Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 28 (2026), Issue 3
journal homepage: http://www.ein.org.pl

Article citation info:
Hu C, Meng D, Shi X, A generalized multi-stage conditional probability inference framework with heterogeneous information fusion, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(3) http://doi.org/10.17531/ein/218335

A generalized multi-stage conditional probability inference framework with heterogeneous information fusion

Indexed by: Web of Science Group

Cenyu Hu[a], Desheng Meng[b], Xianming Shi[a,*]

[a] Army Engineering University, China
[b] 31693 unit of the People's Liberation Army, China

Highlights

- Proposes a generalized Bayesian framework for small-sample multi-level performance assessment.
- Reduces complexity via multi-stage conditional probability inference modeling.
- Introduces system contribution degree for dynamic heterogeneous information fusion.
- Establishes sequential constraints to validate performance improvements across batches.
- Demonstrates superior accuracy in capturing material performance evolutionary trends.

Abstract

In complex system evaluation and decision-making under uncertainty, accurately assessing multi-level performance states using small-sample data remains a fundamental challenge. Traditional multinomial models for multi-grade classification are plagued by high-dimensional parameter spaces and computational intractability, whereas standard binomial approaches oversimplify the inherent hierarchical characteristics of degradation or improvement processes. To address this gap, this paper proposes a novel, generalized Bayesian inference framework built upon a newly proposed multi-stage conditional probability model. Concurrently, the system contribution degree is introduced as the weight for Bayesian fusion, thereby enabling dynamic integration of information from each experimental batch. The Gibbs sampling Markov chain Monte Carlo (MCMC) algorithm is adopted for posterior inference, with strict sequential constraints established to validate performance improvements across successive batches. The effectiveness of the proposed method under small-sample scenarios is verified through multiple batches of performance enhancement tests conducted during the iterative development of the key strategic material. Results demonstrate that this framework more accurately and stably captures the evolutionary trend of material performance, offering the scientific, systematic, and universally applicable solution for multi-batch performance evaluation with small samples.

Keywords
bayesian inference framework, multiple batch test, conjugate prior, system contribution degree, MCMC

1. Introduction

In the field of modern complex systems engineering, assessing the state of multi-stage systems with progressive evolutionary characteristics presents a core challenge. Such systems are prevalent in industrial production, supply chain management, equipment health management, and other domains [1]. For instance, during the full lifecycle performance validation of high-value assets—such as precision instruments or critical spare parts—their functional state progressively improves through iterative batch testing. Following each test, the R&D team implements targeted optimizations to the material's design parameters, manufacturing processes, or usage strategies based on the evaluation results of preceding batches. This aims to

(*) Corresponding author.
E-mail addresses: C. Hu (ORCID: 0009-0003-9827-7525) 18604183824@163.com, D. Meng (ORCID: 0009-0006-9833-2952) 78790944@qq.com, X. Shi (ORCID:0000-0003-2165-0013) 3635557436@qq.com

enhance its ultimate reliability and performance metrics. Consequently, the probability distributions governing the test data from each batch are not homogeneous but exhibit significant dynamic variation—specifically, the system's performance progressively improves with each successive batch.

Existing state assessment studies predominantly focus on single-observation scenarios, utilizing only data from a single time point or batch of experiments, and assuming all data originate from the same population. For such problems, mainstream methods include Monte Carlo simulation, binomial success-failure analysis, and multinomial distribution models. Monte Carlo simulation approximates true probabilities through extensive random sampling, but its computational cost is high, particularly when accuracy is difficult to guarantee under small sample conditions. Traditional binomial distribution models can only categorize results as "success" or "failure," failing to reflect the gradual evolutionary changes in system performance. To address this issue, the multinomial distribution model has been proposed to describe multiple discrete states. Liu et al. [2] estimated the hit probability of ammunition based on the multinomial distribution can more comprehensively characterize the damage effects of different levels. However, due to its high-dimensional parameter space, the multinomial distribution model faces severe computational complexity in practical solutions [3]. Particularly under small-sample conditions, its posterior integral is difficult to solve analytically, leading to slow convergence of the Markov Chain Monte Carlo (MCMC) method, which fails to meet the requirements of real-time evaluation. In addition, traditional modeling often assumes that all historical information has equal credibility, which is prone to causing posterior estimation bias in a multi-source heterogeneous environment and undermining the robustness of evaluation results [4].

Therefore, making full use of incremental information from multi-batch tests is the key to achieving accurate evaluation under small-sample conditions [5]. Due to its powerful information fusion capability, the Bayesian method has shown great potential in this field, integrating multi-source heterogeneous information such as historical test data, simulation prediction results, and expert experience into a unified probabilistic reasoning framework [6]. Li et al. [7] predicted ammunition requirements under small-sample conditions by selecting the conjugate Dirichlet prior of the multinomial distribution; Papananias et al. [8] regarded the previous posterior distribution as a new prior distribution to realize the dynamic update of end-product quality. In recent years, Bayesian optimization has been widely applied in adaptive experimental design to maximize information gain and guide the allocation of subsequent test resources [9]. To address the fusion of multi-source uncertain information, the GEJS measure proposed by Xiao et al. [10] introduces the generalized Jensen-Shannon divergence to handle more complex multi-source conflicting information; Huang et al. [11] proposed the belief divergence measure in the complex domain for dealing with more complex multi-source conflicting information.

However, it is unreasonable to directly fuse all historical information on an equal basis. In practical applications, test data closer to the current evaluation reflects the performance state more accurately and has higher information value [12]. Therefore, it is necessary to scientifically quantify the relative importance of information from different sources [13]. This paper introduces the concept of "system contribution degree" and dynamically weights the reliability of prior information from different sources through expert scoring and the intuitionistic fuzzy membership function method. Hu et al. [14] constructed the Bayesian prior information fusion framework based on evidence theory, which is capable of integrating different types of prior information; Xiao et al. [10] proposed a generalized evidence divergence measure to evaluate the degree of conflict between information sources; Meanwhile, Tao et al. [15] have shown that effectively improving the robustness of the fusion system can be achieved by differentially modeling the statistical characteristics of different information sources. In addition, Li et al. [16] proved that the distributed Bayesian fusion algorithm has the property of consistent convergence, providing convergence guarantee for the multi-batch tests in this paper.

To reduce the complexity introduced by the multinomial distribution, this paper constructs a multi-stage binomial distribution model, converting the evaluation of k discrete states into a chain inference of k-1 continuous conditional probabilities. By dimensionality reduction, the high-dimensional problem is decomposed into multiple low-dimensional binomial distribution sub-problems, which greatly reduces the complexity of model solution. Importance sampling [17], shared prior fusion [18], and the constrained Bayesian decision framework [19] further provide a new paradigm for efficient inference of complex models. Fan et al. [20] proposed a global probability distribution to reduce the complexity in vector regression; Tuia et al. [21] introduced conditional random fields (CRF) to solve the high-dimensional remote sensing image classification problem; de Oude et al. [22] and Glinton et al. [23] demonstrated methods that can significantly improve the accuracy of high-order information fusion.

In summary, this paper proposes a general and reusable multi-stage dynamic evaluation framework for system states [24]. It fully draws on and develops recent advanced achievements in generalized evidence measures, heterogeneous Bayesian fusion, sequential inference, and adaptive experimental design, providing the more scientific, accurate, and robust solution for the effectiveness evaluation of complex systems under small-sample conditions.

2. Multi-stage binomial distribution material efficacy test

During the performance test of materials, it is necessary to first establish a clear classification of target performance states, and reasonably set assumptions during the testing process.

2.1. Multi-level target performance status classification

According to the scope of the material's effect and the degree of impact on the target, different levels of performance status changes will occur [2]. The performance status of the target is divided into five grades, as shown in Table 1.

Table 1. Classification of performance status for different target levels.

Performance improvement	Performance outcome description
No performance improvement	The materials did not have a significant effect or only caused minor impacts, and the loss of target functionality and performance was within 5%.
Mild performance improvement	The product produces a minor effect, resulting in 5% to 20% improvement in the target function performance.
Moderate performance improvement	The target function has been significantly improved. Components need to be replaced for optimization, and the performance has been enhanced by 20% to 50%.
Significant performance improvement	The target function has been fundamentally enhanced and requires a factory-level overhaul and the performance of the function has been improved by 50% to 80%.
Completely upgraded	The target has been completely restructured, achieving new functions and increasing combat effectiveness by over 80%.

2.2. Material effectiveness verification model based on multi-stage binomial distribution

The existing materials are mainly analyzed from the perspectives of success/failure and multiple distributions. That is, assuming the probability of achieving the performance standard is p, and n batches of materials are used, then the probability of achieving the performance standard is expressed as:

$$f(x) = C_n^x p^x (1-p)^{n-x} \tag{1}$$

where x represents the number of batches that have achieved the efficiency standard., and $C_n^x = \frac{n!}{x!(n-x)!}$ denotes the combination, p represents the probability of achieving the standard for material efficiency.

The multiple distribution is divided into five different performance states in Table 1. Each outcome has a certain probability of occurrence, which is denoted as $P_i, i = 1,2,3,4$. The probability of achieving the performance standard is expressed as:

$$P(X_0 = x_0, X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4) = \frac{n!}{x_0! x_1! x_2! x_3! x_4!} p_0^{x_0} p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4} \tag{2}$$

where x_0, x_1, x_2, x_3, x_4 represents the number of materials in different performance states., satisfying $x_0 + x_1 + x_2 + x_3 + x_4 = n$; $p = \{p_0, p_1, p_2, p_3, p_4\}$ denotes the probability set for each level.

Although multi-distribution models can more comprehensively reflect performance outcomes, their complexity makes model solving challenging. In such cases, the model can be converted to a multi-stage binomial distribution for analysis. The mild performance improvement standard encompasses mild and higher performance improvement levels; the moderate standard includes moderate and higher levels; the severe standard covers severe and higher levels; and the complete reconstruction standard corresponds to the complete reconstruction level.

