Heliyon 9 (2023) e21208

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Conversion and fusion method of multi-source and different populations maintainability prior data

Cheng Zhou, Da Xu *, Zhaoyang Wang

Department of Arms and Control, Army Academy of Armored Forces, Beijing, 100072, China

ARTICLE INFO

Keywords:
Maintainability verification
Different populations data
Conversion factor
Data conversion
Multi-source data fusion

ABSTRACT

Maintainability is an important universal quality characteristic that reflects the convenience, speed and economy of weapon and equipment maintenance. Making full use of multi-source data to accurately verify the degree to which the developed equipment meets the maintainability requirements is an important basis for equipment identification and acceptance. To solve the low reliability of equipment maintainability verification results caused by inaccurate comprehensive prior distribution obtained by fusing multi-source and different populations' prior data, a method of data conversion and fusion is proposed. A data conversion model based on the mean value ratio of failure mode maintenance data is constructed. The conversion factor is defined according to objective data to convert the different populations' prior data to the same populations. Next, a comparison of the prior distribution fitting performance of Bayes bootstrap, bootstrap, and two improved sample-resampling methods to are used obtain the closest fitting distribution to the true distribution. By constructing a multi-source data fusion model based on improved KL divergence, a symmetrical KL divergence is constructed to describe the similarity between each prior distribution and the field distribution for the weighted fusion of multi-source prior distribution in addition to determining and testing the normal comprehensive prior distribution. The results show that the conversion and fusion method effectively converts the multi-source and different populations' maintainability prior data and obtains an accurate, comprehensive prior distribution by fusion, laying the foundation for applying the Bayes test method to verify the quantitative index of equipment maintainability.

1. Introduction

Maintainability is an important universal quality characteristic that reflects the convenience, speed and economy of equipment maintenance, related to the time, working hours and other material consumption and costs required for maintenance, which is defined as: the ability of the product to maintain and restore its prescribed state when it is repaired under the specified conditions and within the specified time, according to the specified procedures and methods. For military equipment, the maintainability in peacetime directly affects the combat readiness of the equipment, while in war time it affects the combat effectiveness of the equipment. With the complex and integrated development of modern weapons and equipment, the problem of maintainability is more prominent, and it has been in an equally important position with traditional performance, and has been highly valued by the military and industrial departments. The key to good maintainability lies in the equipment design, whose core is the design analysis and verification of

* Corresponding author.
E-mail address: ljzjbxy2022@163.com (D. Xu).

https://doi.org/10.1016/j.heliyon.2023.e21208
Received 10 July 2023; Received in revised form 27 August 2023; Accepted 18 October 2023
Available online 24 October 2023

maintainability. In order to more directly prove the degree of conformity between the maintainability level achieved by the equipment and the prescribed maintainability requirements, it is necessary to test and evaluate the equipment under the representative actual or near-actual conditions of use and operation. The purpose is to assess and verify the degree of the developed equipment to meet the maintainability requirements, to use it as the basis for equipment identification and acceptance, to find and identify the design defects related to equipment maintainability, in order to take corrective measures to achieve maintainability growth.

The Bayes test method [1–3] and the classical mathematical statistics method [4–6] are the primary means to verify the current maintainability of quantitative indicators of weapons and equipment. Based on comprehensive and sample information, the Bayes test method introduces the prior information, which is more suitable for verifying quantitative maintainability indicators for complex equipment systems under constraints such as the test cycle and cost in the type test stage and less field test data. The basis for applying the Bayes test is obtaining accurate prior distribution, especially under multi-source prior data. Effective fusion of multi-source data into accurate, comprehensive prior distribution is the key to ensuring the reliability of maintainability verification [7]. Current relevant research mostly focuses on the consistency test of multi-source prior and field test data, which can be approximately regarded as a fusion problem of the same population [8–10]. However, in engineering practice, multi-source prior data and field data are often different populations, and direct fusion of multi-source prior data will increase the deviation of the comprehensive prior distribution and reduce the credibility of maintainability verification results.

Multi-source information fusion technology plays an important role in various fields and practical applications, and has achieved fruitful research results [11–13]. For example, Xiao [11] introduced DS evidence theory to fuse multi-source information, and defined a generalized evidence Jensen Shannon (GEJS) divergence to measure the conflicts and differences among multi-source evidence, aiming at the problem that the results may be counterintuitive when DS evidence theory fuses highly conflicting evidence. This method can assign appropriate weights to evidence from different sources, modify the evidence body based on the corresponding weights, and combine DS evidence theory to effectively achieve the fusion of multi-source information. The case study combining fault diagnosis and sensitivity analysis further proves that the proposed method is effective and robust in resolving conflict situations. For the fusion of multi-source heterogeneous prior data, current research is primarily seen in the field of reliability. The idea of fusion is to introduce inheritance factors to describe the similarity between multi-source prior data and field test data, then combine the non-information prior distribution to fuse multi-source heterogeneous prior data into a mixed prior distribution [14–19]. For example, Ming et al. [14] used the chi-square goodness of fit method to determine the inheritance factor between historical data and field data and proposed the Bayes evaluation test scheme based on a mixed prior beta distribution to determine the reliability of the success or failure of products. In the reliability evaluation of small-sample high-value ammunition, Zhang et al. [16] quantified the uncertainty of reliability based on information entropy and conditional entropy theory, and determined the inheritance factor of multi-source and heterogeneous historical data using the uncertainty reduction ratio before constructing a mixed prior distribution of reliability. By introducing the inheritance factor, Kong et al. [17] realized the effective fusion of expert experience information, component subsystem information, and reliability growth information based on information fusion theory. Compared with the Bayes method, which ignores the heterogeneity of multi-source prior information, the reliability evaluation results of complex equipment systems are more reasonable. However, this type of method is aimed at the reliability evaluation of the success or failure of products, which are generally subject to a binomial distribution and thus not suitable for the fusion of the prior data of maintainability time parameters. In the field of maintainability, there has been little research on multi-source heterogeneous prior data fusion, and only a few documents provide a fusion method. The fusion idea establishes a conversion model of multi-source heterogeneous prior data to field data by defining the conversion factor and then carrying out the weighted fusion of the converted prior data distribution [20–22]. Also, Xu et al. [20] established the inner product model of the conversion of maintenance time historical data to field data and the linear model of the conversion of similar equipment data to field data of equipment to be evaluated then calculated the fusion weight of each prior distribution based on the difference of mean parameters. This realized the effective fusion of historical and similar equipment data of different populations. Based on literature 20, Xu et al. [21] and Miao et al. [22] also comprehensively considered three kinds of prior heterogeneous data–including expert experience data, virtual simulation data, and test data under different environments. Revising the data conversion model proposed in literature 20, the fusion weight was determined based on support, and the fusion accuracy of multi-source heterogeneous prior data was improved. This type of method solves the difficulty, to a certain extent, of fusing multi-source and different overall maintainability prior data. However, because the determination of the conversion factor in the data conversion model depends on expert experience, the workload is large, the subjectivity is strong, and the credibility of the fusion result cannot be guaranteed.

