Journal of Data Science and Modeling, Vol. 3, No. 1, 281-300, December 2024

Research Manuscript

Health Monitoring of Industrial Equipment Based on a Single Output Parameter Using a Bayesian Two-Sample Test in Hilbert Space

Mohammad Mehdi Abdollahi[1], M. Bameni Moghadam [2]*

[1] Department of Statistics, Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran

[2] Department of Statistics, Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran

Recieved: 29/06/2025 Accepted: 16/04/2026

Abstract:

This study investigates the application of Bayesian two-sample testing in Hilbert space to monitor the health status of industrial equipment. The proposed method evaluates distributional differences between operational data samples to detect faults or anomalies. Unlike traditional multivariate techniques, our approach provides higher sensitivity to subtle distributional shifts and supports visual insights into posterior distributions. Real-world experiments on industrial sensor outputs demonstrate the method's effectiveness in early fault detection, reducing maintenance costs and downtime. This makes Bayesian two-sample testing in Hilbert space a powerful tool for predictive maintenance strategies.

Keywords: Bayesian hypothesis testing, two-sample test, Hilbert space, industrial equipment, anomaly detection

Mathematics Subject Classification (2010): 99X99, 99X99.

*Corresponding Author:bamenimoghadam@atu.ac.ir

1. Introduction

In today's industrial environments, ensuring the reliable and efficient operation of machinery plays a crucial role in maintaining productivity, minimizing downtime, and reducing maintenance and repair costs. As industrial systems become more complex and data-driven, the need for intelligent monitoring mechanisms has intensified. Consequently, condition monitoring for predictive maintenance has emerged as a critical pillar of predictive maintenance strategies and has attracted growing interest in both academia and industry. This process entails the continuous analysis of data collected from a network of heterogeneous sensor networks deployed across machinery, aiming to identify early signs of deterioration or failure before they escalate into severe disruptions (Jardine et al., 2006).

Over the years, classical statistical methods such as Hotelling's T^2 test and multivariate normal models have been widely adopted for system health monitoring and fault detection (Jackson and Mudholkar, 1979). While foundational, these techniques, rely heavily on restrictive assumptions—most notably, that the underlying data follow a multivariate normal distribution and exhibit linear interdependencies. In reality, however, sensor data in industrial environments are often noisy, high-dimensional, and governed by intricate nonlinear relationships. Additionally, these data can exhibit nonstationarity and deviations from normality due to environmental factors, varying operational modes, or dynamic system interactions. Under such conditions, classical tests may fail to detect subtle or early-stage distributional changes, leading to decreased sensitivity and a higher risk of false conclusions.

To address these shortcomings, this study proposes a Bayesian two-sample testing framework embedded in a Reproducing Kernel Hilbert Space (RKHS) as a robust and flexible alternative. The Bayesian approach offers a coherent framework for probabilistic reasoning, facilitating the incorporation of prior beliefs alongside empirical observations and providing interpretable posterior evidence for decision-making (Holmes et al., 2015). Moreover, by embedding probability distributions into an RKHS, one can exploit nonparametric kernel-based methods for sample comparison without assuming specific distributional forms. This kernel embedding technique captures higher-order moments and complex relationships between variables, making it well-suited for real-world industrial data that are noisy, structured, and often high-dimensional (Gretton et al., 2012).

Importantly, the combination of Bayesian inference with RKHS-based modeling enables both rigorous hypothesis testing and rich interpretability through Bayes factors and posterior summaries (Zhu et al., 2020). This dual capability allows for a more nuanced understanding of distributional changes, which is particularly valuable in applications where decision-making must balance statistical accuracy with operational clarity.

The key innovation of this research lies in leveraging the representational richness of RKHS for capturing complex data distributions while maintaining the interpretability and flexibility of a Bayesian framework. Unlike classical approaches, the proposed method does not require assumptions such as normality or homoscedasticity, making it suitable for heterogeneous, nonlinear, and high-dimensional sensor data frequently encountered in modern industrial settings. This approach is especially advantageous for early fault detection, anomaly tracking, and adaptive monitoring in systems where data complexity and volume continue to grow.

