6}$$

And the variance is given by

$$\hat{\sigma}_{\tilde{\boldsymbol{c}}^*}^2 = S_{\tilde{\boldsymbol{c}}^*}^2 = \frac{1}{B-1}\sum_{i=1}^{B}(\tilde{c}_i^* - \bar{c}^*)^2 \tag{7}$$

In order to unify dimension with the AEE, the variance is usually replaced by the RMSE, i.e.,

$$\mathrm{RMSE}(\tilde{\boldsymbol{c}}^*) = \left[\frac{1}{B}\sum_{i=1}^{B}(\tilde{c}_i^*)^2\right]^{1/2} \tag{8}$$

As we all know, the main nice feature of the RMSE is that it is the most nature finite-sample approximation of the standard error $\sqrt{E[\cdot^2]}$, which is closely related with standard deviation.

In fact, the above statistics are proved to be unbiased, consistent and effective without system errors and outliers in the observation. In particular, with the Gaussian distribution, the mean and variance are the optimal estimation. However, it is usually existed the outliers in practical application, which leads to the limitation of the mean and variance. In other words, whatever in the large or small sample size, they have a common drawback, i.e., the mean and variance are easily dominated by the outliers. For example, if all 100 terms of weapon system errors are around 1 except for one terms of 500, then RMSE ≈ 50, AEE $= 5.99$.

To solve the above problem, several robust statistics were proposed to complement for the above traditional statistics, such as the harmonic average error, the geometric average error, the median error, the error mode and the iterative mid-range error, which are defined as in the following.

3.2. Geometric average error

Unfortunately, the AEE and RMSE are still dominated by large error terms although the domination is alleviated. To avoid this limitation, the GAE is presented in [16]:

$$\mathrm{GAE}(\tilde{\boldsymbol{c}}^*) = \left(\prod_{i=1}^{B}\tilde{c}_i^*\right)^{1/B} = \exp\left[\frac{1}{B}\sum_{i=1}^{B}\ln(\tilde{c}_i^*)\right] \tag{9}$$

Clearly, the GAE is a balance measure because it is less affected by extreme values.

3.3. Harmonic average error

Certainly, in case small errors are of interest, the HAE may be a good choice, this is defined by

$$\mathrm{HAE}(\tilde{\boldsymbol{c}}^*) = \left(\frac{1}{B}\sum_{i=1}^{B}(\tilde{c}_i^*)^{-1}\right)^{-1} \tag{10}$$

Clearly, the HAE is an optimistic measure, which is suit to answer the hit-or-miss question. However, all of the above measures, i.e., the AEE, the RMSE, the GAE and the HAE, can only reflect one part of the performance.

3.4. Median error, error mode and iterative mid-range error

In [16], the RMSE and AEE are pessimistic since they focus on the bad performance; the HAE is optimistic because it is focused on the good performance; the GAE is balanced since it is neither dominated by larger error nor affected by small error. In addition, there exist some other robust statistics, such as the ME, the EM and the IMRE. The ME is the middle error term or the arithmetic average of the middle two error terms, which is defined as

$$\mathrm{ME}(\tilde{\boldsymbol{c}}^*) = \begin{cases} \tilde{c}_{B/2}^* & \text{if } B \text{ is an odd number} \\ \frac{\tilde{c}_{B/2}^* + \tilde{c}_{(B+1)/2}^*}{2} & \text{if } B \text{ is an even number} \end{cases} \tag{11}$$

Obviously, the ME is dominated by the middle error.

For some case, we concentrate on the most frequently error. Therefore, the EM is proposed to solve the above problem, which is the location of the highest peak of the histogram for the given error. Obviously, the ME and EM are decided by the “normal” or “typical” error (i.e., the ME is dominated by the middle error and the EM affected by the location of the highest peak of the histogram for the given error), which are still not robust. Thus, the IMRE, a novel measure of central tendency, is presented to overcome the drawbacks of the median [17]

$$\mathrm{IMRE}(\tilde{\boldsymbol{c}}^*) = \frac{\min(\tilde{\boldsymbol{c}}^*) + \max(\tilde{\boldsymbol{c}}^*)}{2} \tag{12}$$

To see this, here is an example of how to calculate IMRE. Given an error set $\boldsymbol{e}_5 = \{0, 1, 3, 4, 5\}$, compute the mid-range value of $\mathrm{IMRE}_5^0 = \frac{\min(\boldsymbol{e}) + \max(\boldsymbol{e})}{2} = \frac{0+5}{2} = 2.5$, then the mid-range value of IMRE_5^0 is applied to replace the minimum and maximum, i.e., $\boldsymbol{e}_4 = \{1, 2.5, 3, 4\}$, furthermore, $\mathrm{IMRE}_4^1 = \frac{\min(\boldsymbol{e}_1) + \max(\boldsymbol{e}_1)}{2} = \frac{1+4}{2} = 2.5$, then $\boldsymbol{e}_3 = \{2.5, 2.5, 3\}$ and $\mathrm{IMRE}_3^2 = \frac{\min(\boldsymbol{e}_2) + \max(\boldsymbol{e}_2)}{2} = \frac{2.5+3}{2} = 2.75$, continuing this process, finally, $\boldsymbol{e}_2 = \{2.5, 2.75\}$ and $\mathrm{IMRE}_2^3 = \frac{\min(\boldsymbol{e}_3) + \max(\boldsymbol{e}_3)}{2} = \frac{2.5+2.75}{2} = 2.625$, thus, $\boldsymbol{e}_1 = \{2.625\}$ and $\mathrm{IMRE}(\boldsymbol{e}_5) = 2.625$.

Unfortunately, all of the above metrics can only reflect one part of the estimator performance. Three comprehensive performance measures, i.e. the ES, desirability level, and relative concentration and deviation measures were proposed in [25–27]. Among these metrics, the ES can reveal more information because it is an aggregation of many incomprehensive metrics, as shown in the following part.

3.5. Error spectrum

As proposed in [26], let $\boldsymbol{e} = ||\tilde{\boldsymbol{c}}^*||_p$, where $||\cdot||_p$ can be taken as 1-norm or 2-norm. Then, for $-\infty < r < +\infty$, the ES for the weapon systems accuracy evaluation is defined as,

$$S(r) = \left(E\left[\boldsymbol{e}^r\right]\right)^{1/r} = \int e^r dF(e)$$
$$= \begin{cases} (\int e^r f(e)de)^{1/r}, & \text{continuous } e \\ (\sum e_i^q p_i)^{1/r}, & \text{discrete } e \end{cases} \tag{13}$$

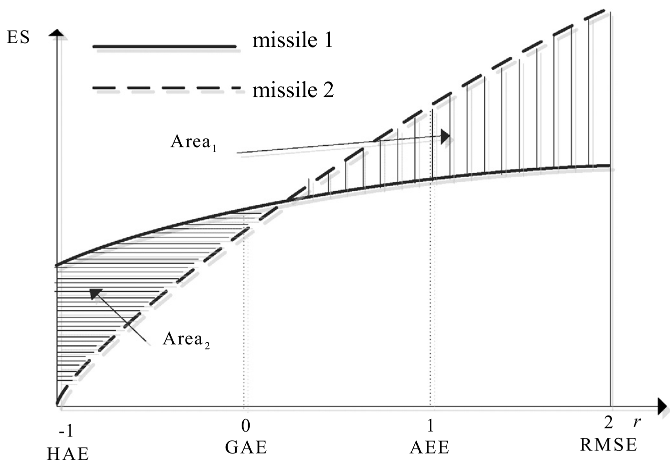


Fig. 2. The ES performance curves of two weapons fall-points accuracy.

where $F(e)$, $f(e)$ and p_i are the CDF, the PDF and the PMF, respectively. The ES has many nice important properties for performance evaluation, which can be found in [14] and [16].