Based on the above analysis, this paper proposes a method of conversion and fusion of multi-source different populations maintainability prior data in view of the fact that the maintenance operation time of complex weapons and equipment follows lognormal distribution or normal distribution in most cases. Based on the field test data, a conversion model based on the mean ratio of fault mode maintenance data is constructed, and the conversion factor is determined by objective data, which solves the problem that the conversion factor determined by expert experience is subjective and the data conversion efficiency is low. The "S" type Boltzmann sigmoid function is used to fit the samples to improve the fitting accuracy of the prior distribution. A multi-source data fusion model based on improved KL divergence is established. The similarity between each prior distribution and the field distribution is described with the help of symmetrical KL divergence to realize a weighted fusion of multi-source data. The normal comprehensive prior distribution was determined and tested.

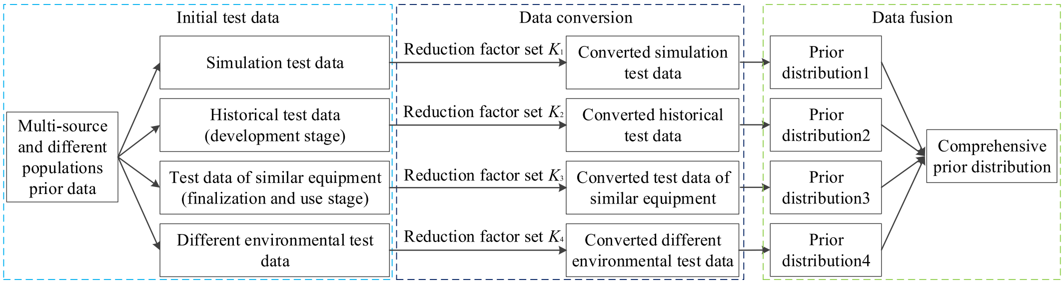

Fig. 1. Conversion and fusion scheme.

2. Data conversion and fusion scheme

2.1. Data classification and feasibility analysis

Weapons and equipment produce maintainability test data over their operational lives, which can provide a variety of prior data for verifying quantitative maintainability indicators based on the Bayes method. According to the source, generation stage, and generation level of maintainability data, multi-source data can be classified from three dimensions: source, time, and level. From the source aspect, the data can be divided into those of expert experience, simulation tests, historical tests, similar equipment tests, and different environmental tests. From the time aspect, the data can be divided into the performance test stage, the operational test stage, and the in-service assessment stage. From the level aspect, the data can be divided into system, subsystem, unit, and component data in the equipment hierarchy. The one based on the source dimension is the most widely used three-dimensional classification method in multi-source data fusion.

Based on preprocessing of multi-source data, we eliminated the impact of data outliers and missing values. We ensured the accuracy and credibility of multi-source prior data for equipment maintainability verification. Our study analyzed the reasons for the different populations of multi-source prior data and the feasibility of data conversion. The simulation test data comes from the simulation process for equipment virtual prototype fault maintenance, a virtual mapping of fault maintenance. The data are rich and reliable and can be converted as a priori data. However, the virtual prototype model, the maintenance process's simulation accuracy, and other simulation system factors may lead to a difference between the simulation test data and the field data. The equipment to be evaluated are the improved versions of equipment of a similar type and the same type of equipment in the historical test stage. The similarity of the two in the composition structure, functional principle, and maintenance support elements gives them the same or similar failure modes and maintenance processes, which provides the possibility for converting similar equipment test data and historical test data to field data. The improvement and perfection of equipment is also the main reason for the data discrepancy. The generation of different environmental test data relies on the same type of equipment in the same test stage. Just because the test environment is different, the failure mode structure and failure maintenance difficulty of equipment may change accordingly, resulting in data heterogeneity. However, these data still have certain similarities and can be used as a kind of prior data. Due to the strong subjectivity of expert experience data, expert data should be minimized in maintainability verification when there are many prior data. Thus, this paper does not consider the conversion and fusion of expert data.

2.2. Conversion and fusion scheme

Through the classification and conversion feasibility analysis of multi-source prior data, the conversion and fusion scheme of multi-source different overall maintainability prior data are proposed, as shown in Fig. 1.

Because the technical status of the same type of equipment in the development stage is the most similar to that of the equipment to be evaluated, to accurately verify the quantitative indicators of equipment maintainability, four kinds of multi-source and different overall prior data are selected as initial test data. The overall prior data include simulation test data, historical test data (development), similar equipment test data (finalization and use), and different environmental test data. We compared and analyzed the different global prior data and field data, determined the corresponding reduction factor set based on the mean ratio of the failure mode maintenance data, and conducted the data reduction. After the reduction, the study then generated the prior distribution of each prior data. Finally, the comprehensive prior distribution was solved by improving the multi-source data fusion model of KL divergence and verifying the prior distribution of equipment maintainability per the Bayes test method.