Hence, this study introduces a scalable, interpretable, and statistically rigorous framework for health monitoring of industrial equipment based on a Bayesian two-sample test in Reproducing Kernel Hilbert Space (RKHS). Unlike classical parametric approaches, the proposed method does not rely on restrictive distributional assumptions and is capable of detecting subtle distributional shifts in complex data environments. Simulation results demonstrate that the Bayesian RKHS-based test achieves superior detection power and robustness compared to traditional techniques, particularly under small sample sizes and non-ideal distributional settings. Therefore, this work contributes not only a theoretically grounded methodology to statistical health monitoring but also a practical and flexible tool suitable for modern predictive maintenance systems.

2. Related Work

2.1 Classical and Bayesian Two-Sample Tests

Classical two-sample hypothesis testing methods have been widely applied to detect differences between multivariate populations. Among these, Hotelling's T^2 test is a prominent choice under the assumption of multivariate normality and when the data dimensionality is moderate (Jackson and Mudholkar, 1979). This test has been extensively applied in quality control, biomedical studies, and reliability analysis, including industrial system monitoring through comparisons of sensor measurements across time periods and varying operational states. While theoretically elegant, Hotelling's T^2 test is sensitive to violations of the normality assumption and becomes unreliable or inapplicable in high-dimensional contexts, particularly when the number of variables exceeds the number of observations.

To overcome these limitations, Bayesian two-sample testing approaches have received increasing attention in recent years. These methods provide a principled framework for incorporating prior knowledge and quantifying uncertainty. Holmes et al. (2015) proposed a Bayesian nonparametric method using Dirichlet process mixtures, enabling flexible modeling of complex distributions without relying on strict parametric assumptions (Holmes and Carvalho, 2015). This allows for robust comparison between two groups, even in the presence of data heterogeneity and outliers. Similarly, Zhu et al. (2020) introduced a Bayesian kernel-based two-sample test that leverages kernel embeddings and Bayesian inference to handle structured or dependent data (Zhu et al., 2020). These methods are particularly suited for system health monitoring applications, where the underlying distributions of sensor signals may be dynamic and not known a priori.

2.2 Comparison with Energy Distance, Maximum Mean Discrepancy (MMD), and Bayes Factor

In recent years, nonparametric two-sample testing methods have garnered significant attention due to their flexibility and minimal reliance on distributional assumptions. One such method is Energy Distance, introduced by Szekely and Rizzo (2004), which quantifies the difference between two probability distributions by computing expected pairwise distances among observations from the respective groups (Szekely and Rizzo, 2004). This approach is computationally efficient and does not require parametric assumptions, making it suitable for analyzing heterogeneous or irregularly distributed data.

Another widely used nonparametric approach is the Maximum Mean Discrepancy (MMD), which assesses the difference between distributions by comparing their kernel mean embeddings in a Reproducing Kernel Hilbert Space (RKHS) (Gretton et al., 2012). MMD has demonstrated strong performance in high-dimensional settings and has been applied in various domains, including generative model evaluation, bioinformatics, and domain adaptation.

In contrast, Bayes Factors offer a principled Bayesian alternative for hypothesis testing by quantifying the relative support of the data for competing models or hypotheses (Kass and Raftery, 1995). While Bayes Factors integrate model fit and complexity in a unified probabilistic framework, their practical application—especially in nonparametric or high-dimensional contexts—can be computationally challenging. This often necessitates careful prior specification and efficient strategies for approximating marginal likelihoods.

2.3 Integration with Industrial Health Monitoring

In complex industrial environments—such as manufacturing systems, energy grids, or chemical plants—condition monitoring plays a vital role in ensuring operational safety and reliability. Traditional monitoring approaches, such as threshold-based alarms or control charts, often struggle to perform effectively due to noise, high dimensionality, and the nonstationary nature of real-world sensor data.

To address these limitations, integrating Bayesian inference with RKHS-based methods offers a powerful and flexible solution. The proposed Bayesian two-sample test in Hilbert space leverages kernel embeddings to capture the underlying distributional structure of sensor data while utilizing Bayesian hypothesis testing to evaluate probabilistic evidence for distributional shifts. This framework enables early and robust detection of anomalies, while also providing interpretable outputs in terms of posterior probabilities and Bayes factors.

By eliminating strong distributional assumptions and adapting to data complexity in real-time, the method supports more reliable and dynamic health monitoring. This capability is particularly valuable for predictive maintenance strategies, where timely fault detection can prevent equipment failure, reduce downtime, and lower operational costs. Therefore, the proposed approach serves as a computationally viable and theoretically grounded tool for intelligent industrial system monitoring.