From (13), it is clear that the ES includes the above several metrics as special cases when r is set to some specific values:

(a) $r = 2$, then $S(2) = (E[\boldsymbol{e}^2])^{1/2}$. For a discrete e_i, $S(2) = \text{RMSE}$.

(b) $r = 1$, then $S(1) = E[\boldsymbol{e}]$. For a discrete e_i, $S(1) = \text{AEE}$.

(c) $r = 0$, then $S(0) = \lim_{r\to 0} S(r) = \exp(E[\ln \boldsymbol{e}])$. For a discrete e_i, $S(0) = \text{GAE}$.

(d) $r = -1$, then $S(-1) = 1/E[1/\boldsymbol{e}]$. For a discrete e_i, $S(-1) = \text{HAE}$.

In view of this, the notation r used in this paper is a real number satisfying $r \in [-1, 2]$.

Especially, for a discrete $\{e_i\}_{i=1}^B$, the ES can be approximately calculated as follows [14,31]

$$S(r) \approx \begin{cases} [\frac{1}{B}\sum_{i=1}^B (e_i)^r]^{1/r} & r \neq 0 \\ [\prod_{i=1}^B e_i]^{1/n} & r = 0 \end{cases} \tag{14}$$

It can be seen from (13) that an ES performance curve is obtained to describe the weapon systems accuracy. To see this, Fig. 2 shows that the ES curves of two evaluated missiles.

Furthermore, we can see that the traditional metric RMSE is just a point on the ES performance curve with $r = 2$. Thus, the above ES performance curve reflects more information because it includes the RMSE, the AEE, the GAE and the HAE. However, the above performance curve has some limitations and drawbacks. On one hand, its calculation without the error distribution is not easy, though in [7] (a further development of [4]), the authors provided analytical formulae for the computation of the ES when the error distribution is given. To overcome this problem, we proposed two approximation algorithms, i.e., the Gaussian mixture and power means error method to calculate the ES performance curve [31]. On the other hand, it is difficult to say which weapon performs better when their ES performance curves intersect with each other. To overcome this problem, the DES was presented in [8,9], which is in fact the average height of the ES performance curve. However, it is still difficult to decide which weapon performs better, when one weapon's DES equals the other. To improve the DES, we proposed two metrics, i.e., the AES and RES in [32] and [33]. For convenience, we quote it in the following.

3.6. Area-error spectrum and range-error spectrum

As we know, the concept of the AES was informally put forward in reference [30]. Here, we proposed two areas of the ES for estimation performance evaluation in [32] and [33], which can also be applied in the weapon systems accuracy evaluation.

The first one is the area under the ES performance curve, which is defined as

$$\text{AES} = \int S(r)dr \tag{15}$$

In this paper, the AES is given by

$$\text{AES} = \int_{-1}^{2} S(r)dr \tag{16}$$

Furthermore, for a discrete $r_i, r_i \in \{r_i\}_{i=1}^m$, the AES can be approximated to

$$\text{AES} = \int_{-1}^{2} S(r)dr \approx \frac{r_m - r_1}{m}\sum_{i=1}^{m} S(r_i) \tag{17}$$

It is clearly seen from (17) that the smaller the area is, the better the performance of a weapon is. Furthermore, the area under the ES performance curve reflects how much better or worse one weapon's accuracy outperforms the other, as shown in Fig. 1.

Furthermore, we presented a range-error spectrum (RES) in [32], which is designed to quantify the concentration of the errors of an objective. Let $S(r) = f(e, r)$, we have [25]

$$g\big(S(r)\big) = f^{-1}(e, r) \tag{18}$$

Similar to (15), the second area, i.e., the RES, is defined as

$$\text{RES} = \int_{S(r_1)}^{S(r_m)} g\big(S(r)\big)dS(r) \approx \frac{S(r_m) - S(r_1)}{m}\sum_{i=1}^{m} g\big(S(r_i)\big) \tag{19}$$

Since $g(S(r_i))$, the inverse function of f, is hard to be obtained analytically, (19) can be approximated as:

$$\text{RES} \approx \big[S(r_m) - S(r_1)\big]\frac{1}{m}\sum_{i=1}^{m}(r_i - r_1) \tag{20}$$

Moreover, we have

$$S(r_m = 2) = \text{RMSE}, \qquad S(r_1 = -1) = \text{HAE} \tag{21}$$

Hence, the above equation is reduced to

$$\text{RES} = (\text{RMSE} - \text{AEE}) \times \frac{1}{m}\sum_{i=1}^{m}(r_i - r_1) \tag{22}$$

Since $\sum_{i=1}^m (r_i - r_1)/m$ is a constant, it is clear that the RES reflects the flatness of the ES performance curve. As mentioned in [32], this property is characterized by the concentration of the errors.

As can be seen in (17) and (22), number m represents the attributes of weapon systems accuracy. For instance, if the number of indices over $\{r_i\}_{i=1}^m$ equals 4 (i.e., $m = 4$), within the interval $r_i \in [-1, 2]$, we have $r_i \in \{-1, 0, 1, 2\}$. That is, the HAE, the GAE, the AEE and the RMSE are altogether applied to describe a weapon's attributes. In some situations, however, we may want to focus on the other error metrics. For this purpose, we set m to different values corresponding to different sets $\{r_i\}_{i=1}^m$. Obviously, different m reflects different attribute of a weapon.

For convenience, we define an attributes distance within the interval $[-1, 2]$ to describe a weapon's attributes:

$$D = \frac{r_m - r_1}{m} \tag{23}$$

As stated before, the above measures reflect all the information of weapon systems accuracy. Therefore, in the following, we propose a WSAA matrix concerning about all the aspects of weapon systems accuracy.