3. Data conversion model establishment

The purpose of converting multi-source and heterogeneous maintainability prior data is to convert heterogeneous prior data into prior data subject to the same population as the field data. This improves the fusion accuracy of the multi-source data. Determining a reasonable conversion factor is the key to data conversion. Because field data best reflect the true maintainability level of equipment, a data conversion model based on the mean ratio of failure mode maintenance data is built based on the field data. The model reflects the difference between the field data and the heterogeneous prior data using the conversion factor and realizes the effective conversion of the heterogeneous prior data to field data.

Assume that there are m types of failure in the field test of equipment, and record the field test data as $Y = (y_1, y_2, \cdots, y_m)$ according to the failure mode, where $y_i = (y_{i1}, y_{i2}, \cdots, y_{is})(i = 1, 2, \cdots, m)$ represents the collection of maintenance data for s times of the ith kind of fault. Similarly, suppose that certain maintainability prior data are $X^* = (x_1^*, x_2^*, \cdots, x_n^*)$ where n is the number of failure modes and $x_i^* = (x_{i1}^*, x_{i2}^*, \cdots, x_{il}^*)(i = 1, 2, \cdots, n)$ is the set of l times of maintenance data for the ith failure. Generally, the failure mode types and maintenance data in the prior information are more than the field test information; that is, $n \geq m, \sum_{i=1}^{n} l_i \geq \sum_{i=1}^{m} s_i$.

Then the reduction factor set $K = (k_1, k_2, \cdots, k_m)$, which represents the reduction factor of the ith failure mode in converting prior data to field data is defined. So,

$$k_i = \frac{\sum_{j=1}^{s} y_{ij} \Big/ s}{\sum_{j=1}^{l} x_{ij}^* \Big/ l}, \tag{1}$$

where $\sum_{j=1}^{s} y_{ij}/s$ is the mean value of the ith failure mode maintenance data of the field data, and $\sum_{j=1}^{l} x_{ij}^*/l$ is the mean value of the ith failure mode maintenance data of the prior data.

Then the data of the ith failure mode x_i^* in the prior data are converted to the field test data as x_i:

$$x_i = k_i x_i^* \quad (i = 1, 2, \cdots, n). \tag{2}$$

where $x_i = (k_i x_{i1}^*, k_i x_{i2}^*, \cdots, k_i x_{il}^*)$.

In the same way, the same failure mode data in the prior data X^* and the field test data can be converted into X one by one with the help of the reduction factor, $X = (k_1 x_1, k_2 x_2, \cdots, k_m x_m)$. For the failure mode data that only exists in the prior data, the method of literature 20 can be used for conversion to make full use of the multi-source prior data. Finally, the Wilcoxon signed-rank test was used to test the consistency between the converted data and the field data. The data conversion was tested according to the consistency results [23,24].

4. Prior distribution fitting of maintainability parameters

Before maintaining multi-source data fusion, it is necessary to convert all prior data into the form of a prior distribution of the maintenance parameters. In engineering practice, bootstrap and Bayes bootstrap methods are generally used to resample small-sample data to obtain new samples. The subsample estimator's overall distribution of the maintenance parameters is statistically inferred [25]. The basic principles of the two methods will not be discussed here. However, the data resampling method and the influence of the two improved methods on the fitting accuracy of the prior distribution will be.

Assume that the maintainability of prior data after data conversion is $X = (x_1, x_2, \cdots, x_n)$, $x_i \sim F(x)$, $i = 1, 2, \cdots, n$, arrange the data in X in ascending order, record them as $x_{(1)}, x_{(2)}, \cdots, x_{(n)}$, and construct the empirical distribution function $F_n(x)$:

$$F_n(x) = \begin{cases} 0 & x < x_{(1)} \\ \dfrac{i}{n} & x_{(i)} \leq x < x_{(i+1)} \quad i \in [1, n-1] \\ 1 & x \geq x_{(n)} \end{cases}. \tag{3}$$

The traditional and Bayes bootstrap methods generate new samples according to the empirical distribution function shown in Eq. (3). Restricted by the $F_n(x)$ structure of the empirical distribution function, it is easy to know that the new sample data are limited within the initial sample interval $[x_1, x_n]$, and the randomness is poor. In particular, the empirical distribution function cannot describe the distribution characteristics outside the initial sample points when the population is continuous, which will reduce the estimation accuracy of the prior distribution. Given the problem that $F_n(x)$ resampling cannot obtain the sample points outside the interval $[x_1, x_n]$, the two most widely used improved bootstrap methods are selected to compare the fitting performance of their prior distribution.

Improvement 1:

Use the following steps to resample the initial sample data to obtain new samples.

Step 1 Generate a group of random numbers $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$ subject to $U(0, 1)$, and make $\eta = (n-1)\xi$, $d = [\eta] + 1$;
Step 2 Extract a new sample $\hat{x}$ containing n data:

$$\hat{x} = x_{(d)} + (\eta - d + 1)\left(x_{(d+1)} - x_{(d)}\right), \tag{4}$$

where $\hat{x} = (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n)$; and

Step 3 Repeat steps 1 and 2 m times to obtain m groups of new samples.

Improvement 2:

It is known that the empirical distribution function $F_n(x)$ approximately follows the "S" type curve and that the Boltzmann sigmoid function is used to fit the sample on the global $(-\infty, +\infty)$. The fitting curve is further modified by Eq. (5) to make up for the deficiency of the upper and lower limit fitting effect of $F_n(x)$ and the modified empirical distribution function $\hat{F}(x)$.

Step 1 Obtain the modified fitting curve function $\hat{f}(x)$:

$$\int_{-\infty}^{+\infty} x\hat{f}(x)\mathrm{d}x = \text{initial sample mean.} \tag{5}$$

Step 2 Generate a group of random numbers $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$ that obey $U(0,1)$, then

$$\hat{x} = \hat{F}^{-1}(\beta), \tag{6}$$

where $\hat{x}$ is a group of new samples containing n data, $\hat{x} = (\hat{x_1}, \hat{x_2}, \cdots, \hat{x_n})$.

Step 3 Repeat steps 1 and 2 m times to obtain m groups of new samples.