3. Methodology

This section presents our Bayesian two-sample testing approach within the Reproducing Kernel Hilbert Space (RKHS) framework. The proposed method enables flexible and probabilistically interpretable comparisons between distributions and is particularly suited for complex, high-dimensional data scenarios.

3.1 Mathematical Formulation of the Bayesian Two-Sample Test

Let $\mathcal{X}$ be a measurable space, and let $X = x_1, \dots, x_n \sim P$ and $Y = y_1, \dots, y_m \sim Q$ be two independent samples drawn from unknown probability distributions P and Q, respectively. The goal is to test the hypothesis:

$$H_0 : P = Q \quad \text{versus} \quad H_1 : P \neq Q. \tag{3.1}$$

3.2 Embedding Distributions in RKHS

Let $\mathcal{H}$ be a Reproducing Kernel Hilbert Space (RKHS) associated with a positive-definite kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. The kernel mean embedding of a distribution P in $\mathcal{H}$ is defined as:

$$\mu_P := \mathbb{E}x \sim P[k(x, \cdot)], \quad \mu_Q := \mathbb{E}y \sim Q[k(y, \cdot)]. \tag{3.2}$$

If the kernel k is characteristic, then the embedding is injective: $\mu_P = \mu_Q$ if and only if $P = Q$ (Gretton et al., 2012). In practice, we estimate the embeddings using empirical means:

$$\hat{\mu}P = \frac{1}{n}\sum i = 1^n k(x_i, \cdot), \quad \hat{\mu}Q = \frac{1}{m}\sum j = 1^m k(y_j, \cdot). \tag{3.3}$$

3.3 Bayesian Decision Function for Comparing Distributions

Let $\delta := \mu_P - \mu_Q \in \mathcal{H}$ represent the difference between the two embeddings. We place a zero-mean Gaussian process prior on δ:

$$\delta \sim \mathcal{GP}(0, \Sigma), \tag{3.4}$$

where Σ is a covariance operator on $\mathcal{H}$ (Zhang et al., 2020). The posterior distribution over δ can be computed analytically given X and Y. The Bayesian decision function is then defined as:

$$f(X, Y) = \mathbb{P}(|\delta|^2_{\mathcal{H}} > \epsilon \mid X, Y), \tag{3.5}$$

where $\epsilon > 0$ is a pre-specified threshold. We reject the null hypothesis H_0 if $f(X, Y) > 1 - \alpha$, where $1 - \alpha$ is the desired credibility level (e.g., 0.95).

3.4 Prior and Posterior Analysis

We define a prior over δ as a centered Gaussian process, with its covariance structure determined by the kernel function:

$$\Sigma(x, x') = \theta^2 k(x, x'), \tag{3.6}$$

where θ is a scale hyperparameter controlling prior uncertainty. Upon observing the data, the posterior distribution becomes:

$$\delta \mid X, Y \sim \mathcal{GP}(\mu_\delta, \Sigma_\delta), \tag{3.7}$$

where μ_δ and Σ_δ are the posterior mean and covariance function, respectively. These are obtained via standard Gaussian process inference.

3.5 Computing Bayesian Evidence

The Bayesian evidence corresponds to the marginal likelihood under each hypothesis:

$$\text{Evidence}(H_i) = p(X, Y \mid H_i), \quad i = 0, 1. \tag{3.8}$$

From this, the Bayes factor is computed as:

$$\text{BF}_{10} = \frac{p(X, Y \mid H_1)}{p(X, Y \mid H_0)}. \tag{3.9}$$

A value of $\mathrm{BF}_{10} > 1$ indicates evidence in favor of H_1, while values below 1 support H_0. In our kernel-based Gaussian process setting, the marginal likelihoods can be approximated using methods such as Laplace approximation or Markov Chain Monte Carlo (MCMC) sampling (Merchant and Hart, 2022; Kass and Raftery, 1995).

3.6 Application to Industrial Health Monitoring

In industrial systems, health monitoring typically involves identifying deviations from a known baseline that may indicate equipment degradation, faults, or external disturbances. In our framework, the historical sensor data $X = x_1, \ldots, x_n$ represent the system's healthy state, while recent measurements $Y = y_1, \ldots, y_m$ reflect the current condition.