4. Weapon systems accuracy attributes matrix

Assume that we have N weapons to be evaluated and a desired weapon (designed in theory). For each weapon, we obtained a corresponding test data (X_j, Z_j), $j = 1, \cdots N$. Then we have the relative distance error data $(\{\tilde{\boldsymbol{c}}_j^*\}_{j=1}^N)$ of the weapon systems accuracy according to the new bootstrap method. Furthermore, the above measures are calculated by the relative distance error data. Finally, the WSAA matrix is defined as

$$A^{\mathrm{WSAA}} = \begin{bmatrix} S_0(r_m=2)=\mathrm{RMSE}_0 & S_1(r_m=2)=\mathrm{RMSE}_1 & \cdots & S_n(r_m=2)=\mathrm{RMSE}_N \\ \vdots & \vdots & \ddots & \vdots \\ S_0(r_p=1)=\mathrm{AEE}_0 & S_1(r_p=1)=\mathrm{AEE}_1 & \cdots & S_n(r_p=1)=\mathrm{AEE}_N \\ \vdots & \vdots & \ddots & \vdots \\ S_0(r_q=0)=\mathrm{GAE}_0 & S_1(r_q=0)=\mathrm{GAE}_1 & \cdots & S_n(r_q=0)=\mathrm{GAE}_N \\ \vdots & \vdots & \ddots & \vdots \\ S_0(r_1=-1)=\mathrm{HAE}_0 & S_1(r_1=-1)=\mathrm{HAE}_1 & \cdots & S_n(r_1=-1)=\mathrm{HAE}_N \\ \mathrm{ME}_0 & \mathrm{ME}_1 & \cdots & \mathrm{ME}_N \\ \mathrm{EM}_0 & \mathrm{EM}_1 & \cdots & \mathrm{EM}_N \\ \mathrm{IMRE}_0 & \mathrm{IMRE}_1 & \cdots & \mathrm{IMRE}_N \\ \mathrm{AES}_0 & \mathrm{AES}_1 & \cdots & \mathrm{AES}_N \\ \mathrm{RES}_0 & \mathrm{RES}_1 & \cdots & \mathrm{RES}_N \end{bmatrix}_{(m+5)\times(N+1)} \tag{24}$$

where the first column represents the attributes of the desired weapon, the rest of the columns stand for the attributes of the corresponding evaluated weapon.

Clearly, the WSAA matrix at least includes two objectives: an evaluation object and a desirability object. Here, the evaluation objects are the weapon systems which we have been tested, and the desirability object (or the design objective) is a virtual weapon system that we desired it.

According to the above WSAA matrix, we use the WSAAC matrix to the weapon systems accuracy evaluation. However, the WSAAC matrix based on the WSAA matrix includes more aspects of weapon systems accuracy.

5. Weapon systems accuracy attributes competition matrix

According to the WSAA matrix, let any two weapons ($j \neq k$ $(j, k \in \{1, \cdots, N+1\})$) compare with each other. To be precise, let the j-th column values and the k-th column values be compared in pairs. Then denote $m_{\mathrm{WSAAC}}(j, k; A_{i(j,k)}^{\mathrm{WSAA}})$ as the competition result with respect to the i-th attribute $A_{i(j,k)}^{\mathrm{WSAA}}$ (the total number of the competitions w.r.t. $A_{i(j,k)}^{\mathrm{WAM}}$ is $m+5$), and

$$m_{\mathrm{WSAAC}}\big(j, k; A_{i(j,k)}^{\mathrm{WSAA}}\big) \triangleq \begin{cases} 1 & \text{if} A_{ij}^{\mathrm{WSAA}} \text{ is better than } A_{ik}^{\mathrm{WSAA}} \\ 0.5 & \text{if } A_{ij}^{\mathrm{WSAA}} \text{ is equal to } A_{ik}^{\mathrm{WSAA}} \\ 0 & \text{if } A_{ik}^{\mathrm{WSAA}} \text{ is better than } A_{ij}^{\mathrm{WSAA}} \end{cases} \tag{25}$$

where A_{ij}^{WSAA} is the i-th element of the j-th weapon systems within the WSAA matrix. In particular, A_{i0}^{WSAA} stands for the i-th attribute of the desired weapon.

Then, the WSAAC is given by

$$M_{\mathrm{WSAAC}}\big(j, k; A_{(1,\cdots,m+5)(j,k)}^{\mathrm{WSAA}}\big) = \frac{1}{m+5}\sum_{i=1}^{m+5} m_{\mathrm{WSAAC}}\big(j, k; A_{i(j,k)}^{\mathrm{WSAA}}\big) \tag{26}$$

where $A_{(1,\cdots,m+5)(j,k)}^{\mathrm{WSAA}}$ is the vector of the WSAA between the j-th weapon system and the k-th weapon system. For $M_{\mathrm{WSAAC}}(j, k; A_{(1,\cdots,m+5)(j,k)}^{\mathrm{WSAA}}) > 0.5$, we think that $A_{(1,\cdots,m+5)j}^{\mathrm{WSAA}}$ is WAC-better than $A_{(1,\cdots,m+5)k}^{\mathrm{WSAA}}$. In addition, we also have the following relation.

$$M_{\mathrm{WSAAC}}\big(j, k; A_{(1,\cdots,m+5)(j,k)}^{\mathrm{WSAA}}\big) + M_{\mathrm{WSAAC}}\big(k, j; A_{(1,\cdots,m+5)(j,k)}^{\mathrm{WSAA}}\big) = 1 \tag{27}$$

Furthermore, the non-transitive problem of the WSAAC is solved based on the WSAAC matrix.

$$X_{\mathrm{WSAAC}} \triangleq \begin{bmatrix} M_{\mathrm{WSAAC}}(1,1; A_{1(1,1)}^{\mathrm{WSAA}}) & \cdots & M_{\mathrm{WSAAC}}(1, N+1; A_{1(1,N+1)}^{\mathrm{WSAA}}) \\ \vdots & \ddots & \vdots \\ M_{\mathrm{WSAAC}}(N+1, 1; A_{(N+1)(N+1,1)}^{\mathrm{WSAA}}) & \cdots & M_{\mathrm{WSAAC}}(N+1, N+1; A_{(N+1)(N+1,N+1)}^{\mathrm{WSAA}}) \end{bmatrix} \tag{28}$$

Clearly, the WSAAC matrix contains the entire pairwise competition results of all compared weapon systems. Moreover, since the WSAAC matrix is a positive matrix, according to the Perron–Frobenius theorem [29], there exists an eigenvector $\mathrm{Eig}_{1\times(N+1)} > 0$, we have

$$X_{\mathrm{WSAAC}} \cdot \mathrm{Eig}_{1\times(N+1)} = \lambda \cdot \mathrm{Eig}_{1\times(N+1)} \tag{29}$$

where λ is the only eigenvalue on the spectral circle of X_{WSAAC}, λ and is called the Perron root. Thus, the above eigenvector (i.e., $\mathrm{Eig}_{1\times(N+1)}$) is applied to evaluate the weapon systems accuracy, because the elements of the eigenvector reflect how good the corresponding weapon systems accuracy is [28].

Remark 1. For the same type weapon systems, for any two tested weapon systems ($p \neq q(p, q \in \{1, \cdots, N+1\})$), we can conclude that the accuracy of the p-th tested weapon system outperforms the q-th tested weapon system if $\mathrm{Eig}_{1\times(N+1)}(p) > \mathrm{Eig}_{1\times(N+1)}(q)$ and vice versa. Certainly, if $\mathrm{Eig}_{1\times(N+1)}(p) = \mathrm{Eig}_{1\times(N+1)}(q)$, we can say that the accuracy of the p-th tested weapon system nearly as well as the q-th tested weapon system. Here $\mathrm{Eig}_{1\times(N+1)}(1)$ is the eigenvalue of the desired weapon system. On the contrary, $\mathrm{Eig}_{1\times(N+1)}(p) < \mathrm{Eig}_{1\times(N+1)}(1)$ implies that the p-th tested weapon systems accuracy cannot satisfy design requirement.