5. Multi-source prior distribution fusion

After obtaining the prior distribution of maintainability multi-source data separately, in order to ensure the accurate verification of equipment maintainability using Bayes test method, it is necessary to effectively fuse the multi-source prior distribution into an accurate comprehensive prior distribution. This section establishes a multi-source prior distribution fusion model, determines the fusion weight according to the similarity between each prior distribution and the field distribution, and obtains the comprehensive prior distribution by weighting, and further determines and tests whether the comprehensive prior distribution is a normal distribution.

5.1. Derivation of fusion model

Kullback–Leibler (KL) divergence is defined based on entropy and can quantitatively describe the degree of difference between probability distributions [26,27]. Based on the obtained maintainability of multi-source prior distribution and field distribution, the KL divergence can be used to measure the difference between each distribution and the field distribution, reflect the similarity degree between them, and determine the fusion weight according to the similarity degree to fuse the multi-source distribution.

With H maintainability multi-source prior data, the known distributions of maintainability parameter θ (θ is the mean value) are $\pi_1(\theta), \pi_2(\theta), \cdots, \pi_H(\theta)$, and the field distribution is $\pi_z(\theta)$, then the KL divergence of the ith distribution $\pi_i(\theta), i = 1, 2, \cdots, H$ to the field distribution $\pi_z(\theta)$ is:

$$\mathrm{KL}(\pi_i \| \pi_z) = \int_{-\infty}^{+\infty} \pi_i(\theta) \log \frac{\pi_i(\theta)}{\pi_z(\theta)} \mathrm{dx}. \tag{7}$$

Due to the asymmetry of KL divergence, that is, $\mathrm{KL}(\pi_i \| \pi_z) \neq \mathrm{KL}(\pi_z \| \pi_i)$, different KL divergence values will be obtained when the prior distribution and the field distribution are taken as benchmarks. At this time, the similarity between the prior distribution and the field distribution cannot be compared according to the KL divergence. To overcome the asymmetry of KL divergence, the KL divergence is symmetrically improved so that:

$$D(\pi_i, \pi_z) = \mathrm{KL}(\pi_i \| \pi_z) + \mathrm{KL}(\pi_z \| \pi_i) = \int_{-\infty}^{+\infty} (\pi_i(\theta) - \pi_z(\theta)) \log \frac{\pi_i(\theta)}{\pi_z(\theta)} \mathrm{dx}, \tag{8}$$

where $D(\pi_i, \pi_z)$ is the symmetry difference between $\pi_i(\theta)$ and $\pi_z(\theta)$. The larger $D(\pi_i, \pi_z)$ is, the greater the difference between the two probability distributions, the smaller the similarity; the converse is true as well. In special cases, when $\pi_i(\theta) = \pi_z(\theta)$ and $D(\pi_i, \pi_z) = 0$, the two probability distributions are identical, and prior data can be used directly as field test data.

Use Eq. (9) to convert the difference $D(\pi_i, \pi_z)$ between $\pi_i(\theta)$ and $\pi_z(\theta)$ into the similarity of interval $[0, 1]$, and record it as $\rho(\pi_i, \pi_z)$:

$$\rho(\pi_i, \pi_z) = 1 - \frac{D(\pi_i, \pi_z)}{1 + D(\pi_i, \pi_z)} = \frac{1}{1 + D(\pi_i, \pi_z)}. \tag{9}$$

Therefore, when performing multi-source prior distribution fusion, the more similar the prior distribution is to the field distribution, the larger the fusion weight should be, and the fusion weight w_i of the prior distribution $\pi_i(\theta)$ is

Table 1
Conversion of maintainability history tests data from different populations.

Failure mode number	MTTR/min (historical data)	MTTR/min (field data)	Reduction factor	MTTR/min (historical data after conversion)
1	27.0000	25.0000	0.9259	25.0000
2	29.0000, 30.0000	27.0000	0.9153	26.5424, 27.4576
3	30.0000, 33.0000	29.0000	0.9206	27.6190, 30.3810
4	35.0000, 38.0000, 39.0000	34.0000	0.9107	31.8750, 34.6071, 35.5179
5	39.0000, 40.0000	36.0000	0.9114	35.5443, 36.4557
6	38.0000, 39.0000, 42.0000	38.0000	0.9580	36.4034, 37.3613, 40.2353
7	43.0000, 45.0000	40.0000	0.9091	39.0909, 40.9091
8	46.0000	42.0000	0.9130	42.0000
9	48.0000, 49.0000	45.0000	0.9278	44.5361, 45.4639
10	70.0000, 72.0000	61.0000	0.8592	61.8592, 60.1406

$$w_i = \frac{\rho(\pi_i, \pi_z)}{\sum_{i=1}^{H} \rho(\pi_i, \pi_z)}, \tag{10}$$

where H is the number of the multi-source prior distribution, $\sum_{i=1}^{H} w_i = 1$.

The integrated prior distribution $\pi(\theta)$ after fusion is

$$\pi(\theta) = \sum_{i=1}^{H} w_i \pi_i(\theta). \tag{11}$$

5.2. Determination and test of synthetic prior distribution

In this paper, given the condition that the maintenance time follows the log-normal distribution and normal distribution, to meet the demand that the prior distribution is uniform when the Bayes sequential test method [28,29] is used for maintainability verification when the mean values of the multi-source prior distribution are not much different, the fused comprehensive prior distribution can be approximately regarded as a normal distribution. However, the normal comprehensive prior distribution needs to be determined and tested.

Take the fusion of two prior distributions as an example. There are two sets of independent maintenance prior data, X_1 and X_2, where $X_1 \sim \pi_1(\theta) = N(\theta_1, \sigma_1^2)$, $X_2 \sim \pi_2(\theta) = N(\theta_2, \sigma_2^2)$. According to Eq. (10), the comprehensive prior distribution is $\pi(\theta) = w_1\pi_1(\theta) + w_2\pi_2(\theta)$. When the mean and of the prior distribution and are not very different, assuming $\pi(\theta) = f(\theta) \sim N(\mu, \sigma^2)$, we need to test whether can be approximated to normal distribution and determine its parameters.