The original multinomial distribution is decomposed into $k-1$ independent conditional binomial distributions. The mild performance improvement standard efficacy test is expressed as:

$$L(s_1|p_1) = C_{n_1}^{s_1} p_1^{s_1} (1-p_1)^{n_1-s_1} \quad (3)$$

where $n_1 = n, s_1 = n - x_0$

The standard performance test for moderate performance improvement is expressed as:

$$L(s_2|p_2) = C_{n_2}^{s_2} p_2^{s_2} (1-p_2)^{n_2-s_2} \quad (4)$$

In this way, the probability of each condition meeting the efficiency standard can be described by an independent binomial distribution.

Specifically, for the j -th stage$(j = 1,2,3,4)$, let n_j be the number of effective trials that enter this stage (i.e., the "success" count from the previous stage), and x_j be the number of "successes" achieved or exceeding the corresponding performance level in this stage. Then, the probability p_j of meeting the performance standards follows the following binomial distribution:

$$L(x_j|p_j) = C_{n_j}^{x_j} p_j^{x_j} (1-p_j)^{n_j-x_j} \quad (5)$$

The specific definitions of the number of trials n_j and the number of successful outcomes s_j at each stage are shown in Table 2.

Table 2. Parameter definition of multi-stage binomial distribution model.

Stage j	Damage	Number of effective trialsn_j
$j = 1$	Mild and above	$n_1 = x_1 + x_2 + x_3 + x_4$
$j = 2$	Moderate and above	$n_1 = x_2 + x_3 + x_4$
$j = 3$	Severely and above	$n_1 = x_3 + x_4$
$j = 4$	Complete reconstruction	$n_1 = x_4$

Table 2 converts the assessment of 5 discrete damage levels into the assessment of 4 continuous conditional probabilities, where each conditional probability p follows an independent binomial distribution $B(n_j, p_j)$.

3. Multi-batch experimental information fusion based on system contribution degree

In multi-batch material performance trials, data from different batches do not carry equal weight. Typically, trials conducted closer to the final evaluation batch reflect material performance states that more closely align with current levels, thereby yielding higher informational value. Therefore, when integrating multi-source information, it is essential to scientifically quantify the relative importance of prior information across each batch.

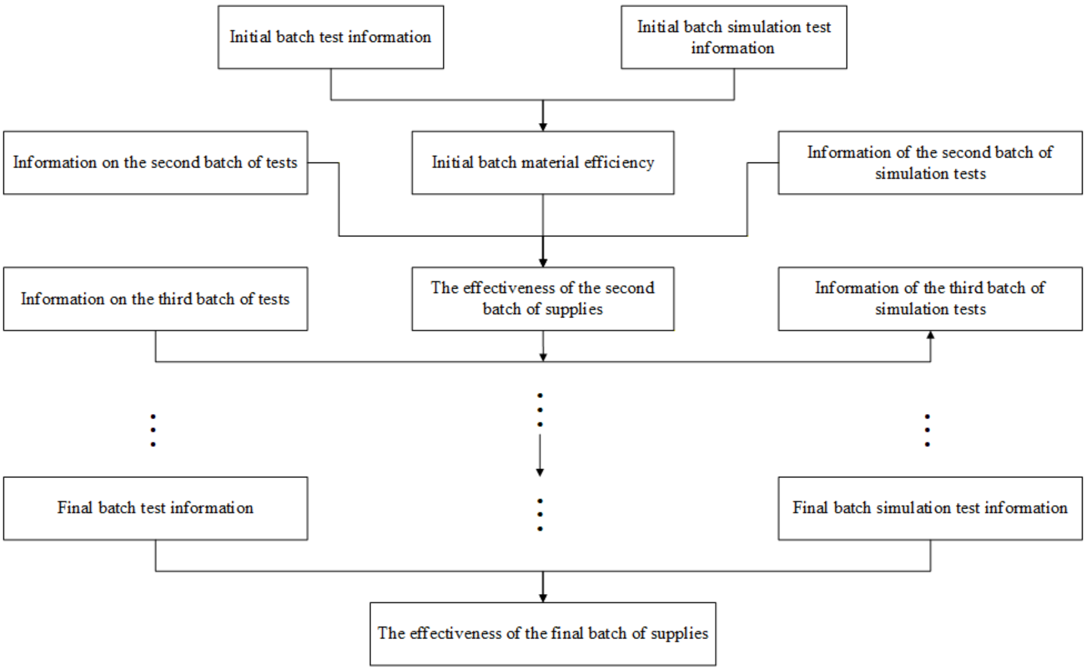


Figure 1. Framework of Information Fusion in Multi-batch Experiments.

3.1. Multi-batch Information Fusion Framework

The sources of information for multi-batch material performance tests mainly include: the actual test data of the current batch, the performance evaluation results of the previous batches, and the simulation test information of this batch. The information fusion process is shown in Figure 1.

3.2. Calculation of fusion weights based on system contribution degree

System contribution can be used to evaluate the importance of materiel within an operational system, and can be extended to assess the influence of various evaluation factors on the overall system. This paper introduces a framework to quantify the influence of prior information from different sources—specifically, results from previous test batches and simulation data from the current batch—on the material effectiveness evaluation system. The strength of this influence primarily depends on the reliability of the information source. By analyzing the credibility of information sources, their weight in the Bayesian fusion process can be determined. More reliable information corresponds to a higher system contribution and is assigned a greater fusion weight.

Specifically, for the analysis of the i -th batch, the required information includes the shooting test data of the current batch, the ammunition hit probability information of the $i-1$-th batch, and the simulation test data of the i -th batch. The prior information for constructing this batch is derived from the ammunition hit probability information of the i -th batch and the simulation test of the i -th batch. Using the method of system contribution degree, the specific fusion weights of these two parts of prior information are determined.

Let $\delta = \{\delta_1, \delta_2\}$ be the prior information set of the ammunition hit probability, where δ_1 represents the hit probability information of the previous batch, and δ_2 represents the simulation test information of the current batch. Invite n_e domain experts to evaluate the reliability of information $\delta_l (l = 1,2)$. Let $w_{\delta_l}^d$ represent the degree of reliability that the d-th expert considers information to be, and $v_{\delta_l}^d$ represent the degree of unreliability that the d-th expert considers information to be. Using the intuitionistic fuzzy membership function method, the membership relationship of the experts to information δ_l can be expressed as follows [12]:

$$\Theta_{\delta_l}^d = \{\delta_l, (w_{\delta_l}^d, v_{\delta_l}^d)\} \tag{6}$$

where $0 \le w_{\delta_l}^d \le 1, 0 \le v_{\delta_l}^d \le 1$ and $w_{\delta_l}^d + v_{\delta_l}^d \le 1$.

When using the intuition-fuzzy membership method to evaluate the reliability of the information sources for the probability of ammunition hitting targets, considering that experts may have uncertainty or indecision in judging the credibility, an intuition indicator is introduced to quantify this degree of hesitation. The expression is:

$$\tau_{\delta_l}^d = 1 - w_{\delta_l}^d - v_{\delta_l}^d \tag{7}$$

When determining the fusion weights of the prior information, half of the expert's hesitation degree is adopted as the correction factor, and the membership degree of the information source δ_l can be calculated. The expression is as follows [25]:

$$\varepsilon_{\delta_l} = \frac{1}{n_e}\sum_{d=1}^{n_e}(w_{\delta_l}^d - v_{\delta_l}^d) + \frac{1-\frac{1}{n_e}\sum_{d=1}^{n_e}(w_{\delta_l}^d - v_{\delta_l}^d)}{2} \tag{8}$$

After computing the unnormalized contribution scores, the final fusion weights are obtained by normalization as follows:

$$\varepsilon_{\delta_l}^\tau = \frac{\varepsilon_{\delta_l}}{\varepsilon_{\delta_1} + \varepsilon_{\delta_2}} \tag{9}$$

where ε_{δ_1} represents the degree of membership of the efficacy evaluation information for the $i-1$ -th batch of ammunition, and ε_{δ_2} represents the degree of membership of the information from the i -th batch of simulation experiments. Based on the fusion weight $\varepsilon_{\delta_l}^\tau$, the mixed prior distribution for the effectiveness evaluation of the i -th batch of materials can be calculated:

$$\pi(P^{(i)}) = \varepsilon_{\delta_1}\pi_1(P^{(i)}) + \varepsilon_{\delta_2}\pi_2(P^{(i)}) \tag{10}$$

where $P^{(i)}$ represents the probability set of achieving different performance levels for the materials; $\pi_1(P^{(i)})$ is the distribution based on the posterior results of the $i-1$-th batch as the prior; $\pi_2(P^{(i)})$ is the prior distribution provided by the simulation experiment of the i -th batch.

4. Bayesian inference for the probability of achieving the test effectiveness standards for multiple batches of materials

Unlike studies that only analyze the data from a single batch of tests, multi-batch material per-formance tests require the integration of all historical information. Starting from the initial batch, the dynamic analysis of the probability of each batch of materials meeting the performance stand-ards is conducted.

4.1. Bayesian inference for the probability of initial batch performance meeting standards

For the initial batch, the assessment of the probability of the material performance meeting the standards is solely based on the actual performance test data $L(x^{(1)}|p^{(1)})$ of this batch and the prior information $\pi(p^{(1)}|\alpha^{(1)})$ from the simulation experiments of this batch.

For the j -th stage ($j = 1,2,3,4$), its likelihood function is:

$$L(x_j^{(1)}|p_j^{(1)}) = \binom{n_j^{(1)}}{x_j^{(1)}} \left(p_j^{(1)}\right)^{x_j^{(1)}} \left(1-p_j^{(1)}\right)^{n_j^{(1)}-x_j^{(1)}} \quad (11)$$

where $n_1^{(1)} = n^{(1)}, s_1^{(1)} = x_1^{(1)} + x_2^{(1)} + x_3^{(1)} + x_4^{(1)}, n_2^{(1)} = s^{(1)}, s_2^{(1)} = x_2^{(1)} + x_3^{(1)} + x_4^{(1)}, n_3^{(1)} = s_2^{(1)}, s_3^{(1)} = x_3^{(1)} + x_4^{(1)}$, $n_4^{(1)} = s_3^{(1)}, s_4^{(1)} = x_4^{(1)}$, $p_j^{(1)}$ represents the probability of achieving the performance standard in the initial batch trial of the j -th stage, and $x_j^{(1)}$ represents the number of "successful" occurrences that reach or exceed the corresponding performance level in this stage.