Remark 2. For the different type weapon systems, for $p \neq q$ $(p, q \in \{1, \cdots, N+1\})$, if $\mathrm{Eig}_{1\times(N+1)}(p) \geq \mathrm{Eig}_{1\times(N+1)}(q)$, we think that the accuracy of the p-th weapon system is superior to the q-th weapon system and vice versa.

Clearly, the above evaluation method can not only evaluate the similar weapon systems but also to evaluate different types.

6. The new evaluation method for the weapon systems accuracy

As stated before, a schematic diagram of the new evaluation method for weapon systems accuracy is summarized as follows.

Step 1: Initializing the desired weapon fall points data $(\boldsymbol{X}_0, \boldsymbol{Z}_0)$, the tested data of the k batch weapon $(\boldsymbol{X}_k, \boldsymbol{Z}_k)$ $(k = 1, 2, \cdots, N$, N is the number of the evaluated weapon), and the attributes distance D;

Step 2: According to (4), transform the above tested data to the relative distance error data $\tilde{\boldsymbol{c}}_k = (\tilde{c}_1, \tilde{c}_2, \cdots, \tilde{c}_B)$;

Step 3: Based on the new bootstrap method, we can obtain the new observations $\tilde{\boldsymbol{c}}_k^* = (\tilde{c}_1^*, \tilde{c}_2^*, \cdots, \tilde{c}_B^*)$;

Step 4: Compute the ES, the AES and the RES according to the new observations $\tilde{\boldsymbol{c}}_k^* = (\tilde{c}_1^*, \tilde{c}_2^*, \cdots, \tilde{c}_B^*)$, we can obtain the WSAA matrix;

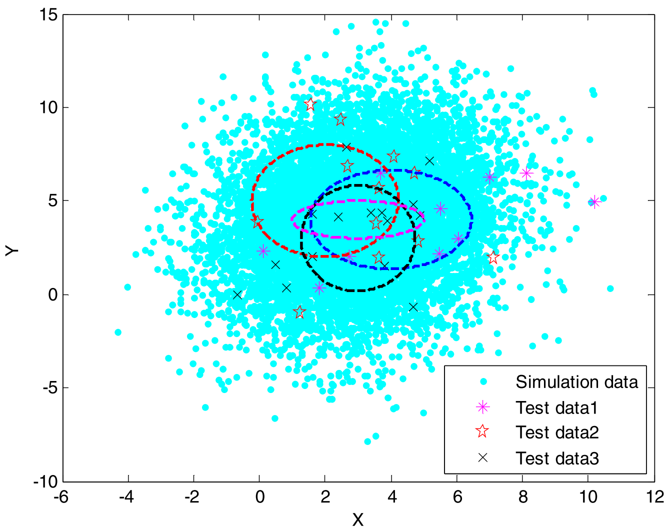


Fig. 3. The simulation data, test data1, test data2, and test data3.

Step 5: Based on the WSAA matrix, we have the WSAAC matrix;

Step 6: Calculate the eigenvector of the WSAAC matrix;

Step 7: According to the eigenvector, we can conclude that whether the tested weapon satisfied design requirement or not.

7. Numerical example

In this section, a numerical example using the synthetic data is provided to illustrate the weapon systems accuracy evaluation process of the proposed method.

7.1. The evaluation process using the proposed method

Step 1. As we know, based on the virtual proving ground, a great amount of simulation is carried out before the design of weapon systems. Therefore, we have a large number of simulation data $(\boldsymbol{X}^s, \boldsymbol{Z}^s) = (x_1^s, z_1^s), (x_2^s, z_2^s), \cdots, (x_{n_s}^s, z_{n_s}^s)$ where n_s is the number of performed simulations. Here, assume that $(\boldsymbol{X}^s, \boldsymbol{Z}^s)$ follows a bivariate normal distribution with mean $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{(\boldsymbol{X}^s,\boldsymbol{Z}^s)} \\ \sigma_1\sigma_2\rho_{(\boldsymbol{X}^s,\boldsymbol{Z}^s)} & \sigma_2^2 \end{bmatrix} \tag{30}$$

In addition, suppose that the target position is $(x_0, z_0) = (0, 0)$ and let $\boldsymbol{\mu} = (3, 4)$, $\Sigma = \left[\begin{smallmatrix} 4 & 1 \\ 1 & 9 \end{smallmatrix}\right]$, over 10 000 Monte Carlo runs, we obtained 10 000 fall points as shown in Fig. 3.

In practical applications, the first real experiment was carried out to evaluate whether the weapon system satisfied our design requirement. Then we obtained a tested data (test data1). Here, we assume that the first tested data are follows the bivariate normal distributions, as shown in Table 1. If the weapon system cannot satisfy our design requirement, we need to improve the weapon system. Assume that two more advanced technologies were used to improve the weapon system. In many real-life, to test the effectiveness of the above technologies, we made two real experiments. At the same time, we obtain two tested data. Here, we named the two tested data as test data2 and test data3, respectively. Similar to the test data1, the two tested data are still generated from the bivariate normal distributions, which are summarized in Table 1.

Similar to the simulation data, we have the above three test data over the Monte Carlo runs, as shown in Fig. 3.

Table 1
Test data of three times experiment.

Experiment classification	Number	Mean	Covariance matrix
Test data1	10	$\boldsymbol{\mu}_1 = (4, 4)$	$\Sigma_1 = \begin{bmatrix} 6 & 1 \\ 1 & 7 \end{bmatrix}$
Test data2	12	$\boldsymbol{\mu}_2 = (2, 5)$	$\Sigma_2 = \begin{bmatrix} 5 & 1 \\ 1 & 9 \end{bmatrix}$
Test data3	15	$\boldsymbol{\mu}_3 = (3, 3)$	$\Sigma_3 = \begin{bmatrix} 3 & 1 \\ 1 & 8 \end{bmatrix}$

Step 2. According to (4), the above simulation data and test data are transformed to the relative distance error data, respectively, which are given by

$$\begin{aligned} \tilde{\boldsymbol{c}}^s &= (\tilde{c}_1^s, \tilde{c}_2^s, \cdots, \tilde{c}_{n_s}^s) \\ \tilde{\boldsymbol{c}}^1 &= (\tilde{c}_1^1, \tilde{c}_2^1, \cdots, \tilde{c}_{10}^1) \\ \tilde{\boldsymbol{c}}^2 &= (\tilde{c}_1^2, \tilde{c}_2^2, \cdots, \tilde{c}_{12}^2) \\ \tilde{\boldsymbol{c}}^3 &= (\tilde{c}_1^3, \tilde{c}_2^3, \cdots, \tilde{c}_{15}^3) \end{aligned} \tag{31}$$

Step 3. The new bootstrap method is applied to enrich $\tilde{\boldsymbol{c}}^1, \tilde{\boldsymbol{c}}^2, \tilde{\boldsymbol{c}}^3$. Thus, we have

$$\begin{aligned} \tilde{\boldsymbol{c}}^{1*} &= (\tilde{c}_1^{1*}, \tilde{c}_2^{1*}, \cdots, \tilde{c}_B^{1*}) \\ \tilde{\boldsymbol{c}}^{2*} &= (\tilde{c}_1^{2*}, \tilde{c}_2^{2*}, \cdots, \tilde{c}_B^{2*}) \\ \tilde{\boldsymbol{c}}^{3*} &= (\tilde{c}_1^{3*}, \tilde{c}_2^{3*}, \cdots, \tilde{c}_B^{3*}) \end{aligned} \tag{32}$$

Furthermore, one replication and one thousand replications are performed to show the superiority of the new bootstrap method, as shown in Fig. 4 and Fig. 5, respectively.