Step 1 Based on the 3σ principle of normal distribution, select the abscissa interval $[t_1, t_2]$, and use the step $\Delta t = (t_2 - t_1)/1000$ as the probability density curve of the prior distribution $\pi_1(\theta)$, $\pi_2(\theta)$ and $\pi(\theta)$, where $t_1 \le \min(\theta_1 - 3\sigma_1, \theta_2 - 3\sigma_2)$, $t_2 \ge \max(\theta_1 + 3\sigma_1, \theta_2 + 3\sigma_2)$.

Step 2 Provisionally observe and judge whether $\pi(\theta)$ approximates a normal distribution. If the curve of $\pi(\theta)$ is in the form of a regular normal distribution probability density curve, it is provisionally determined that $\pi(\theta)$ is a normal distribution; otherwise, it is directly determined that $\pi(\theta)$ cannot be approximated by a normal distribution. In the provisional case, if $f(\mu) = 1/(\sqrt{2\pi}\sigma) = \max(\pi(\theta))$, the standard deviation σ of $f(\theta)$ can be obtained. In addition, when $\pi(\theta)$ takes the maximum value $\max(\pi(\theta))$, the corresponding pointer I can be obtained, and the mean value of $f(\theta)$ is $\mu = t_1 + \Delta t \times (I - 1)$.

Step 3 Select the abscissa interval $[t_1', t_2']$ again, where $t_1' \le \min(\mu - 3\sigma, t_1)$, $t_2' \ge \max(\mu + 3\sigma, t_2)$, and use the step $\Delta t' = (t_2' - t_1')/1000$ for the curve of $\Delta\pi(\theta) = \pi(\theta) - f(\theta)$, where $\Delta\pi(\theta)$ is the difference between the comprehensive prior distribution $\pi(\theta)$ and the normal distribution $f(\theta)$.

Step 4 Check whether $\pi(\theta)$ can be approximated by a normal distribution. To ensure the accuracy of verifying equipment maintainability quantitative indicators and reduce the verification error caused by approximating $\pi(\theta)$ to $f(\theta)$, set the inspection threshold $\varepsilon = 0.1 \max(\pi(\theta))$. If $\max(|\Delta\pi(\theta)|) \le \varepsilon$, accept $\max(|\Delta\pi(\theta)|) \le \varepsilon$, and vice versa.

6. Results and discussion

For example, suppose that a certain type of armored vehicle has multi-source maintainability prior data that differs from the small sample of field test data in the type test stage. To realize the effective fusion of multi-source heterogeneous maintainability prior data and obtain accurate and comprehensive prior distribution, this paper takes the conversion and fusion of the heterogeneous data of the mean time to repair (MTTR) index as an example and verifies the feasibility and effectiveness of the proposed method by combining the historical test data in the equipment development stage and the similar equipment test data in the design and use stage (these two types of data have already been processed).

Table 2
Estimation results of MTTR logarithmic mean.

Method	Estimated value of μ	Confidence interval of μ	Interval length of μ	Estimated value of σ	Confidence interval of σ	Interval length of σ
Classical statistical method	3.6073	[3.5133,3.7012]	0.1879	0.2429	[0.1928,0.3328]	0.1400
Bootstrap method	3.6085	[3.6075,3.6096]	0.0021	0.0531	[0.0524,0.0538]	0.0014
Bayes bootstrap method	3.6078	[3.6068,3.6088]	0.0020	0.0511	[0.0504,0.0518]	0.0014
Improvement 1	3.6097	[3.6088,3.6107]	0.0019	0.0485	[0.0479,0.0492]	0.0013
Improvement 2	3.6070	[3.6061,3.6078]	0.0017	0.0434	[0.0428,0.0440]	0.0012

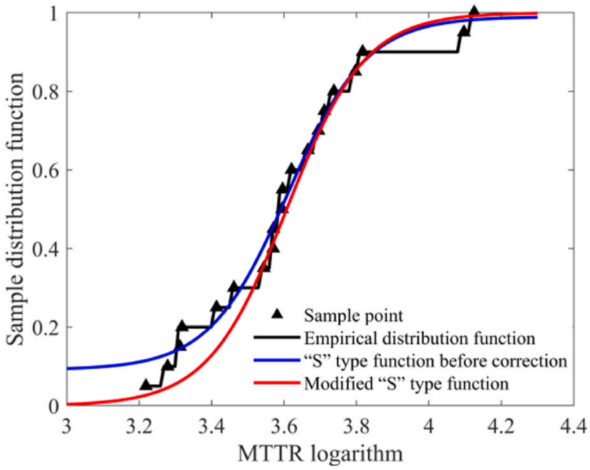

Fig. 2. Sample distribution function.

6.1. Conversion of different populations' prior data

Taking the conversion of historical test data of different populations as an example, the historical data and field data are divided into ten groups according to the different failure modes, with a one-to-one group correspondence. Then, based on the field data, the conversion factor of each failure mode data are calculated based on the mean ratio of failure mode maintenance data, and the conversion of historical data to field data is realized. The results are shown in Table 1.

Under the condition of significance level $\alpha = 0.05$, the Wilcoxon method was used to test the consistency of the converted historical data and similar equipment data with the field data. The results passed the consistency test, and the probability of the historical data being equal to the overall mean of the field data before and after the conversion was 0.3108 and 0.9474, respectively. The probability of the similar equipment data being equal to the overall mean of the field data before and after the conversion was 0.2903 and 0.6918, respectively. From these results, we can observe that the conversion model effectively converts the different population prior data to the field data and significantly improves the probability that the prior data are equal to the overall mean of the field data.

6.2. Determination of prior distribution

It is known that MTTR follows the log-normal distribution, and the field data, converted historical data, and similar equipment data are recorded as X_z, X_1 and X_2 after taking the logarithm.

$$X_z = (3.2189, 3.2958, 3.3673, 3.5264, 3.5835, 3.6376, 3.6889, 3.7377, 3.8067, 4.1109)$$

$$X_1 = (3.2189, 3.2787, 3.3126, 3.3185, 3.4138, 3.4618, 3.5441, 3.5700, 3.5708, 3.5961, \\ 3.5947, 3.6206, 3.6947, 3.6659, 3.7114, 3.7377, 3.7963, 3.8169, 4.0967, 4.1249$$

$$X_2 = (3.1429, 3.2895, 3.2752, 3.3160, 3.2932, 3.3673, 3.4024, 3.4024, 3.4340, 3.5835, \\ 3.5835, 3.6278, 3.6278, 3.6568, 3.6889, 3.7377, 3.7837, 3.7837, 3.8291, 4.1109$$

Taking X_1 as an example, two improved bootstrap methods are used to resample and obtain new samples, with 10,000 samples and 20 data in each new sample. The statistical inference results are shown in Table 2 by the classical statistical method, bootstrap method, and Bayes bootstrap method.