Following the Bayesian conjugate prior principle, the Beta distribution is chosen as the prior distribution:

$$\pi(p_j^{(1)}|\alpha_j^{(1)},\beta_j^{(1)}) = \frac{\Gamma(\alpha_j^{(1)}+\beta_j^{(1)})}{\Gamma(\alpha_j^{(1)})\Gamma(\beta_j^{(1)})}(p_j^{(1)})^{\alpha_j^{(1)}-1}(1-p_j^{(1)})^{\beta_j^{(1)}-1} \quad (12)$$

where $\alpha_j^{(1)},\beta_j^{(1)}$ represent the hyperparameters of the prior distribution of the initial batch in the j -th stage.

By using the Bayesian method to integrate the above information, the posterior distribution of the probability of the initial batch of materials meeting the efficiency standards can be calculated:

$$\pi\left(p_j^{(2)}|s_j^{(2)}\right) = \frac{L\left(s_j^{(1)}\middle|p_j^{(1)}\right)\pi\left(p_j^{(1)}|\alpha_j^{(1)},\beta_j^{(1)}\right)}{\int_0^1 L\left(s_j^{(1)}\middle|p_j^{(1)}\right)\pi\left(p_j^{(1)}|\alpha_j^{(1)},\beta_j^{(1)}\right)dp_j^{(1)}} =$$

$$= \frac{\frac{n_j^{(i)}!}{s_j^{(1)}!\left(n_j^{(i)}-s_j^{(1)}\right)}\cdot\frac{\Gamma\left(\alpha_j^{(1)}+\beta_j^{(1)}\right)}{\Gamma\left(\alpha_j^{(1)}\right)\Gamma\left(\beta_j^{(1)}\right)}(p_j^{(1)})^{\alpha_j^{(1)}+s_j^{(1)}-1}(1-p_j^{(1)})^{\alpha_j^{(1)}+n_j^{(i)}-s_j^{(1)}-1}}{\frac{n_j^{(i)}!}{s_j^{(1)}!\left(n_j^{(i)}-s_j^{(1)}\right)}\cdot\frac{\Gamma\left(\alpha_j^{(1)}+\beta_j^{(1)}\right)}{\Gamma\left(\alpha_j^{(1)}\right)\Gamma\left(\beta_j^{(1)}\right)}\frac{\Gamma\left(\alpha_j^{(1)}+s_j^{(1)}\right)\Gamma\left(\beta_j^{(1)}+n_j^{(i)}-s_j^{(1)}\right)}{\Gamma\left(\alpha_j^{(1)}+s_j^{(1)}+\beta_j^{(1)}+n_j^{(i)}-s_j^{(1)}\right)}} = \frac{\Gamma\left(\alpha_j^{(1)}+s_j^{(1)}+\beta_j^{(1)}+n_j^{(i)}-s_j^{(1)}\right)}{\Gamma\left(\alpha_j^{(1)}+s_j^{(1)}\right)\Gamma\left(\beta_j^{(1)}+n_j^{(i)}-s_j^{(1)}\right)}(p_j^{(1)})^{\alpha_j^{(1)}+s_j^{(1)}-1}(1-p_j^{(1)})^{\alpha_j^{(1)}+n_j^{(i)}-s_j^{(1)}-1} \quad (13)$$

where $n_j^{(1)}$ represents the effective sample size of the initial batch trial in the j -th stage. From the above equation, it can be seen that the posterior distribution also follows the Beta distribution:

$$\pi\left(p_j^{(1)}|s_j^{(1)}\right) = Beta\left(\alpha_j^{(1)}+s_j^{(1)},\beta_j^{(1)}+n_j^{(1)}-s_j^{(1)}\right) \quad (14)$$

Based on the posterior distribution, the Bayesian expected value of the probability of the initial batch material performance reaching the standard in the j -th stage can be calculated:

$$E\left(p_j^{(1)}\right) = \frac{\alpha_j^{(1)}+s_j^{(1)}}{\alpha_j^{(1)}+\beta_j^{(1)}+n_j^{(1)}} \quad (15)$$

4.2. Bayesian inference for the probability of achieving performance standards in the second batch

After completing the initial batch of tests and optimizing the design parameters, the second batch of performance tests will be conducted. To evaluate the performance of this batch, it is necessary to fully consider and utilize the information accumulated from the previous batches. Therefore, the analysis information of the second batch includes three parts: the measured performance test data of the second batch $L(x^{(2)}|p^{(2)})$, the simulation experiment information of the second batch $\pi(p^{(2)}|\alpha^{(2)})$, and the posterior results of the performance evaluation after the initial batch $\pi(p^{(1)}|s^{(1)})$. Compared to the initial batch, the second batch incorporated historical trial information as prior knowledge. According to the system contribution degree method proposed in Section 3, the fusion weights $\varepsilon_{\delta_1}^{\tau}$ and $\varepsilon_{\delta_2}^{\tau}$ of historical information and simulation information can be determined. Therefore, the prior distribution of the j -th stage in the second batch is constructed as a mixed Beta distribution:

$$\pi\left(p_j^{(2)}\right) = \varepsilon_{\delta_1}^{\tau}Beta\left(a_j^{(1)},b_j^{(1)}\right) + \varepsilon_{\delta_2}^{\tau}Beta\left(\alpha_j^{(2)},\beta_j^{(2)}\right) \quad (16)$$

Combining the likelihood function $L(s_j^{(2)}|p_j^{(2)})$ of this batch, according to Bayes' theorem, the posterior distribution of the j -th stage of the second batch is:

$$\pi(p_j^{(2)}|s_j^{(2)}) = \frac{L(s_j^{(2)}|p_j^{(2)})\pi(p_j^{(2)})}{\int_0^1 L(s_j^{(2)}|p_j^{(2)})\pi\left(p_j^{(2)}\right)dp_j^{(2)}}$$

$$= \frac{L(s_j^{(2)}|p_j^{(2)})\varepsilon_{\delta_1}^{\tau}Beta(\alpha_j^{(1)},\beta_j^{(1)}) + \varepsilon_{\delta_2}^{\tau}Beta(\alpha_j^{(2)},\beta_j^{(2)})}{\int_0^1 L(s_j^{(2)}|p_j^{(2)})\pi\left(p_j^{(2)}\right)dp_j^{(2)}} \quad (17)$$

If the complexity of the mixing term is ignored, it can be approximately assumed that the posterior distribution remains the Beta distribution. Then the posterior is:

$$\pi\left(p_j^{(2)}|s_j^{(2)}\right) = Beta\left(a_j^{(1)}, b_j^{(1)}\right) + Beta\left(\alpha_j^{(1)} + s_j^{(2)}, b_j^{(1)} + n_j^{(2)} - s_j^{(2)}\right) \quad (18)$$

Based on the above posterior distribution, the posterior expected value of the probability of achieving the performance standard in the j -th stage of the second batch is:

$$E\left(p_j^{(2)}\right) = \frac{\alpha_j^{(1)} + s_j^{(2)}}{\alpha_j^{(1)} + \beta_j^{(1)} + n_j^{(2)}} \quad (19)$$

This estimation result will be input as the prior information for the subsequent batch (the third batch), enabling the iterative update of the knowledge for performance evaluation.

4.3. Bayesian inference for the probability of achieving efficiency standards in the i-th batch

For the analysis of the i -th batch ($i > 2$), the information composition follows the same pattern, including: the measured efficacy test data $L(x^{(i)}|p^{(i)})$ for this batch, the simulation experiment information $\pi(p^{(i)}|\alpha^{(i)})$ for this batch, and the posterior result of the efficacy assessment from the previous batch (the $i-1$-th batch) $\pi(p^{(i-1)}|s^{(i-1)})$). Here, $\pi(p^{(i-1)}|s^{(i-1)})$ has integrated all the historical assessment information from the initial batch to the $i-1$-th batch [26].

The probability of the effectiveness of the materials in the i -th batch and the j -th stage meeting the standards is expressed by the Bayesian formula as follows:

$$\pi\left(p_j^{(i)}|s_j^{(i)}\right) = \frac{L(s_j^{(i)}|p_j^{(i)})\pi\left(p_j^{(i)}\right)}{\int_0^1 L(s_j^{(i)}|p_j^{(i)})\pi\left(p_j^{(i)}\right)dp_j^{(i)}} \quad (20)$$

The likelihood function for the i -th batch and the j -th stage is:

$$L(s_j^{(i)}|p_j^{(i)}) = \binom{n_j^{(i)}}{s_j^{(i)}}\left(p_j^{(i)}\right)^{s_j^{(i)}}\left(1 - p_j^{(i)}\right)^{n_j^{(i)} - s_j^{(i)}} \quad (21)$$

where

$$n_1^{(i)} = n^{(i)}, s_1^{(1)} = \sum_{k=1}^{4} x_k^{(i)}, n_j^{(i)} = s_{j-1}^{(i)}, j = 2,3,4, s_j^{(i)} = \sum_{k=j}^{4} x_k^{(i)}, j = 2,3,4$$

The prior distribution is obtained through weighted fusion based on the contribution degree of the system:

$$\pi(p_j^{(i)}) = \varepsilon_{\delta_1}^{\tau\ (i-1)} Beta(a_j^{(i-1)}, b_j^{(i-1)}) + \varepsilon_{\delta_2}^{\tau\ (i-1)} Beta(\alpha_j^{(i)}, \beta_j^{(i)}) \quad (22)$$

According to Bayes' formula, the posterior distribution for the i -th batch and the j -th stage is:

$$\pi\left(p_j^{(i)}|s_j^{(i)}\right) = Beta\left(a_j^{(i-1)} + s_j^{(i)}, b_j^{(i-1)} + n_j^{(i)} - s_j^{(i)}\right) \quad (23)$$