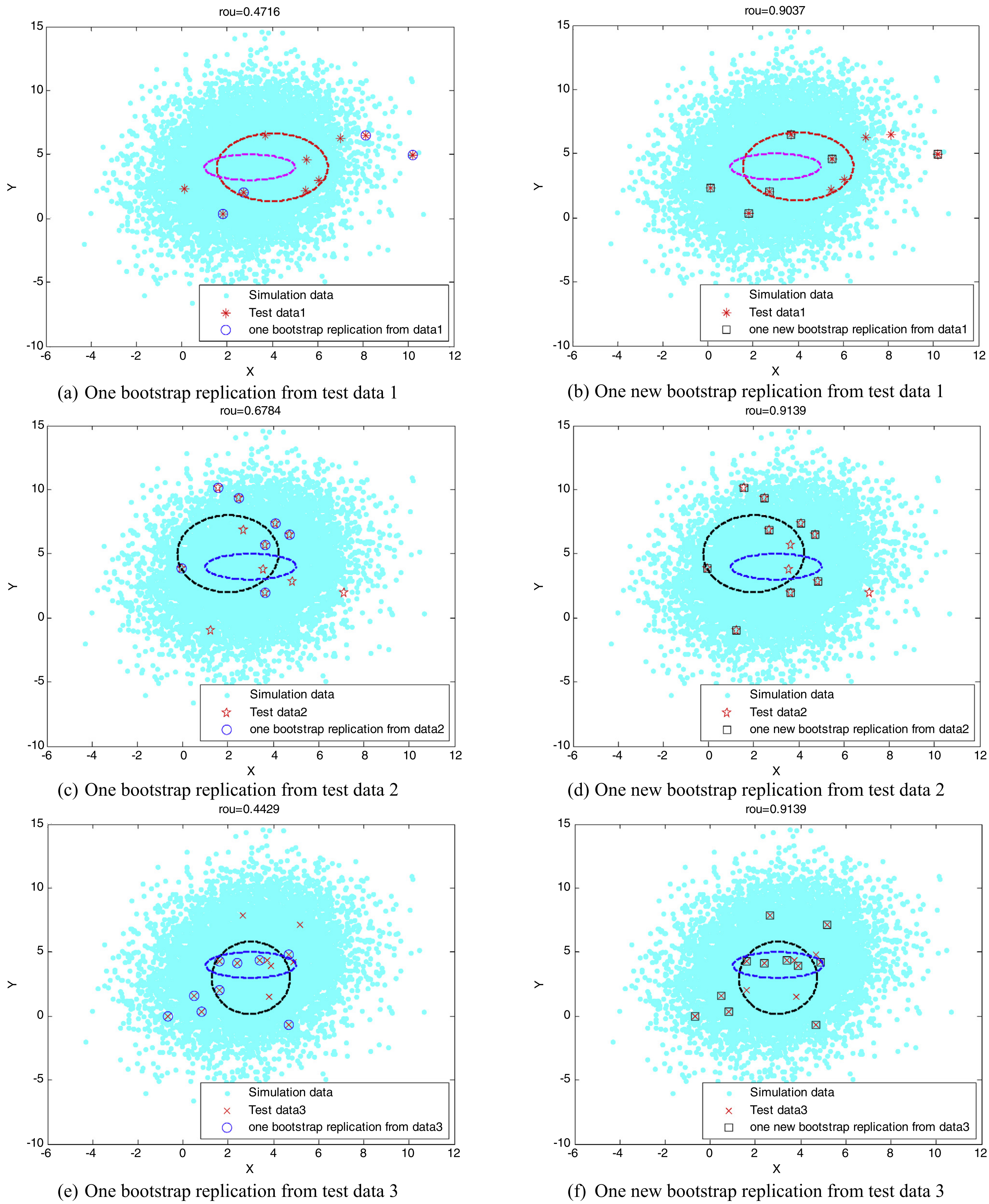
Figs. 4(a), 4(c) and 4(e) show that one bootstrap replication from the above three test data, respectively. Similarly, we can see three groups of a new bootstrap replication from Figs. 4(b), 4(d) and 4(f), respectively.

Before performing Step 4, we need to decide the attributes distance D. As discussed in Section 3.6, for $r \in [-1, 2]$, if $D = 0.75$, then $m = 4$. Thus, the WSAA matrix reduces to

$$A^{\mathrm{WSAA}} = \begin{bmatrix} S_0(r_m=2)=\mathrm{RMSE}_0 & S_1(r_m=2)=\mathrm{RMSE}_1 & \cdots & S_n(r_m=2)=\mathrm{RMSE}_N \\ S_0(r_p=1)=\mathrm{AEE}_0 & S_1(r_p=1)=\mathrm{AEE}_1 & \cdots & S_n(r_p=1)=\mathrm{AEE}_N \\ S_0(r_q=0)=\mathrm{GAE}_0 & S_1(r_q=0)=\mathrm{GAE}_1 & \cdots & S_n(r_q=0)=\mathrm{GAE}_N \\ S_0(r_1=-1)=\mathrm{HAE}_0 & S_1(r_1=-1)=\mathrm{HAE}_1 & \cdots & S_n(r_1=-1)=\mathrm{HAE}_N \\ \mathrm{ME}_0 & \mathrm{ME}_1 & \cdots & \mathrm{ME}_N \\ \mathrm{EM}_0 & \mathrm{EM}_1 & \cdots & \mathrm{EM}_N \\ \mathrm{IMRE}_0 & \mathrm{IMRE}_1 & \cdots & \mathrm{IMRE}_N \\ \mathrm{AES}_0 & \mathrm{AES}_1 & \cdots & \mathrm{AES}_N \\ \mathrm{RES}_0 & \mathrm{RES}_1 & \cdots & \mathrm{RES}_N \end{bmatrix}_{(4+5)\times(N+1)} \tag{33}$$

Clearly, the WSAA matrix only used the HAE, the GAE, the AEE and the RMSE, which is the four special points in the ES curve. In this case, much information of the ES curve will be lost. What's worse, we will obtain the different evaluation result because of the different attributes distance. Therefore, an appropriate attributes distance is designed to use more information of the ES curve and to obtain a steady evaluation result.

In this is numerical example, the attributes distance is given by the following simulation. According to (23)–(29), for any fixed $D \in [0, 1]$, substituting the above simulation data and the test data1 into the above measures, we obtained the WSAA matrix.