It can be seen from Table 2 that the improved method II of fitting empirical distribution function $F_n(x)$ with "S" type Boltzmann sigmoid function has the highest precision of parameter estimation results, indicating that the fitting results of prior distribution are

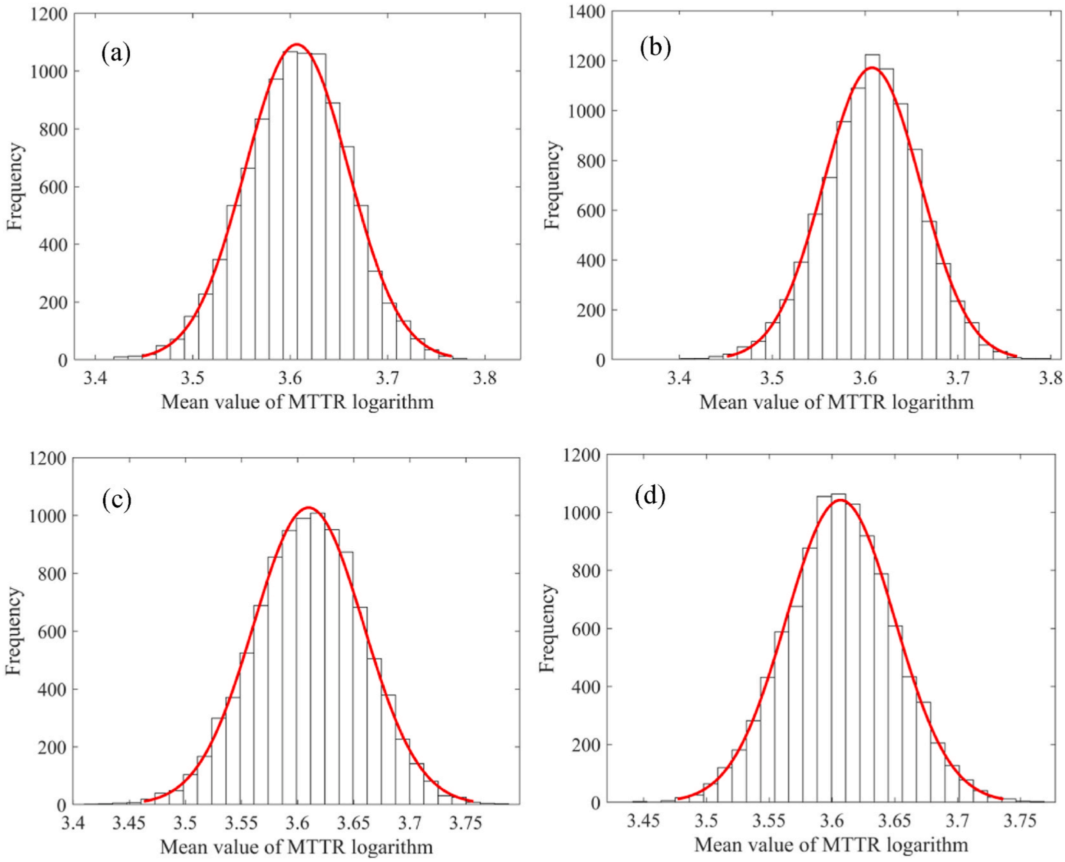

Fig. 3. (a) Distribution fitting curve of bootstrap method. (b) Distribution fitting curve of Bayes bootstrap method. (c) Distribution fitting curve of improved bootstrap method I. (d) Distribution fitting curve of improved bootstrap method II.

closest to the true distribution, and the sample distribution function is shown in Fig. 2.

The Boltzmann sigmoid distribution function $F(x)$ and its density function $f(x)$ of the sample are respectively

$$F(x)=0.0858+\frac{0.9058}{1+e^{(-x+3.6087)/0.1067}} \text{ and } f(x)=8.4892\times\frac{e^{(-x+3.6087)/0.1067}}{\left(1+e^{(-x+3.6087)/0.1067}\right)^2}.$$

The modified empirical distribution function $\hat{F}(x)$ and its density function $\hat{f}(x)$ are respectively

$$\hat{F}(x)=\frac{1}{1+e^{(-x+3.6073)/0.1067}} \text{ and } \hat{f}(x)=9.3721\times\frac{e^{(-x+3.6073)/0.1067}}{\left(1+e^{(-x+3.6073)/0.1067}\right)^2}.$$

The distribution fitting curve of the logarithmic mean of MTTR is shown in Fig. 3.

Then the prior distribution of the converted historical data is $\pi_1(\theta) = N(3.6070, 0.0434^2)$.

Similarly, the prior distribution of similar equipment data after conversion is: $\pi_2(\theta) = N(3.5511, 0.0466^2)$, and the distribution of field data is $\pi_z(\theta) = N(3.5875, 0.0432^2)$.