Based on the above posterior distribution, the posterior expected value of the probability of achieving the performance standard in the i -th batch and the j - th stage is:

$$E(p_j^{(i)}) = \frac{a_j^{(i-1)} + s_j^{(i)}}{a_j^{(i-1)} + b_j^{(i-1)} + n_j^{(i)}} \quad (24)$$

Ultimately, based on the Bayesian inference results of the i -th batch, the final probabilities of the materials achieving each performance level (such as no improvement, mild improvement, moderate improvement, severe improvement, and complete reconfiguration) can be derived, and the following sequential constraint relationships can be established to verify the improvement effect of the material performance:

$$\begin{gathered} E\left(p_1^{(i)}\right) \geq E\left(p_1^{(i-1)}\right) \\ E(p_1^{(i)}p_2^{(i)}) \geq E(p_1^{(i-1)}p_2^{(i-1)}) \\ E(p_1^{(i)}p_2^{(i)}p_3^{(i)}) \geq E(p_1^{(i-1)}p_2^{(i-1)}p_3^{(i-1)}) \\ E(p_1^{(i)}p_2^{(i)}p_3^{(i)}p_4^{(i)}) \geq E(p_1^{(i-1)}p_2^{(i-1)}p_3^{(i-1)}p_4^{(i-1)}) \end{gathered} \quad (2)$$

These constraints respectively indicate that the probabilities of the i -th batch of materials achieving mild and above, moderate and above, severe and above, or even complete reconfiguration are all no less than those of the $i-1$-th batch. Thus, this quantitatively proves that the improvement in material performance is effective.

5. Solution to the probability of achieving performance standard for multiple batches of materials based on the MCMC method

The Bayesian model constructed in Section 3 involves multiple batches of material performance compliance probability parameter $P^{(i)} = (p_1^{(i)}, p_2^{(i)}, p_3^{(i)}, p_4^{(i)})$. Due to the high dimensionality of the joint posterior distribution, an analytical solution is difficult to obtain. Therefore, in this paper, the MCMC method is adopted to numerically simulate and solve the posterior distribution of the parameters through Gibbs sampling. The specific implementation steps are as follows [27]:

Step 1: Initial value setting

To initiate the iterative process, initial values $[P_j^{(i)}]^0$ need to be set for the probability parameters of the performance targets for each batch. To accelerate the convergence of the Markov chain, the expected value of its mixed prior distribution can be selected as the initial point. For the i -th batch and the j -th stage, the prior distribution is the mixed Beta distribution [28]:

$$\pi(p_j^{(i)}) = \varepsilon_{\delta_1}^{\tau\ (i)} Beta(a_j^{(i-1)}, b_j^{(i-1)}) + \varepsilon_{\delta_2}^{\tau\ (i)} Beta(a_j^{(i-1)}, b_j^{(i-1)}) \quad (26)$$

The expected value of this mixed distribution is:

$$E_{prior}[p_j^{(i)}] = \varepsilon_{\delta_1}^{\tau}{}^{(i)} \frac{a_j^{(i-1)}}{a_j^{(i-1)}+b_j^{(i-1)}} + \varepsilon_{\delta_2}^{\tau}{}^{(i)} \frac{a_j^{(i)}}{a_j^{(i)}+b_j^{(i)}} \quad (27)$$

Therefore, the expected value of the mixed prior, denoted as $E_{prior}[p_j^{(i)}]$, can be taken as the initial value of the parameter, with $[p_j^{(i)}]^{(0)} = E_{prior}[p_j^{(i)}]$.

Step 2: Iterative sampling process Iterative sampling process

Starting from the initial value $[p_j^{(i)}]^{(0)}$ as the base, begin the Gibbs iteration. Assuming that the value of parameter $[p_j^{(i)}]^{(t-1)}$ at the t -th iteration is:

$[p_j^{(i)}]^{(t-1)} = ([p_1^{(i)}]^{(t-1)}, [p_2^{(i)}]^{(t-1)}, [p_3^{(i)}]^{(t-1)}, [p_4^{(i)}]^{(t-1)})$, then the t -th iteration process is as follows:

$$\begin{aligned}\pi\left(p_j^{(i)} \mid s_j^{(i)}\right) &\propto L(s_j^{(i)} \mid p_j^{(i)})\pi\left(p_j^{(i)}\right) \\ &= \varepsilon_{\delta_1}^{\tau}{}^{(i)} Beta\left(a_j^{(i-1)} + s_j^{(i)}, b_j^{(i-1)} + n_j^{(i)} - s_j^{(i)}\right) + \varepsilon_{\delta_2}^{\tau}{}^{(i)} Beta\left(\alpha_j^{(i)} + s_j^{(i)}, \beta_j^{(i)} + n_j^{(i)} - s_j^{(i)}\right)\end{aligned} \quad (28)$$

Sampling from the mixed distribution can be done using the component sampling method: First, draw a sample from the first $Beta(a_j^{(i-1)} + s_j^{(i)}, b_j^{(i-1)} + n_j^{(i)} - s_j^{(i)})$ using the probability weight $\varepsilon_{\delta_1}^{\tau}{}^{(i)}$; Then, using the probability weight $\varepsilon_{\delta_2}^{\tau}{}^{(i)}$, the sample is drawn from the second distribution $Beta(\alpha_j^{(i)} + s_j^{(i)}, \beta_j^{(i)} + n_j^{(i)} - s_j^{(i)})$.

Update all parameters in sequence according to this rule:

Draw $[p_1^{(i)}]^{(t)}$ from the posterior univariate distribution $\pi_1\left(p_1^{(i)}|s_1^{(i)}, [p_2^{(i)}, p_3^{(i)}, p_4^{(i)}]^{(t-1)}\right)$ of material performance parameters meeting the target specifications.

Draw $[p_2^{(i)}]^{(t)}$ from the posterior univariate distribution $\pi_2\left(p_2^{(i)}|s_2^{(i)}, [p_1^{(i)}, p_3^{(i)}, p_4^{(i)}]^{(t-1)}\right)$ of material performance parameters meeting the target specifications.

Draw $[p_3^{(i)}]^{(t)}$ from the posterior univariate distribution $\pi_3\left(p_3^{(i)}|s_3^{(i)}, [p_1^{(i)}, p_2^{(i)}, p_4^{(i)}]^{(t-1)}\right)$ of material performance parameters meeting the target specifications.

Draw $[p_4^{(i)}]^{(t)}$ from the posterior univariate distribution $\pi_4\left(p_4^{(i)}|s_4^{(i)}, [p_1^{(i)}, p_2^{(i)}, p_3^{(i)}]^{(t-1)}\right)$ of material performance parameters meeting the target specifications.

Thus, the new parameter samples generated in the t -th iteration are:

$$\left[p_j^{(i)}\right]^{(t)} = \left(\left[p_1^{(i)}\right]^{(t)}, \left[p_2^{(i)}\right]^{(t)}, \left[p_3^{(i)}\right]^{(t)}, \left[p_4^{(i)}\right]^{(t)}\right) \quad (29)$$

When the mean value of the hit probability parameter $1/k \cdot \sum_{j=1}^{k}[p_j^{(i)}]^{(t)}$ reaches the convergence state, it is determined that the Markov chain has reached a stable state.

Step 3: Generate posterior samples Generate posterior samples

Since the parameters $p_j^{(i)}$ at each stage are independent of each other, the posterior marginal condition $\pi(p_j^{(i)}|s^{(i)})$ is equivalent to the marginal posterior distribution $\pi(p_j^{(i)}|s_j^{(i)})$.

The form of the posterior distribution $\pi(p_j^{(i)}|s_j^{(i)})$ depends on Eq. (21). Depending on the different forms of the prior, the sampling strategy is as follows:

a) If the prior is a pure Beta (i.e., either $\varepsilon_{\delta_1}^{\tau}{}^{(i)} = 1$ or $\varepsilon_{\delta_2}^{\tau}{}^{(i)} = 1$), then the posterior will also be Beta and can be sampled directly from the posterior Beta distribution.

b) If the prior is a mixture of Beta distributions, then the posterior will also be the mixture of Beta distributions:

Continue the iterative transfer process, generating a new sequence of hit probability parameter points, namely $[p_j^{(i)}]^{(k+1)}, [p_j^{(i)}]^{(k+2)}, \cdots, [p_j^{(i)}]^{(m)}$.

Step 4: Calculate posterior estimation

Remove the first k points from the Markov chain convergence state, and use the new point sequence as the sampling value for the performance evaluation parameter $p_j^{(i)}$. This can yield the posterior expectation of the parameter $E(p_j^{(i)})$.

$$E\left(p_j^{(i)}\right) = \frac{1}{m-k}\sum_{t=k+1}^{m}\left[p_j^{(i)}\right]^{(t)} \quad (30)$$

Step 5: Calculate the final effectiveness probability

After determining the posterior estimates for the probability parameters of performance attainment at each stage, substituting these values into Eq. (31) yields the Bayesian estimates for the probability of achieving each performance level based on the multi-batch growth test:

$$\begin{cases}\left[\hat{p}_1^{(i)}\right]' = \frac{1}{m-k}\sum_{t=k+1}^{m}\left[p_1^{(i)}\right]^{(t)} \\ \left[\hat{p}_2^{(i)}\right]' = \frac{1}{m-k}\sum_{t=k+1}^{m}\left[p_2^{(i)}\right]^{(t)} \\ \left[\hat{p}_3^{(i)}\right]' = \frac{1}{m-k}\sum_{t=k+1}^{m}\left[p_3^{(i)}\right]^{(t)} \\ \left[\hat{p}_4^{(i)}\right]' = \frac{1}{m-k}\sum_{t=k+1}^{m}\left[p_4^{(i)}\right]^{(t)}\end{cases} \quad (30)$$

In addition, according to the sampling results, if it is found that the sequential constraint relationship is not satisfied (that is, the probability of the i -th batch of materials meeting the performance standard is less than that of the $i-1$-th batch), it indicates that the improvement of the material performance before this batch of tests is not satisfactory, and further in-depth research on the material performance is required [29].