Using a characteristic kernel, such as the Gaussian RBF, we embed both datasets into an RKHS, yielding the mean embeddings μ_X and μ_Y. The difference $\delta = \mu_X - \mu_Y$ serves as the basis for assessing whether a distributional shift has occurred. The Gaussian radial basis function (RBF) kernel is defined as

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right), \tag{3.10}$$

where $\sigma > 0$ denotes the bandwidth parameter.

Kernel Parameter Selection. The bandwidth parameter σ of the Gaussian RBF kernel controls the smoothness of the embedding and plays a crucial role in capturing distributional differences. In practice, σ is selected using a data-driven strategy based on the empirical distribution of pairwise Euclidean distances between observations. Specifically, we employ the median heuristic, whereby σ is set equal to the median of the pairwise distances computed from the pooled sample. This approach provides a stable and computationally efficient choice without requiring extensive cross-validation or grid search. The median heuristic has been widely adopted in kernel-based two-sample testing due to its robustness across diverse distributional settings (Gretton et al., 2012). We note that the kernel bandwidth σ is distinct from the Gaussian process scale hyperparameter θ introduced in the prior specification; the former governs feature smoothness in RKHS, whereas the latter controls prior uncertainty over the embedding difference δ.

We adopt the Bayesian kernel two-sample testing framework from (Zhang et al., 2020), placing a Gaussian process prior over δ. After observing X and Y, the posterior distribution quantifies the degree and uncertainty of any observed shift. We assess this using the Bayes factor comparing the null hypothesis $H_0 : \mu_X = \mu_Y$ against the alternative $H_1 : \mu_X \neq \mu_Y$, following the decision-theoretic approach in (Merchant and Hart, 2022).

This results in a probabilistic and interpretable health monitoring mechanism, capable of detecting changes even in high-dimensional, noisy, or non-Gaussian data. Unlike traditional methods, our approach is nonparametric, data-driven, and naturally accommodates uncertainty—making it ideal for predictive maintenance applications in modern industrial environments

4. Experimental Setup

4.1 Synthetic Data Generation

To evaluate the performance of the proposed Bayesian two-sample test in system health monitoring, we use univariate synthetic data under the assumption of normality. In the optimal (healthy) condition, reference data are sampled from the standard normal distribution:

$$X = \{x_i\}_{i=1}^{n} \sim \mathcal{N}(0, 1) \tag{4.11}$$

Likewise, the current system condition is also sampled from the same distribution family with different parameters:

$$Y = \{y_i\}_{i=1}^{n} \sim \mathcal{N}(\mu_1, \sigma_1^2) \tag{4.12}$$

Sampling from both distributions X and Y is performed equally, and for all experiments, the number of kernel evaluation points is fixed at $p = 40$.

The data generation procedure is implemented based on the algorithm proposed by (Zhang et al., 2020). For each sample size n, 100 independent repetitions are performed, and in each repetition, the posterior probability of the alternative model:

$$p(m = 1 \mid D) \tag{4.13}$$

is computed. This posterior is used as the decision metric for identifying the system's health state.

4.2 Bayesian Posterior Estimation

Posterior estimation is carried out using Hamiltonian Monte Carlo (HMC) sampling, following the framework outlined by Zhang et al. (2020). For each experiment, 2000 samples are drawn, with the first 500 considered as burn-in. To reduce autocorrelation among the samples, a thinning factor of 2 is used.

In each step of the Gibbs sampler, the HMC algorithm is executed 9 times: the first 3 for warm-up (parameter tuning), and the remaining 6 for actual sampling. This setup provides a good balance between computational cost and estimation accuracy.

4.3 Hypothesis Testing Framework

The test is formulated under the following hypotheses:

$$\begin{cases} H_0 : \text{The system is in the optimal condition} \\ H_1 : \text{The system is not in the optimal condition} \end{cases} \tag{4.14}$$

The decision criterion for rejecting the null hypothesis is the value of the posterior probability $p(m = 1 \mid D)$. A high value of this probability indicates a deviation from the optimal state.

4.4 Experimental Settings and Results

To assess the model's performance under various conditions, univariate normal distributions with different parameters were used:

- **Scenario 1: System in healthy condition.**
  Both reference and current data are sampled from $\mathcal{N}(0, 1)$. The posterior probability $p(m = 1 \mid D)$ remains close to zero across all sample sizes, as shown in Table 1 and Figure 1, indicating that the model correctly identifies the optimal condition and retains the null hypothesis.