6.3. Fusion of multi-source prior distribution

Combine Eq. (7) and Eq. (8) to calculate the symmetry difference of historical distribution and similar equipment distribution with field distribution respectively:

$$D(\pi_1,\pi_z)=0.2177, D(\pi_2,\pi_z)=0.6211.$$

According to Eq. (9), calculate the similarity of historical distribution and similar equipment distribution with field distribution respectively:

$$\rho(\pi_1,\pi_z)=0.8212, \rho(\pi_2,\pi_z)=0.6169.$$

According to Eq. (10), calculate the fusion weights of historical distribution and similar equipment distribution respectively:

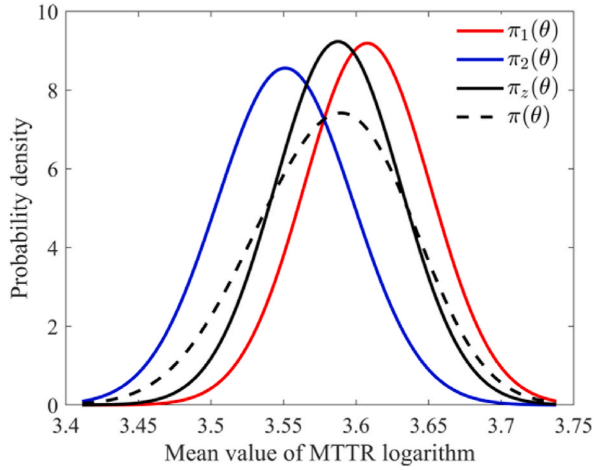

Fig. 4. Prior distribution and field distribution curves.

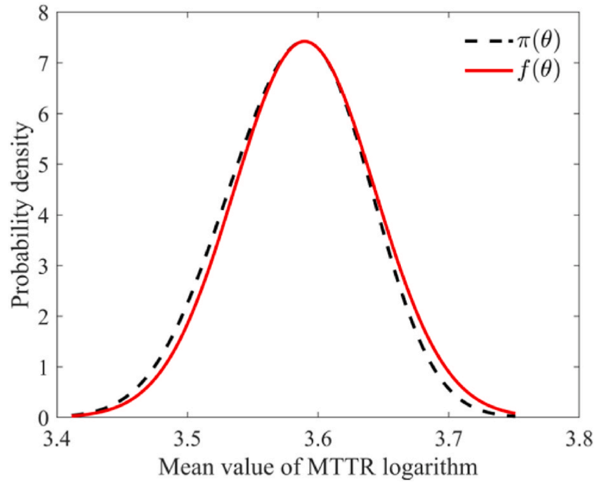

Fig. 5. Comprehensive prior distribution and approximate normal distribution curves.

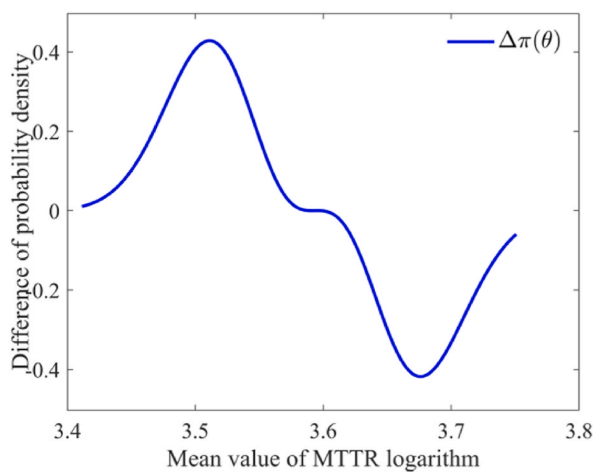

Fig. 6. Difference between comprehensive prior distribution and approximate normal distribution.

$w_1 = 0.5710, w_2 = 0.4290.$

According to Eq. (11), calculate the weighted fusion distribution of historical distribution and similar equipment distribution:

$$\pi(\theta) = 0.5710N\left(3.6070, 0.0434^2\right) + 0.4290N\left(3.5511, 0.0466^2\right).$$

Let us test whether $\pi(\theta)$ is approximately a normal distribution. Based on the 3σ principle of normal distribution, take the abscissa interval $[t_1, t_2] = [3.4113, 3.7379]$ and step $\Delta t = 3.266 \times 10^{-4}$ as the curves of $\pi_1(\theta)$, $\pi_2(\theta)$, $\pi_z(\theta)$ and $\pi(\theta)$, as shown in Fig. 4.

From Fig. 4, it can be provisionally assumed that the comprehensive prior distribution $\pi(\theta)$ is a normal distribution. Let $\pi(\theta) = f(\theta) \sim N(\mu, \sigma^2)$, and calculate the standard deviation of $f(\theta)$ from $f(\mu) = 1/(\sqrt{2\pi}\sigma) = \max(\pi(\theta))$, $\sigma = 0.0537$. When $\pi(\theta)$ takes the maximum value $\max(\pi(\theta))$, find the corresponding pointer $I = 547$, then the mean value of $f(\theta)$ is $\mu = t_1 + \Delta t \times (I - 1) = 3.5896$. Reselect the interval $[t_1', t_2'] = [3.4113, 3.7509]$ and step $\Delta t' = 3.396 \times 10^{-4}$, and make the probability density curve of $\pi(\theta)$ and $f(\theta)$, as shown in Fig. 5. Plot the curve of $\Delta\pi(\theta) = \pi(\theta) - f(\theta)$, as shown in Fig. 6.

As seen from Fig. 5, the probability density curves of $\pi(\theta)$ and $f(\theta)$ are closely similar, and the test threshold $\varepsilon = 0.1 \max(\pi(\theta)) = 0.7422$ is taken. From $\max(|\Delta\pi(\theta)|) = 0.4284 \leq 0.7422$, accept $\pi(\theta) = f(\theta) \sim N(3.5896, 0.0537^2)$.

7. Conclusion

(1) A conversion model based on the mean ratio of the failure mode maintenance data is established. Based on the field test data, the conversion factor is defined according to the mean ratio of the failure mode maintenance data, which solves the problem that the determination of the conversion factor by expert experience is subjective and that the data conversion efficiency is low.
(2) The prior distribution fitting performance of the Bayes bootstrap method, bootstrap method, and two improved sample-resampling methods are compared. The results show that using the "S" type Boltzmann sigmoid function can better fit the samples to obtain a prior distribution.
(3) The multi-source data fusion model is established to realize the weighted fusion of multi-source prior distribution by introducing KL divergence, and the normal comprehensive prior distribution is further determined and tested.

Author contributions

Cheng Zhou conceived and designed the experiments, Cheng Zhou and Da Xu performed the experiments, Cheng Zhou and Zhaoyang Wang analyzed and interpreted the data, all authors wrote the paper.