The Gibbs sampling process of the aforementioned MCMC method is illustrated in Figure 2.

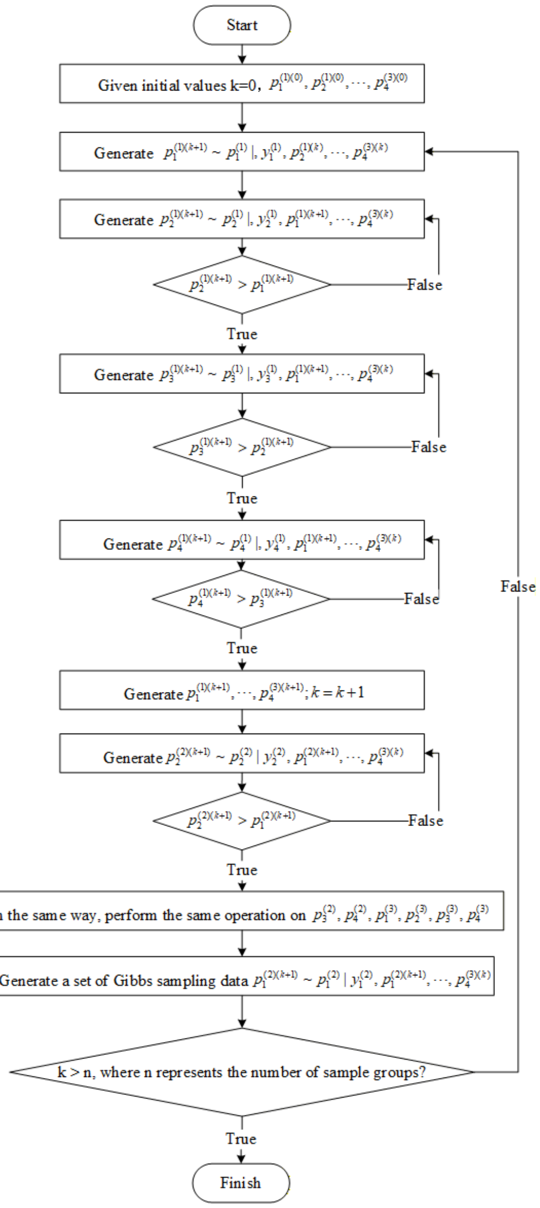


Figure 2. Gibbs Sampling Procedure.

6. Example Analysis

To validate the effectiveness of the proposed method under small-sample conditions, multiple batches of performance enhancement tests were conducted for the iterative development process of a critical strategic material. The tests were divided into three batches, with each batch involving functional testing of 10 material lots. Based on the evaluation results from preceding batches, the R&D team implemented targeted optimizations to the material design, anticipating that its overall functional performance would progressively improve with each subsequent batch [30].

The actual observed frequencies $x^{(i)}$ of each batch and the prior information provided by the simulation experiment are shown in Table 3.

Table 3. Observed frequencies and simulation prior means for each batch.

Batch	Observation frequency	Simulation prior mean
Batch 1	(4, 2, 2, 1, 1)	(0.60, 0.50, 0.40, 0.30)
Batch 2	(3, 1, 3, 2, 1)	(0.70, 0.60, 0.50, 0.40)
Batch 3	(2, 1, 2, 3, 2)	(0.80, 0.70, 0.60, 0.50)

Based on the evaluation criteria for system contribution, six experts in the same field were invited to score the reliability of the post-evaluation information δ_1 for batch $i-1$ and the simulation test information δ_2 for batch i.

The scoring results are shown in Table 4.

Table 4. Expert scoring results for system contribution degree evaluation.

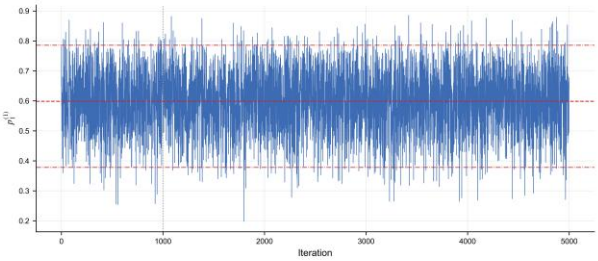
Expert Number	$w_{\delta_1}^d$	$v_{\delta_1}^d$	$w_{\delta_2}^d$	$v_{\delta_2}^d$
1	0.6	0.3	0.4	0.3
2	0.7	0.2	0.5	0.3
3	0.6	0.3	0.4	0.4
4	0.5	0.3	0.5	0.2
5	0.7	0.2	0.4	0.3
6	0.6	0.3	0.5	0.3

All six experts are senior researchers with over 10 years of experience in materials performance evaluation. Prior to scoring, they participated in a consensus workshop to align evaluation criteria. The coefficient of variation (CV) for each information source's membership scores was below 15%, indicating good agreement, and no outlier was removed.

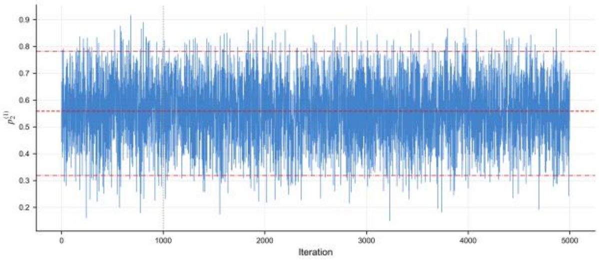
Based on the evaluation criteria for system contribution, six experts in the same field were invited to score the reliability of the post-evaluation information for batch $i-1$ and the simulation test information for batch i. After calculation, the fusion weights for each batch are as follows: Batch 2: $\varepsilon_{\tau\delta_1}^{(2)} = 0.60, \varepsilon_{\tau\delta_2}^{(2)} = 0.40$, Batch 3: $\varepsilon_{\tau\delta_1}^{(3)} = 0.65, \varepsilon_{\tau\delta_2}^{(3)} = 0.35$.

This paper employs the PyMC3 library on the Python platform to implement the MCMC method for calculating the probability of material performance meeting standards. The Gibbs sampling algorithm is used to iterate parameters for each batch, running 5,000 iterations per batch. The first 1,000 iterations are discarded as burn-in to ensure sufficient convergence of the Markov chain.

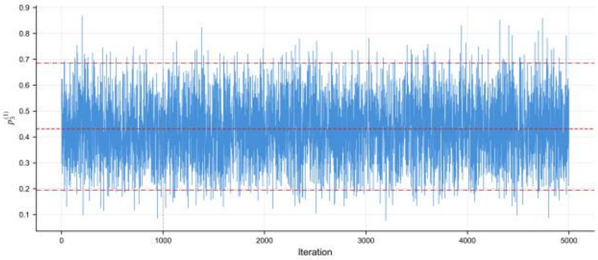
Based on the initial batch test data and initial simulation results, an iterative trace plot of parameter $P^{(1)} = (P_1^{(1)}, P_2^{(1)}, P_3^{(1)}, P_4^{(1)})$ was generated, with the outcome shown in Figure 3.



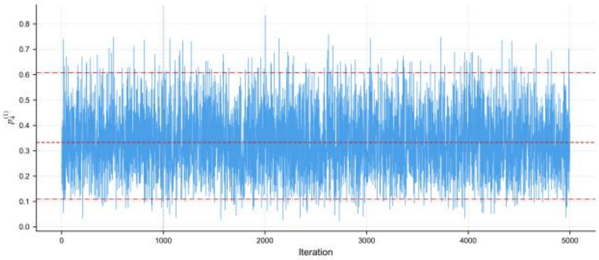
(a) Parameter $p_1^{(1)}$ iteration trace graph



(b) Parameter $p_2^{(1)}$ iteration trace graph

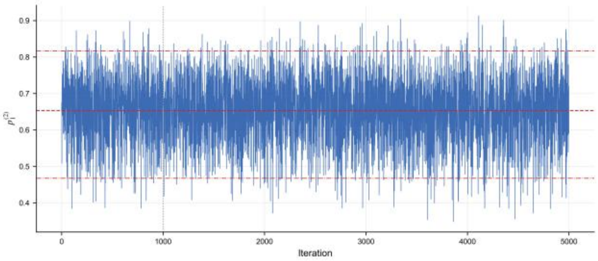


(c) Parameter $p_3^{(1)}$ iteration trace graph

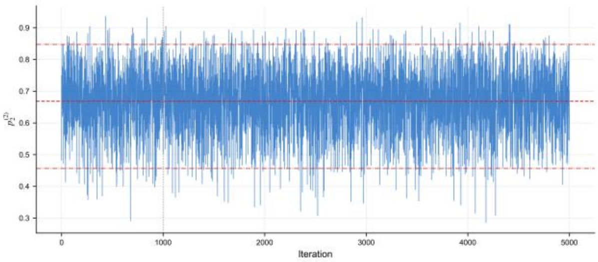


(d) Parameter $p_4^{(1)}$ iteration trace graph

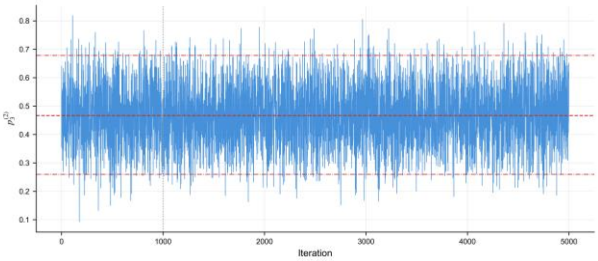
Figure 3. Iteration trace diagram of efficiency-achieving parameters for the first batch.