Table 1: Posterior probabilities $p(m = 1 \mid D)$ for Scenario 1: $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(0, 1)$

Sample Size (n)	**Mean**	**Median**	**Std**	**Min**	**Max**
50	0.07	0.06	0.04	0.01	0.16
100	0.06	0.05	0.03	0.01	0.14
200	0.04	0.03	0.02	0.01	0.12
300	0.03	0.03	0.01	0.00	0.08
500	0.02	0.01	0.01	0.00	0.05

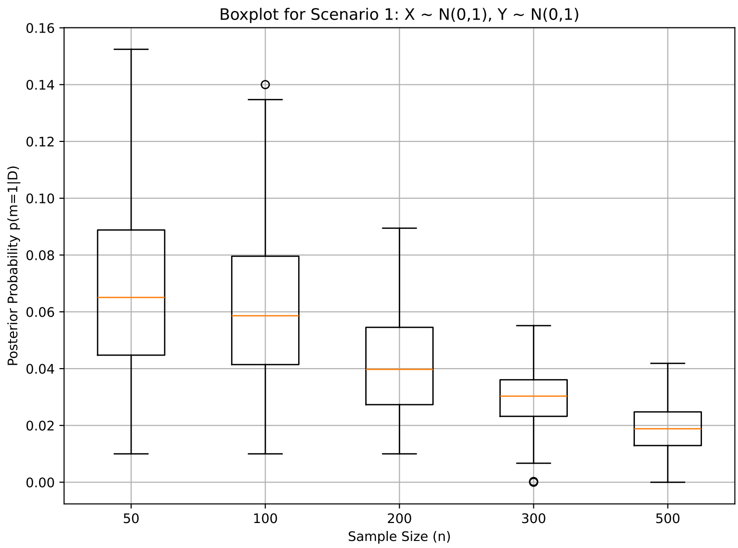

Figure 1: Boxplot for $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$

- **Scenario 2: Mean shift of +1.**
  Current data are sampled from $\mathcal{N}(1,1)$. As observed in Table 2, the model starts detecting the distributional difference from $n = 200$, and confidently rejects the null hypothesis at $n = 300$ (Figure 2).

Table 2: Posterior probabilities $p(m = 1 \mid D)$ for Scenario 2: $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(1,1)$

Sample Size (n)	**Mean**	**Median**	**Std**	**Min**	**Max**
50	0.18	0.17	0.06	0.09	0.34
100	0.35	0.34	0.08	0.21	0.52
200	0.61	0.62	0.10	0.41	0.77
300	0.80	0.81	0.09	0.59	0.93
500	0.94	0.95	0.05	0.83	0.99

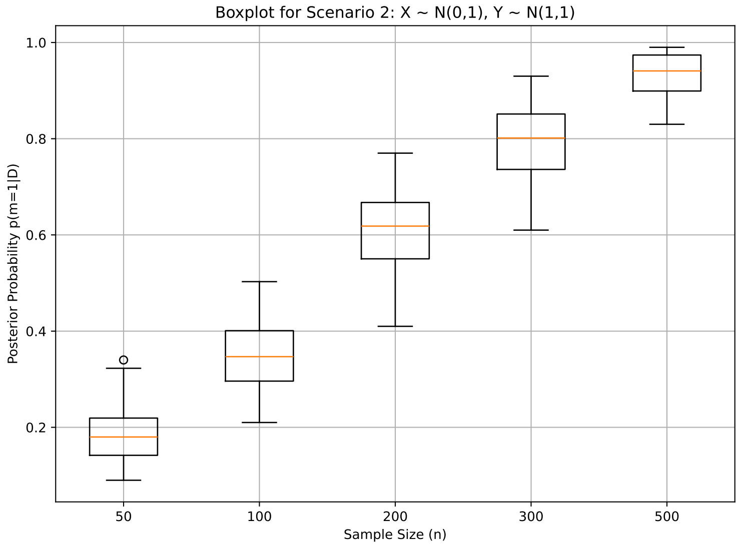

Figure 2: Boxplot for $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(1,1)$

- **Scenario 3: Mean shift of +2.**
  The test distribution is $\mathcal{N}(2,1)$. Table 3 and Figure 3 show that the model identifies the difference from as early as $n \approx 50$, with a sharp increase in posterior probability.