Data availability statement

Data included in article/supp. material/referenced in article.

Funding

This study was funded by the Defense Industrial Technology Development Program(2020ZB20)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] F.A.M. Ahmed, G. Mootaz, H. Amr, Bayesian estimation of parameters of reliability and maintainability of a component under imperfect repair and maintenance, Arabian J. Sci. Eng. 44 (3) (2019) 2497–2502.
[2] L.Z. Wang, R. Pan, X.H. Wang, W.H. Fan, J.Q. Xuan, A Bayesian reliability evaluation method with different types of data from multiple sources, Reliab. Eng. Syst. Saf. 167 (2017) 128–135.
[3] W.J. Zhang, H.B. Yang, S.F. Zhang, Reliability evaluation of success or failure products based on Bayes mixed prior distribution, Acta Armamentarii 37 (3) (2016) 505–511.
[4] R. Lev, S. Oksana, Development of modern models and methods of the theory of statistical hypothesis testing, E. Eur. J. Enterprise Technol. 5 (4) (2020) 11–18.
[5] Z.Y. Wu, J.P. Hao, H.Q. Chen, J.Q. Li, J.S. Tan, Review on the verification methods of equipment mean time to repair, Syst. Eng. Electr. Technol. 42 (12) (2020) 2931–2943.
[6] S. Geaici, P.K. Varshney, On the optimality of likelihood ratio test for prospect theory-based binary hypothesis testing, IEEE Signal Process. Lett. 25 (12) (2018) 1845–1849.
[7] J. Du, B.J. Fei, Y. Liu, G.Z. Yao, Research on reliability assessment for the armored equipment based on multi-source information fusion, Adv. Mater. Res. 2542 (2013) 2091–2094.
[8] B. Wang, P. Jiang, B. Guo, Reliability evaluation of aerospace valves based on multi-source information fusion, Acta Armamentarii 43 (1) (2022) 199–206.
[9] B.Y. Wei, Y.H. Shou, W. Jiang, A visual weighted average information fusion method, J. Xi'an Jiaot. Univ. 52 (4) (2018) 145–149.
[10] S.S. Shao, H.T. Liu, Z.H. Zhang, Reliability evaluation of ship spare parts based on multi-source prior information fusion, Syst. Eng. Electr. Technol. 41 (12) (2019) 2905–2910.
[11] F. Xiao, Y. GEJS, A generalized evidential divergence measure for multisource information fusion, IEEE transactions on systems man cybernetics systems 53 (4) (2023) 2246–2258.

[12] F.Y. Xiao, W. Pedrycz, Negation of the quantum mass function for multisource quantum information fusion with its application to pattern classification, IEEE Trans. Pattern Anal. Mach. Intell. 45 (2) (2023) 2054–2070.
[13] F.Y. Xiao, Z.H. Cao, C.T. Lin, A complex weighted discounting multisource information fusion with its application in pattern classification, knowledge and engineering 35 (8) (2023) 7609–7623.
[14] Z.M. Ming, J.Y. Tao, X. Chen, Y.A. Zhang, Research on bayes reliability evaluation test scheme of success or failure products based on mixed Beta distribution, Acta Armamentarii (2) (2008) 204–207.
[15] J. Guo, Z.J. Li, T. Keyser, A Bayesian approach for integrating multilevel priors and data for aerospace system reliability assessment, Chin. J. Aeronaut. 31 (1) (2018) 41–53.
[16] L.L. Zhang, L. Jie, X. Gao, Study on the comprehensive evaluation method of high value ammunition reliability, Syst. Eng. Electr. Technol. 43 (11) (2021) 3399–3404.
[17] Y. Kong, D. Li, Y. Sun, Comprehensive evaluation of complex system reliability based on information fusion, Ordnance Industry Automation 38 (12) (2019) 68–73.
[18] H.P. Wang, F.H. Duan, Bayesian network method for reliability analysis of complex uncertainty systems, Acta Armamentarii 41 (1) (2020) 171–182.
[19] L. Tang, J.Y. Tang, S.J. Cheng, Bayesian statistical fusion reliability evaluation model for multi-source heterogeneous data, Mechanical Strength 44 (1) (2022) 126–132.
[20] T.X. Xu, Y. Liu, J.Z. Zhao, Y.T. Han, Fusion method of maintainability prior information, Syst. Eng. Electr. Technol. 36 (9) (2014) 1887–1892.
[21] D. Xu, Z. Guan, C. Li, Maintainability multi-source data fusion method based on superposition degree, Fire Control Command Control 45 (4) (2020) 35–39.
[22] J.Y. Miao, P.C. Yan, G.Y. Lian, X.S. Zhang, Information fusion method before maintainability test, Fire Control Command Control 42 (8) (2017) 145–149.
[23] A. Richard, Chechile, A Bayesian analysis for the Wilcoxon signed-rank statistic, Commun. Stat. Theor. Methods 47 (21) (2018) 5241–5254.
[24] D. Xu, C. Guan, C. Zhou, Research on consistency checking method of equipment maintainability multi-source pre-test data, Aerospace Control 38 (6) (2020) 61–66.
[25] J. Sun, G.L. Li, C. Xu, Evaluation method for test accuracy of anti-ship missile range, Fire Control Command Control 39 (10) (2014) 95–98.
[26] X.J. Duan, H.Y. Huang, Weighted fusion evaluation of supplementary samples based on information divergence, Acta Armamentarii 127 (10) (2007) 1276–1280.
[27] A.L.M. Levada, On the numerical computation of Fisher–Rao based distances and KL-divergences between Gaussian random fields, Journal of Computational Science 62 (2022).
[28] B. Guo, P. Jiang, Y.Y. Xing, A censored sequential posterior odd test (SPOT) method for verification of the mean time to repair, IEEE Trans. Reliab. 57 (2) (2008) 243–247.
[29] K. Wang, X.J. Shi, L. Qin, X.H. Nie, Y.F. Long, Design of weapon equipment testability verification test scheme based on Bayes small sample theory and sequential network diagram inspection, Acta Armamentarii 40 (11) (2019) 2319–2328.