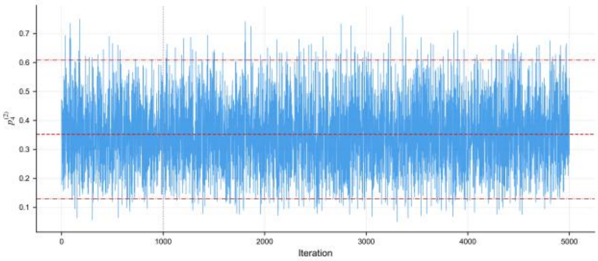
(a) Parameter $\boldsymbol{p}_1^{(2)}$ iteration trace graph



(b) Parameter $\boldsymbol{p}_2^{(2)}$ iteration trace graph



(c) Parameter $\boldsymbol{p}_3^{(2)}$ iteration trace graph



(d) Parameter $\boldsymbol{p}_4^{(2)}$ iteration trace graph

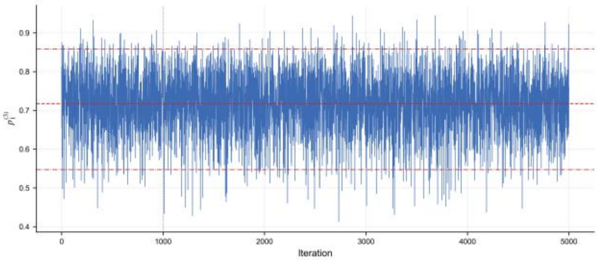
Figure 4. Iteration trace diagram of efficiency-achieving parameters for the second batch.

As can be seen from the figure, all parameter sequences stabilize after approximately 1000 iterations, with stable fluctuations and no obvious trends. This indicates that the Markov chain has reached a convergent state, and the sampling results are reliable.

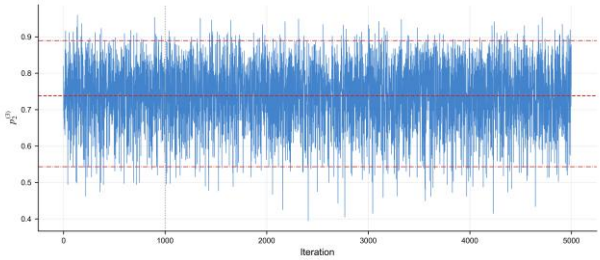
Subsequently, using the Bayesian estimates from the initial batch and simulation data from the second batch as prior information, the efficacy of the second batch was calculated. The resulting parameter iteration trace plot is shown in Figure 4.

Using the Bayesian estimation results of the second batch and the simulation experiment information of the third batch as the priors, the efficacy of the third batch was calculated, and the iterative trace diagram of the generated parameter x was obtained, as shown in Figure 5.

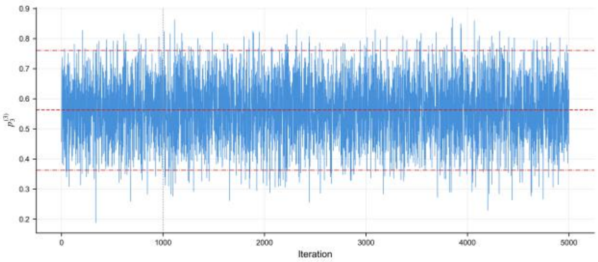
The results in Figures 4 and 5 show that the iterative traces of all parameters exhibit excellent convergence and stability, verifying that the MCMC algorithm can operate stably within the framework of integrating historical information and simulation priors. Furthermore, the posterior probability density function graphs of each batch of parameters were plotted, as shown in Figure 6.



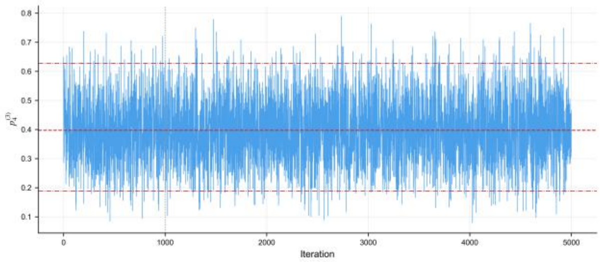
(a) Parameter $\boldsymbol{p_1^{(3)}}$ iteration trace graph



(b) Parameter $\boldsymbol{p_2^{(3)}}$ iteration trace graph

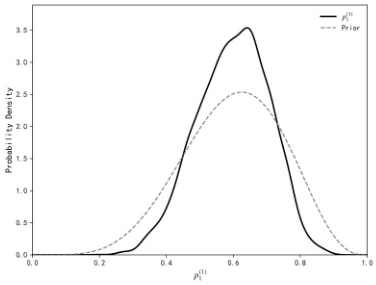


(c) Parameter $\boldsymbol{p_3^{(3)}}$ iteration trace graph

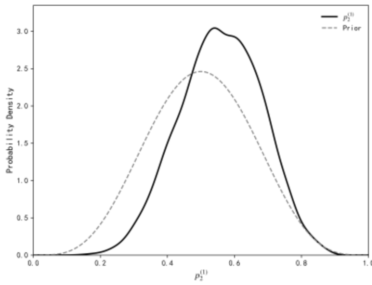


(d) Parameter $\boldsymbol{p_4^{(3)}}$ iteration trace graph

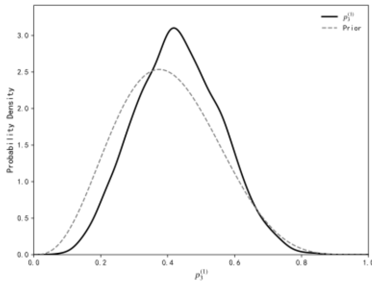
Figure 5. Iteration trace diagram of efficiency-achieving parameters for the third batch.



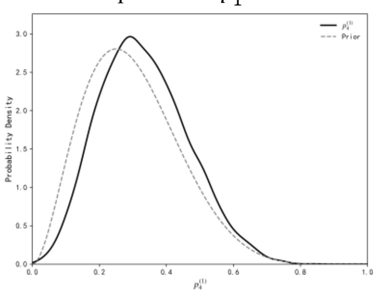
(a) The probability density function diagram of parameter $p_1^{(1)}$



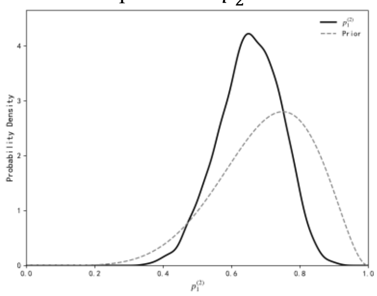
(b) The probability density function diagram of parameter $p_2^{(1)}$



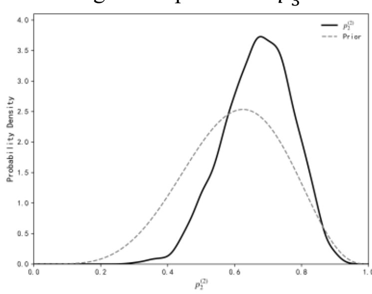
(c) The probability density function diagram of parameter $p_3^{(1)}$



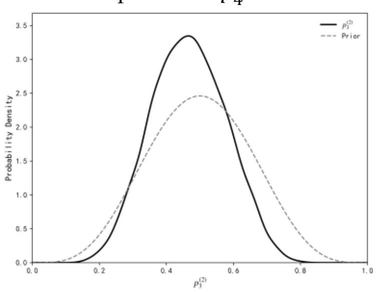
(d) The probability density function diagram of parameter $p_4^{(1)}$



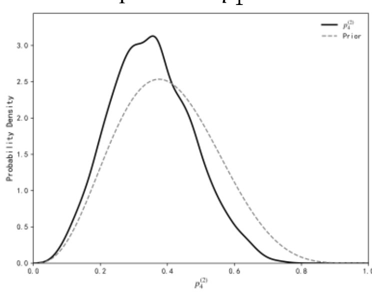
(e) The probability density function diagram of parameter $p_1^{(2)}$



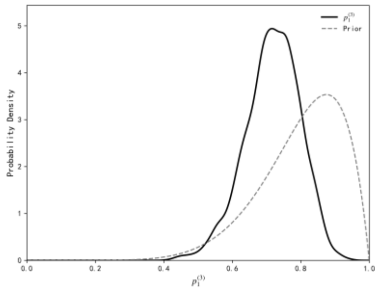
(f) The probability density function diagram of parameter $p_2^{(2)}$



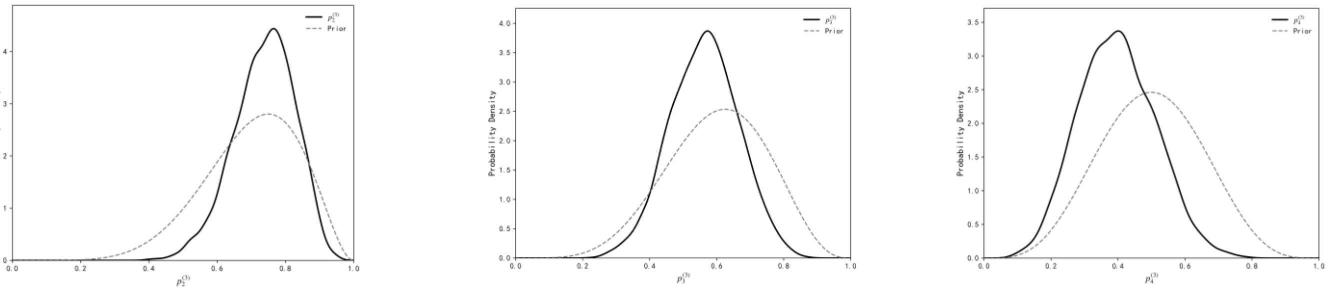
(g) The probability density function diagram of parameter $p_3^{(2)}$



(h) The probability density function diagram of parameter $p_4^{(2)}$



(i) The probability density function diagram of parameter $p_1^{(3)}$



(j) The probability density function diagram of parameter $p_2^{(3)}$ (k) The probability density function diagram of parameter $p_3^{(3)}$ (l) The probability density function diagram of parameter $p_4^{(3)}$

Figure 6. The posterior probability density functions of each batch of parameters.

It can be seen that as the batches progress, the density peaks of $p_1^{(i)}$、$p_2^{(i)}$、$p_4^{(i)}$shift significantly to the right, indicating that the success probability of the corresponding stage has significantly increased; while the distribution of $p_3^{(i)}$ is relatively concentrated, suggesting that the performance improvement in this stage is relatively stable. This graph intuitively reflects the batch-by-batch optimization process of the material efficiency. Figure 6 visually demonstrates the sequential optimization process of the material efficiency.