(a) One bootstrap replication from test data 1

(b) One new bootstrap replication from test data 1

(c) One bootstrap replication from test data 2

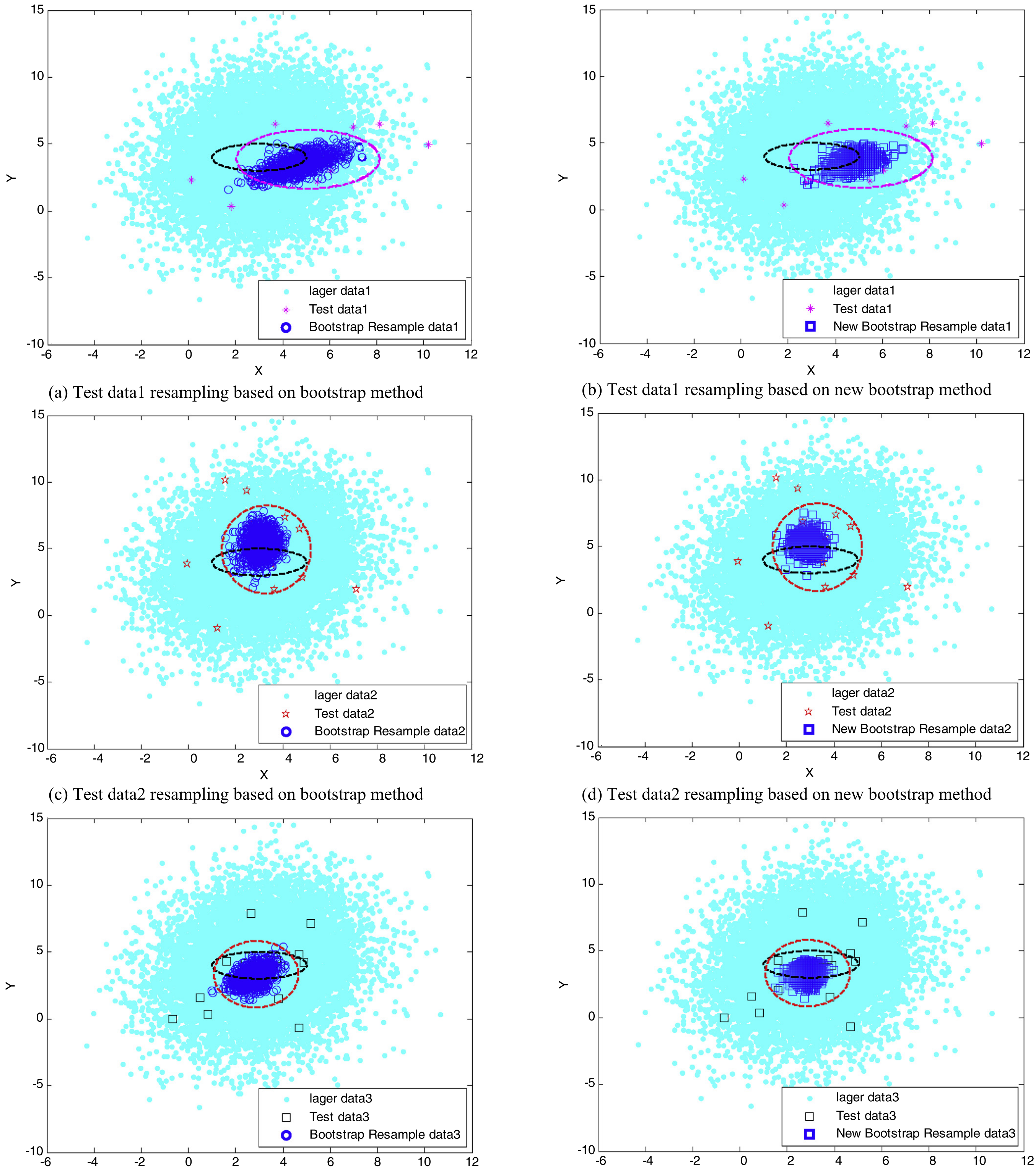
(d) One new bootstrap replication from test data 2

(e) One bootstrap replication from test data 3

(f) One new bootstrap replication from test data 3

Fig. 4. One replication from three test data based on bootstrap method and new bootstrap method.

Then the WSAAC matrix is calculated by the WSAA matrix. Furthermore, the eigenvector of the WSAAC matrix is computed to evaluate the weapon systems accuracy. Actually, the eigenvector of the WSAAC matrix depends on the elements of the WSAAC matrix. Therefore, if all the elements of the WSAAC matrix tends to stable, the eigenvector of the WSAAC matrix will be stable which leads to a steady evaluation result. Therefore, for convenience, take the first element $X_{\mathrm{WAC}}(1,2,A^{\mathrm{WAM}}_{(1,2,\cdots,m+5(1,2))})$ as an example. Then the $X_{\mathrm{WAC}}(1,2,A^{\mathrm{WAM}}_{(1,2,\cdots,m+5(1,2))})$ curve with the attributes distance is shown in Fig. 6. Furthermore, the computation time curve with attributes distance is shown in Fig. 7.



(a) Test data1 resampling based on bootstrap method

(b) Test data1 resampling based on new bootstrap method

(c) Test data2 resampling based on bootstrap method

(d) Test data2 resampling based on new bootstrap method

(e) Test data3 resampling based on bootstrap method

(f) Test data3 resampling based on new bootstrap method

Fig. 5. Three test data resampling by bootstrap method and new bootstrap method.

It can be clearly seen from Fig. 6 that if $D \in [0.2, 1]$, then $X_{\mathrm{WAC}}(1, 2, A^{\mathrm{WAM}}_{(1,2,\cdots,m+5(1,2)}) \in [0.2, 0.5]$, if, however, $D \in [0, 0.2]$, then $X_{\mathrm{WAC}}(1, 2, A^{\mathrm{WAM}}_{(1,2,\cdots,m+5(1,2)}) \in [0.35, 0.4]$. Clearly, when D is approaching zero, the range of the $X_{\mathrm{WAC}}(1, 2, A^{\mathrm{WAM}}_{(1,2,\cdots,m+5(1,2)})$ tends to a steady. Besides that, Fig. 7 shows that the computation time decreases as the increasing of the distance. Therefore, in this paper, we let the attributes distance equals to 0.005 ($D = 0.005$).

Step 4. Compute the ES, the AES and the RES according to the new observations $\tilde{\mathbf{c}}^*_k = (\tilde{c}^*_1, \tilde{c}^*_2, \cdots, \tilde{c}^*_B)$, we can obtain the WSAA matrix as follows:

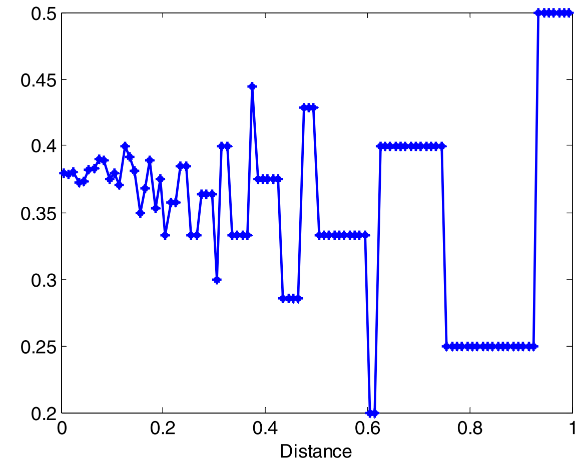


Fig. 6. The $X_{\mathrm{WAC}}(1,2,A^{\mathrm{WAM}}_{(1,2,\cdots,m+5(1,2)})$ curve.

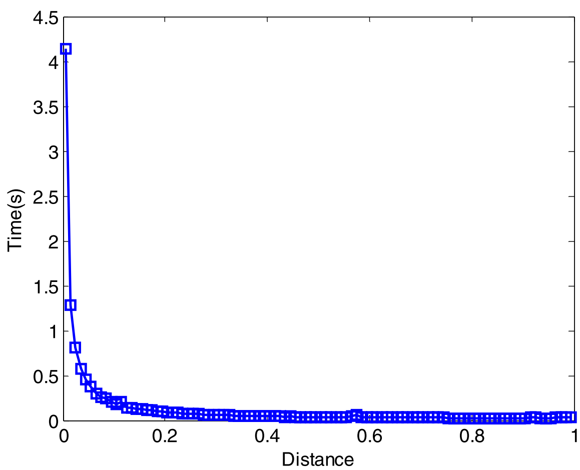


Fig. 7. The computation time curve.

$$A^{\mathrm{WSAA}} = \begin{bmatrix} \mathrm{RMSE}_0 = 6.1854 & \mathrm{RMSE}_1 = 6.0038 & \mathrm{RMSE}_2 = 6.0618 & \mathrm{RMSE}_3 = 4.3320 \\ \vdots & & & \\ \mathrm{AEE}_0 = 5.6106 & \mathrm{AEE}_1 = 5.9678 & \mathrm{AEE}_2 = 6.0410 & \mathrm{AEE}_3 = 4.3094 \\ \vdots & & & \\ \mathrm{GAE}_0 = 4.9187 & \mathrm{GAE}_1 = 5.9321 & \mathrm{GAE}_2 = 6.0208 & \mathrm{GAE}_3 = 4.2872 \\ \vdots & & & \\ \mathrm{HAE}_0 = 3.9867 & \mathrm{HAE}_1 = 5.8966 & \mathrm{HAE}_2 = 6.0014 & \mathrm{HAE}_3 = 4.2656 \\ \mathrm{ME}_0 = 5.4744 & \mathrm{ME}_1 = 5.9868 & \mathrm{ME}_2 = 6.0449 & \mathrm{ME}_3 = 4.3293 \\ \mathrm{EM}_0 = 5.3733 & \mathrm{EM}_1 = 6.0728 & \mathrm{EM}_2 = 6.1356 & \mathrm{EM}_3 = 4.4429 \\ \mathrm{IMRE}_0 = 5.6047 & \mathrm{IMRE}_1 = 5.9729 & \mathrm{IMRE}_2 = 6.0416 & \mathrm{IMRE}_3 = 4.3121 \\ \mathrm{AES}_0 = 15.7422 & \mathrm{AES}_1 = 17.8540 & \mathrm{AES}_2 = 18.0955 & \mathrm{AES}_3 = 12.8978 \\ \mathrm{RES}_0 = 3.2981 & \mathrm{RES}_1 = 0.1607 & \mathrm{RES}_2 = 0.0906 & \mathrm{RES}_3 = 0.0995 \end{bmatrix} \tag{34}$$

Step 5: Based on the WSAA matrix, we have the WSAAC matrix:

$$X^{NB}_{\mathrm{WSAAC}} = \begin{bmatrix} 0.5000 & 0.8725 & 0.9118 & 0.0850 \\ 0.1275 & 0.5000 & 0.9967 & 0.0000 \\ 0.0882 & 0.0033 & 0.5000 & 0.0033 \\ 0.9150 & 1.0000 & 0.9967 & 0.5000 \end{bmatrix} \tag{35}$$

Since the above matrix exists zero elements, to satisfy the Perron–Frobenius theorem, it can be replaced by very small positive numbers. Therefore, (35) can be rewritten as

$$X^{NB}_{\mathrm{WSAAC}} = \begin{bmatrix} 0.5000 & 0.8725 & 0.9118 & 0.0850 \\ 0.1275 & 0.5000 & 0.9967 & 0.0001 \\ 0.0882 & 0.0033 & 0.5000 & 0.0033 \\ 0.9150 & 0.9999 & 0.9967 & 0.5000 \end{bmatrix} \tag{36}$$

Step 6: Calculate the eigenvector of the WSAAC matrix.

According to the Perron–Frobenius theorem [29], the eigenvector of the (36) is given by

$$\mathrm{ERV}^{NB}_{\mathrm{WSAAC}} = [\,0.2715 \quad 0.1453 \quad 0.0520 \quad 0.5312\,] \tag{37}$$

Similar to (34)–(37), we can obtain the following result based on the bootstrap method.

$$\mathrm{ERV}^{B}_{\mathrm{WSAAC}} = [\,0.2500 \quad 0.1576 \quad 0.0366 \quad 0.5558\,] \tag{38}$$

Step 7: According to the eigenvector, we can conclude whether the tested weapon satisfied design requirement or not.

Thus, according to (37), the evaluation results are:

$$W_{\text{Third test}} > W_{\text{desired}} > W_{\text{Second test}} > W_{\text{First test}} \tag{39}$$

Similarly, according to (38), the evaluation results are:

$$W_{\text{Third test}} > W_{\text{desired}} > W_{\text{Second test}} > W_{\text{First test}} \tag{40}$$

7.2. Analysis of the simulation results

Fig. 4 shows that the bootstrap replication data lost more information than the new bootstrap replication data. That is, it can be seen from Figs. 4(a) and 4(b) that one of bootstrap replication data1 only used four fall points information. The information of the other six fall points was lost. However, based on the correlation coefficient criterion, one of new bootstrap replication data1 includes six fall points. Clearly, more information was considered in the new bootstrap replication data1. Similarly, as shown in Figs. 4(c)–4(f), both the new bootstrap replication data2 and the new bootstrap replication data3 are consider more information of the initial test data than the bootstrap replication data2 and the bootstrap replication data3, respectively.

Fig. 5 shows that the new bootstrap replication data are more concentration than the bootstrap replication data. Obviously, the reason is that each the new bootstrap replication data includes many different fall points by using the correlation coefficient criterion, while the bootstrap replication data is generated by the Monte Carlo method that leads to the randomness of the replication data.

According to (39) and (40), when the three tested data are the same weapon, we can conclude that after the first experiment, the weapon cannot satisfy the design requirement. Although the second experiment is still not meeting the design requirement, it shows that the second tested weapon outperforms the first tested weapon. In other words, the improve technology used in the second tested weapon is effective. Furthermore, the third tested weapon is superior to the desired weapon. Therefore, we think that the third tested weapon satisfy the design requirement. Certainly, when the three data are the different weapons, it can be easily concluded that the third weapon is the best; the next is the desired weapon and the second weapon, the first weapon is the poor.