Statistical analysis was conducted on the Markov chain samples of each batch of parameters, and the posterior means obtained are shown in Table 5.

Table 5. Bayesian posterior mean of the performance parameters for each batch of materials.

Parameter	Initial batch Bayesian estimation of mean	The second batch Bayesian estimation of mean	The third batch Bayesian estimation of mean
$p_1^{(i)}$	0.5986	0.6533	0.7174
$p_2^{(i)}$	0.6674	0.8552	0.9018
$p_3^{(i)}$	0.3603	0.5002	0.7143
$p_4^{(i)}$	0.4998	0.6667	0.8000

Table 5 summarizes the Bayesian posterior means of parameters across each batch. Results indicate that from the initial batch to the third batch, the mean of the key parameter $p_1^{(i)}$ increased from 0.5986, 0.3603, and $p_3^{(i)}$ increased from 0.4998 to 0.7174, 0.7143, and $p_4^{(i)}$ increased from 0.8000, respectively, fully demonstrating the effectiveness of the design optimization.

Taking the third batch as the final evaluation batch, the Bayesian estimation results are shown in Table 6.

Table 6 presents the estimated results of the final functional status probability for the third batch. Among them, the probability of complete upgraded is 0.119, and the probability of significant improvement is 0.179. The combined probability of these two is 29.8%, significantly higher than the level of the initial batch (about 10%). This indicates that the new scheme has significantly enhanced the high-level functional output capability.

To quantitatively evaluate the improvement effect of material performance, an ordered constraint relationship was constructed for verification. The cumulative probabilities of causing mild and above, moderate and above, severe and above, and complete reconstruction for each batch were calculated, and the results are shown in Table 7.

Table 6. The Bayesian estimation results of the functional status probability for the third batch of materials.

Function status level	Initial batch Bayesian estimation of mean	The second batch Bayesian estimation of mean
No performance improvement	0.283	0.080
Mild performance improvement	0.188	0.067
Moderate performance improvement	0.232	0.067
Significant performance improvement	0.179	0.055
Completely upgraded	0.119	0.046

Table 7. Cumulative function for enhancing probability and sequential constraint verification.

Parameter	Batch 1	Batch 2	Batch 3	Meet incremental demand
Mild and above	0.5986	0.6533	0.7174	√
Moderate and above	0.3345	0.4368	0.5296	√
Severely and above	0.1442	0.2039	0.2982	√
Complete reconstruction	0.0480	0.0718	0.1185	√

In Table 7, all the cumulative functional improvement probabilities show a strict sequential increase trend, fully meeting the preset order constraints, indicating that the improvement in material performance after each batch of tests is indeed effective.

From the results, it can be seen that the method proposed in this paper can determine the probability distribution of the different performance status levels of the materials. Compared with the binomial distribution method based on success or failure, the obtained probabilities are more comprehensive.

The Bayesian Markov chain results for the probability parameters of the multi-batch and multi-stage binomial distribution of material effectiveness compliance were statistically analyzed. The results are shown in Table 8.

Table 8. Bayesian Markov Chain statistical results of multi-batch growth experiments.

Batch	Parameters	Sample	Mean	Variance	Monte Carlo error	2.5% CI Lower	Median	97.5% CI Upper
Batch1	$p_1^{(1)}$	4000	0.599	0.011	1.690E-03	0.378	0.604	0.786
	$p_2^{(1)}$	4000	0.559	0.015	1.917E-03	0.319	0.560	0.782
	$p_3^{(1)}$	4000	0.431	0.016	2.003E-03	0.195	0.428	0.685
	$p_4^{(1)}$	4000	0.332	0.017	2.058E-03	0.109	0.323	0.608
Batch2	$p_1^{(2)}$	4000	0.653	0.008	1.440E-03	0.468	0.656	0.817
	$p_2^{(2)}$	4000	0.669	0.011	1.624E-03	0.457	0.674	0.848
	$p_3^{(2)}$	4000	0.467	0.012	1.747E-03	0.260	0.465	0.678
	$p_4^{(2)}$	4000	0.352	0.015	1.939E-03	0.129	0.348	0.608
Batch3	$p_1^{(3)}$	4000	0.717	0.006	1.262E-03	0.547	0.722	0.858
	$p_2^{(3)}$	4000	0.738	0.008	1.403E-03	0.543	0.745	0.890
	$p_3^{(3)}$	4000	0.563	0.010	1.611E-03	0.363	0.565	0.761
	$p_4^{(3)}$	4000	0.397	0.013	1.804E-03	0.188	0.395	0.628

Based on the Bayesian estimation results from the multi-batch growth experiments in Table 8, it can be observed that the mean values of the probability parameters $p_j^{(i)}$ for material performance compliance across batches exhibit an increasing trend with each subsequent batch. This indicates that design optimization after each batch of testing effectively enhances material performance across all stages. Concurrently, both parameter variance and Monte Carlo error decrease as the number of batches increases, demonstrating reduced uncertainty in the posterior distribution and improved precision of the estimation results.

In the first batch, the mean of parameter $p_1^{(1)}$ was 0.5986 with a variance of 0.0114; by the third batch, its mean had increased to 0.7174 while its variance significantly decreased to 0.0064, clearly reflecting the steady improvement in material performance. Similarly, the mean of parameter p_2 increased from 0.5588 to 0.7381, while its variance decreased from 0.0147 to 0.0079. This trend was observed across all four parameters, validating the model's effectiveness.

Table 9. Bayesian Markov Chain statistical results of the probability parameter for single batch ammunition damage.

Parameters	Sample	Mean	Variance	Monte Carlo error	2.5% CI Lower	Median	97.5% CI Upper
p_1	4000	0.682	0.007	1.368E-03	0.611	0.812	0.938
p_2	4000	0.701	0.009	1.522E-03	0.560	0.784	0.928
p_3	4000	0.521	0.012	1.763E-03	0.413	0.654	0.849
p_4	4000	0.352	0.016	1.979E-03	0.230	0.470	0.711

The Bayesian estimation results for the probability parameter of efficacy attainment in the final batch incorporate both historical test data from preceding batches and prior simulation-based information from the current batch, yielding a more comprehensive analysis. Therefore, the Bayesian estimation results for the probability parameter of efficacy attainment in the final batch are selected as the final evaluation value for the multi-batch escalation trial.

After determining the final Bayesian estimates for the probability parameters of achieving performance targets at each stage, substituting these values into Eq. (31) yields the Bayesian estimates for the probability of achieving each performance state level for the material based on the multi-batch growth test. Taking the third batch as an example, the mean probabilities of performance states at each level derived from the inversion are as follows: no performance improvement at 0.283, slight improvement at 0.188, moderate improvement at 0.232, significant improvement at 0.179, and complete reconstruction at 0.119.

To validate the effectiveness of the multi-batch fusion method proposed in this paper, single-batch analysis was also conducted. As shown in Table 9, for the third batch of data, the mean parameter values obtained from single-batch analysis were 0.682, 0.701, 0.521, and 0.352, while those from multi-batch analysis were 0.717, 0.738, 0.563, and 0.397. It is evident that the multi-batch analysis results better align with the expected trend of "performance improvement across batches," with all parameter means exceeding those from single-batch analysis. This indicates that incorporating historical information more accurately reflects the improvement trend in material performance.

Select the iteration trace of parameter $p_2^{(2)}$、$p_2^{(3)}$ from the multi-batch growth experiment and compare it with the iteration trace of parameter p_2 from the single-batch experiment. Observe the iterations from the 2500th to the 2550th iteration. The results are shown in Figure 7.

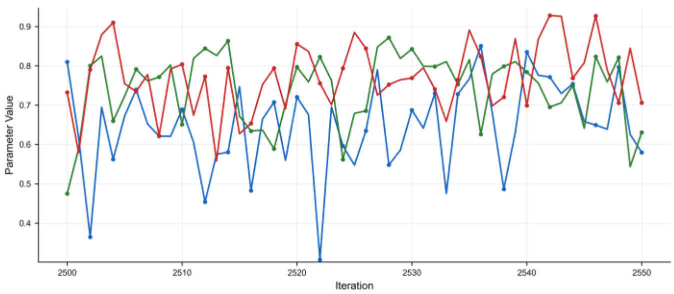


Figure 7. Comparison of iteration traces of parameters $p_2^{(2)}$,$p_2^{(3)}$and p_2.

The scenario where single-batch test parameters are used solely based on information from the third batch of tests fails to align with the assumptions of multi-batch incremental testing. This occurs because no sequential constraints were established, thereby failing to meet the characteristics required for multi-batch incremental testing of material performance.

Simultaneously, both single-batch and multi-batch analyses were conducted on the third batch of data. The Bayesian estimates for the probability parameters of material effectiveness compliance under the two methods are compared as shown in Table 10.

Table 10. Comparison of single and multiple batch Bayesian estimation results.

Method	Parameters	Mean	Standard deviation
Single batch	p_1	0.682	0.086
	p_2	0.701	0.096
	p_3	0.521	0.111
	p_4	0.352	0.125
Multiple batches	$p_1^{(3)}$	0.717	0.080
	$p_2^{(3)}$	0.738	0.089
	$p_2^{(3)}$	0.563	0.102
	$p_4^{(3)}$	0.397	0.114

As shown in Table 10, the conditional probability estimates for each stage of the single-batch analysis, which solely rely on the current batch's experimental data and simulation prior, are generally lower than those of the multi-batch analysis. However, the multi-batch analysis method, by integrating historical test information, strictly satisfies the sequential constraint relationship of "the performance of materials improves progressively" in its estimation results, and the standard deviation of the reconstructed probability (0.114) is significantly lower than that of the single-batch method (0.125), indicating that its estimation is more robust.

In single-batch analysis, parameter estimates may yield unreasonable results due to the lack of integration and constraints from historical information. Furthermore, this approach completely disconnects the relationships between batches, failing to fully leverage information accumulated from prior trials and resulting in inadequate information utilization. Under small-sample conditions, this information waste significantly increases the uncertainty of efficiency estimates and reduces the accuracy of evaluations.

Furthermore, when the total number of batches is large, the performance state reflected by early batches may differ significantly from that of later ones. Directly incorporating these early, weaker-performing data points into the final assessment could dilute the true performance level. Therefore, the weighted fusion method based on system contribution proposed in this paper assigns greater weight to recent test data. This approach more accurately reflects the current true effectiveness of the material, effectively enhancing the precision of performance estimation.

7. Conclusions

This paper proposes a Bayesian dynamic evaluation method based on performance increment constraints to address the dynamic increase in probability of meeting performance standards during multi-batch testing of materials, which arises from iterative optimization of design parameters. By constructing a multi-stage binomial distribution model, the method transforms complex multi-level performance state assessments into a series of sequential conditional probability inferences. It scientifically integrates historical test data with simulation results using system contribution measures, effectively enhancing the information content of small-sample tests. Simultaneously, considering the inherent patterns of multi-batch testing for material performance, the Bayesian inference process incorporates sequential constraints for performance incrementality, efficiently solved via the MCMC method. Research findings demonstrate that this approach more accurately and robustly reflects the improvement trends in material performance, providing a scientific, systematic, and universally applicable solution for multi-batch performance evaluation under small-sample conditions.

While this study assumes a monotonic performance improvement trend—typical in iterative R&D—the proposed framework can be generalized to non-sequential or even regressive scenarios by relaxing the ordered constraint in Eq. (30). Moreover, the core idea of multi-stage conditional decomposition combined with system contribution–weighted fusion is domain-agnostic. Potential applications include equipment health degradation monitoring, multi-phase clinical trials, and supply chain quality evolution, where small-sample, multi-source, and staged-state assessment challenges are similarly prevalent.

References

1. Anderson D, Elmore P, Petry F. Evaluation of heterogeneous uncertain information fusion. Journal of Ambient Intelligence and Humanized Computing 2020; 11(2): 799–811. https://doi.org/10.1007/s12652-019-01320-3.
2. Ding Y J, Wang Z C, Chen G D, Ren W X, Xin Y. Markov Chain Monte Carlo-Based Bayesian Method for Nonlinear Stochastic Model Updating. Journal of Sound and Vibration 2022; 520: 116595. https://doi.org/10.1016/j.jsv.2021.116595.
3. Liu H, Shi X. Damage Effectiveness Calculation of Hitting Targets with Ammunition Based on Bayesian Multinomial Distribution. Symmetry 2022; 14(5): 892.https://doi.org/10.3390/sym14050892.
4. Cinquini T L, Dagan O, Ahmed N R. Nonlinear heterogeneous Bayesian decentralized data fusion. In Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023; 9262–9268. https://doi.org/10.1109/IROS55552.2023.10342177.
5. Li S, Li A. Efficient estimation under data fusion. Biometrika 2023; 110(4): 1041–1054. https://doi.org/10.1093/biomet/asad007.
6. Bladon P, Hall R. J. Information fusion using Bayesian multinets. In Multisensor, multisource information fusion: Architectures, algorithms, and applications 2004; 5434: 91–102. SPIE. https://doi.org/10.1117/12.543661.
7. Li K, Shi X M, Li J, Zhao M, Zeng C. Bayesian estimation of ammunition demand based on multinomial distribution. Discrete Dynamics in Nature and Society 2021; 14(2): 56–67. https://doi.org/10.1155/2021/5575335.
8. Papanasios M, McLeay T E, Mahfouf M, Kadirkamanathan V. A Bayesian information fusion approach for end product quality estimation using machine learning and on-machine probing. Journal of Manufacturing Processes 2022; 6(3): 37–54. https://doi.org/10.1016/j.jmapro.2022.01.020.
9. Greenhill S, Rana S, Gupta S, Vellanki P, Venkatesh S. Bayesian optimization for adaptive experimental design: A review. IEEE Access 2020; 8: 13937–13948. https://doi.org/10.1109/ACCESS.2020.2966228.
10. Xiao F. GEJS: A generalized evidential divergence measure for multisource information fusion. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2023; 53(4): 2246–2258. https://doi.org/10.1109/TSMC.2022.3211498.
11. Huang J, Xiao F. Complex belief divergence measures for multisource information fusion. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2025; 55(2): 948–962. https://doi.org/10.1109/TSMC.2024.3490183.
12. Li Y, Jin B. Pairwise fusion approach incorporating prior constraint information. Communications in Mathematics and Statistics 2020; 8(1): 47–62. https://doi.org/10.1007/s40304-018-0168-3.
13. Yang W, Chen B, Yu L. Bayesian-wavelet-based multisource decision fusion. IEEE Transactions on Instrumentation and Measurement

2021; 70: 2511810. https://doi.org/10.1109/TIM.2021.3094829.
14. Hu J, Huang H, Li Y, Gao H. Bayesian prior information fusion for power law process via evidence theory. Communications in Statistics - Theory and Methods 2022; 51(14): 4921–4939. https://doi.org/10.1080/03610926.2020.1828464.
15. Tao Y, Li X, Chen H, Yang J. Solution for the robust estimation of heterogeneous data fusion based on classification estimation. Journal of Surveying Engineering 2024; 150(3): 04024004. https://doi.org/10.1061/JSUED2.SUENG-1492.
16. Li Y, Zhou E, Zhang F. A distributed Bayesian fusion algorithm with uniform consistency. IEEE Transactions on Automatic Control 2024; 69(9): 6176–6182. https://doi.org/10.1109/TAC.2024.3375254.
17. Saucan A A, Elvira V, Varshney P K, Win M. Z. Information fusion via importance sampling. IEEE Transactions on Signal and Information Processing over Networks 2024; 10: 376–389. https://doi.org/10.1109/TSIPN.2023.3299512.
18. Wu P, Imbiriba T, Elvira V, Closas P. Bayesian data fusion with shared priors. IEEE Transactions on Signal Processing 2024; 72: 275–288. https://doi.org/10.1109/TSP.2023.3343564.
19. Traganitis P A, Giannakis G B. Bayesian constrained decision fusion. In Proceedings of the 2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) 2021; 166–170. https://doi.org/10.1109/SPAWC51858.2021.9593097.
20. Fan C, Yan X. Probability distribution and deviation information fusion driven support vector regression model and its application. Mathematical Problems in Engineering 2017; 2017: 9650769. https://doi.org/10.1155/2017/9650769.
21. Tuia D, Volpi M, Moser G. Decision fusion with multiple spatial supports by conditional random fields. IEEE Transactions on Geoscience and Remote Sensing 2018; 56(6): 3277–3289. https://doi.org/10.1109/TGRS.2018.2797316.
22. de Oude P, Pavlin, G, Hood T. A modular approach to adaptive Bayesian information fusion. In Proceedings of the 2007 10th International Conference on Information Fusion (FUSION) 2007; 1056–1063.
23. Glinton R, Giampapa J, Sycara K. A Markov random field model of context for high-level information fusion. In Proceedings of the 2006 9th International Conference on Information Fusion (FUSION) 2006; 1535–1542.
24. Anava Y, Shtok A, Kurland O, Rabinovich, E. A probabilistic fusion framework. In Proceedings of the 2016 ACM Conference on Information and Knowledge Management 2016; 1463–1472. https://doi.org/10.1145/2983323.2983739.
25. Stutz L T, Teixeira J S, Rangel I C S S, Knupp D C, Silva Neto A J. Population-based Markov Chain Monte Carlo Method with Delayed-Acceptance Applied to Structural Damage Identification. Journal of the Brazilian Society of Mechanical Sciences and Engineering 2025; 47: 277. https://doi.org/10.1007/s40430-025-05572-3.
26. Yang K, Bar-Shalom Y, Hunt S. Hybrid probabilistic information matrix fusion. In Signal processing, sensor/information fusion, and target recognition XXX 2021; 11756: 1175603. SPIE. https://doi.org/10.1117/12.2585951.
27. Hu C, Ling F, Shi X, Wang Y. Bayesian inference for zero-inflated negative binomial lindley model of over-dispersed count data with excess zeros. International Journal of Industrial Engineering Computations 2025; 16(4): 1077–1088. https://doi.org/10.5267/j.ijiec.2025.7.001.
28. Hu C, Shi X, Wang Y. Equipment reliability modeling. Eksploatacja i Niezawodnosc 2026; 28(1): 216106. https://doi.org/10.17531/ein/216106.
29. Hu C, Shi X M. Bayesian evaluation of multi-grade damage efficiency of ammunition using multi-stage binomial distribution. International Journal of Industrial Engineering Computations 2026; 17(1): 67–82. https://doi.org/10.5267/j.ijiec.2025.12.002.
30. Hu C, Fang L, Shi X M. Bayesian hierarchical negative binomial-Lindley model for ammunition malfunction analysis: Theory and empirical validation. Communications in Statistics-Theory and Methods 2025; Advance online publication. https://doi.org/10.1080/03610926.2025.2592780.

Nomenclature

Symbol	Description
$p^{(i)} = \left(p_1^{(i)}, p_2^{(i)}, p_3^{(i)}, p_4^{(i)}\right)$	The distribution of specific success probabilities at each stage
$x_k^{(i)}$	The frequency of observed damage at the k -th level
$s_j^{(i)}$	The number of "successes" at stage j in the i -th batch
$n_j^{(i)}$	The effective sample size entering stage j in the i -th batch
$\pi_k^{(i)}$	The absolute probability of achieving exactly damage level k in the i -th batch

Symbol	Description
$\alpha_j^{(i)}, \beta_j^{(i)}$	The hyperparameters of the Beta prior distribution for stage parameter
$\delta = \{\delta_1, \delta_2\}$	The set of prior information sources
$w_{\delta_l}^d, v_{\delta_l}^d$	The membership degree (reliability) and non-membership degree (unreliability) assigned
$\tau_{\delta_l}^d$	Degree of hesitation
$\varepsilon_{\delta_l}^{\tau}$	Integration weights