8. Conclusion

The primary contribution of this paper is that a new evaluation method, based on error spectrum, is presented to the weapon systems accuracy. First, based on the correlation coefficient criterion, the new bootstrap method has faster convergence speed and lower square error than the bootstrap method under the small sample case. Second, more attributes of the weapon systems will be considered by using the error spectrum. Third, the above evaluation method can not only evaluate the similar weapon systems before and after improvement but also to rank the accuracy of the designed weapon systems to the other. Finally, simulation shows that the proposed evaluation method can efficiently handle the weapon systems accuracy problem. It is worth noting that the limitation of the proposed method is that the new evaluation method depends on the test data. As a future work, a possible direction is to put the presented method to the weapon systems evaluation.

Conflict of interest statement

The authors declared that they do not have any conflicts of interest to this work.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China through grant 61502534, the National Natural Science Foundation of China through grant 51506222, the Natural Science Foundation of Shaanxi Province of China through grant 2014JQ8339 and the Natural Science Foundation of Shaanxi Province of China through grant 2016JQ5085.

References

[1] D.L. Mon, C.H. Cheng, J.C. Lin, Evaluating weapon system using fuzzy analytical hierarchy process based on entropy weight, Fuzzy Sets Syst. 62 (2) (1994) 127–134.
[2] S.M. Chen, Evaluating weapon systems using fuzzy arithmetic operations, Fuzzy Sets Syst. 77 (3) (1996) 265–276.
[3] C.H. Cheng, Evaluating weapon systems using ranking fuzzy numbers, Fuzzy Sets Syst. 107 (1) (1999) 25–35.
[4] Q.M. Li, H.W. Wang, J. Liu, Small sample Bayesian analyses in assessment of weapon performance, J. Syst. Eng. Electron. 18 (3) (2007) 545–550.
[5] M. Dağdeviren, S. Yavuz, N. Kılınç, Weapon selection using the AHP and TOPSIS methods under fuzzy environment, Expert Syst. Appl. 36 (4) (2009) 8143–8151.
[6] J. Jiang, X. Li, Z.J. Zhou, et al., Weapon system capability assessment under uncertainty based on the evidential reasoning approach, Expert Syst. Appl. 38 (11) (2011) 13773–13784.
[7] J. Zhang, W.L. An, Assessing circular error probable when the errors are elliptical normal, J. Stat. Comput. Simul. 82 (4) (April 2012) 565–586.
[8] M. Jerome, P. Rudy, L.G. Francois, Missile target accuracy estimation with importance splitting, Aerosp. Sci. Technol. 25 (1) (2013) 40–44.
[9] A. Wald, Sequential tests of statistical hypotheses, Ann. Math. Stat. 16 (2) (Jun. 1945) 117–186.
[10] T.L. Saaty, Models, Methods, Concepts & Applications of the Analytic Hierarchy Process, second edition, Springer Science Business Media, New York, 2012, Chaps. 1, 2.
[11] P. Wang, P. Meng, B.W. Song, Response surface method using grey relational analysis for decision making in weapon system selection, J. Syst. Eng. Electron. 25 (2) (2014) 265–272.
[12] Y.D. He, H.Y. Chen, L.G. Zhou, et al., Intuitionistic fuzzy geometric interaction averaging operators and their application to multi-criteria decision making, Inf. Sci. 259 (3) (2014) 142–159.
[13] P.J. Huber, Robust Statistics, John Wiley & Sons, New York, 1981, Chaps. 3, 5.
[14] P.S. Bullen, Handbook of Means and Their Inequalities, Kluwer Academic, Dordrecht, the Netherlands, 2003, Chaps. 2, 3.
[15] X.R. Li, Z.L. Zhao, Measures of performance for evaluation of estimators and filters, in: Proceedings of the 2001 SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, in: Proc. SPIE, vol. 4473, July–Aug. 2001, pp. 530–541.
[16] X.R. Li, Z.L. Zhao, Evaluation of estimation algorithms, part I: incomprehensive measures of performance, IEEE Trans. Aerosp. Electron. Syst. 42 (4) (2006) 1340–1358.
[17] H.L. Yin, X.R. Li, J. Lan, Iterative mid-range with application to estimation performance evaluation, IEEE Signal Process. Lett. 22 (11) (Nov. 2015) 2044–2048.
[18] B. Efron, Boorstrap methods: another look at the Jackknife, Ann. Stat. 7 (1) (1979) 1–26.
[19] B. Efron, The Jackknife, the Bootstrap and Other Resampling Plans, SIAM, Philadelphia, 1982, Chaps. 5, 9.
[20] V. Picheny, N.H. Kim, R.T. Haftka, Application of bootstrap method in conservative estimation of reliability with limited samples, Struct. Multidiscip. Optim. 41 (2) (2010) 205–217.
[21] W.L. Hung, E.S. Lee, S.C. Chuang, Balanced bootstrap resampling method for neural model selection, Comput. Math. Appl. 62 (12) (2011) 4576–4581.
[22] D.B. Rubin, The Bayesian bootstrap, Ann. Stat. 9 (1) (Jan. 1981) 130–134.
[23] Z.G. Zheng, Random weighting method, Acta Math. Appl. Sin. 10 (2) (April 1987) 247–253.
[24] S.S. Gao, L. Xue, Y.M. Zhong, C.F. Gu, Random Weight method for estimation of error characteristics in SINS/GPS/SAR/ integrated navigation system, Aerosp. Sci. Technol. 46 (1) (2015) 22–29.
[25] Z.L. Zhao, X.R. Li, Interaction between estimators and estimation criteria, in: Proc. Int. Conf. Information Fusion, Philadelphia, PA, USA, July 2005, pp. 311–316.
[26] X.R. Li, Z.L. Zhao, Z.S. Duan, Error spectrum and desirability level for estimation performance evaluation, in: Proc. Workshop on Estimation, Tracking and Fusion: A Tribute to Fred Daum, Monterey, CA, USA, 24 May, 2007, pp. 1–7.
[27] Z.L. Zhao, X.R. Li, Two classes of relative measures of estimation performance, in: Proc. Int. Conf. Information Fusion, Québec City, Canada, July 2007, pp. 1432–1440.
[28] H.L. Yin, J. Lan, X.R. Li, Measures for ranking estimation performance based on single or multiple performance metrics, in: Proc. 16th International Conference on Information Fusion, Istanbul, Turkey, 9–12 July, 2013.
[29] R.B. Bapat, T.E.S. Rnghavan, Nonnegative Matrices and Applications, Cambridge University Press, 1997, Chapt. 2.
[30] Y.H. Mao, C.Z. Han, Z.S. Duan, Dynamic error spectrum for estimation performance evaluation: a case study on interacting multiple model algorithm, IET Signal Process. 8 (2) (2014) 202–210.
[31] W.S. Peng, Y.W. Fang, R.J. Zhan, et al., Two approximation algorithms of error spectrum for estimation performance evaluation, Optik 127 (5) (Dec. 2015) 2811–2821.
[32] W.S. Peng, Y.W. Fang, C. Dong, Enhanced dynamic error spectrum for estimation performance evaluation in target tracking, Optik 127 (8) (Mar. 2016) 3943–3949.
[33] W.S. Peng, Y.W. Fang, Z.S. Duan, et al., Enhanced error spectrum for estimation performance evaluation, Optik 127 (12) (Jun. 2016) 5084–5091.