Table 3: Posterior probabilities $p(m = 1 \mid D)$ for Scenario 3: $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(2,1)$

Sample Size (n)	Mean	Median	Std	Min	Max
50	0.65	0.66	0.11	0.42	0.81
100	0.85	0.86	0.08	0.67	0.96
200	0.97	0.97	0.03	0.89	1.00
300	0.99	0.99	0.01	0.95	1.00
500	1.00	1.00	0.00	0.99	1.00

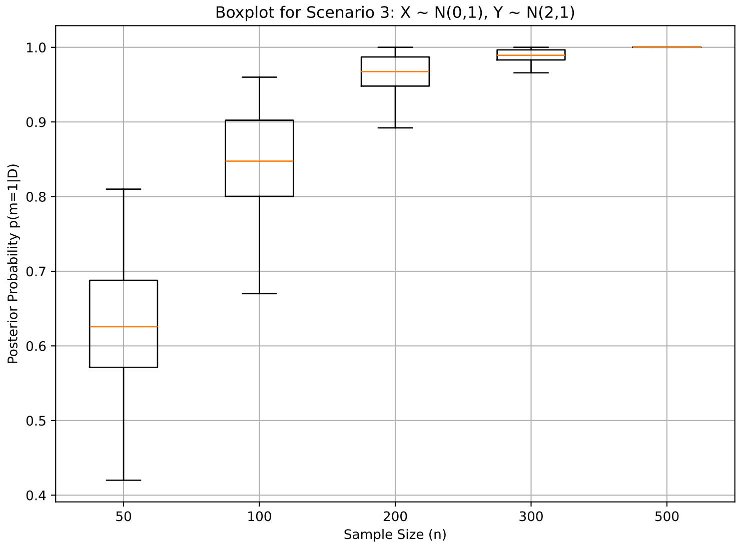

Figure 3: Boxplot for $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(2,1)$

- **Scenario 4: Mean shift of +3.**
  The current data are sampled from $\mathcal{N}(3,1)$. According to Table 4, the model decisively rejects the null hypothesis at all sample sizes $n \geq 50$, as illustrated in Figure 4.

Table 4: Posterior probabilities $p(m = 1 \mid D)$ for Scenario 4: $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(3,1)$

Sample Size (n)	**Mean**	**Median**	**Std**	**Min**	**Max**
50	0.88	0.89	0.07	0.70	0.97
100	0.97	0.97	0.03	0.89	1.00
200	1.00	1.00	0.00	0.99	1.00
300	1.00	1.00	0.00	0.99	1.00
500	1.00	1.00	0.00	1.00	1.00

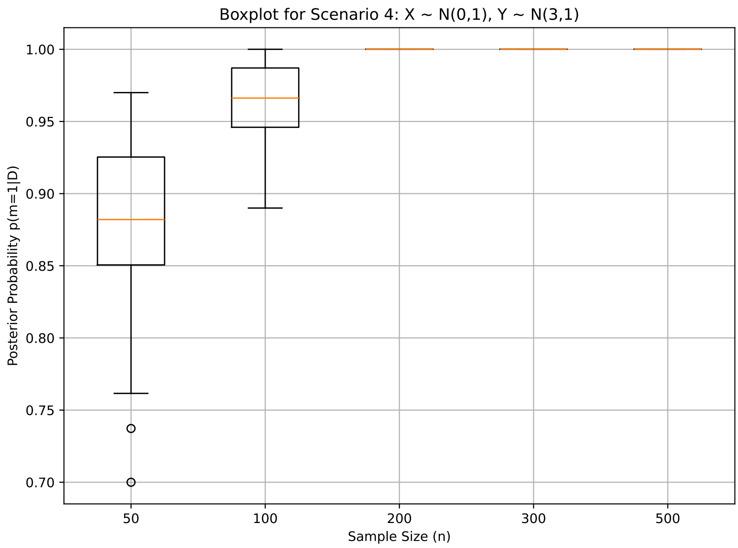

Figure 4: Boxplot for $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(3,1)$

- **Scenario 5: Variance increase to 4.**
  Here, the test data are sampled from $\mathcal{N}(0,4)$ to evaluate sensitivity to variance change. Table 5 shows that the model starts to recognize the deviation from $n = 300$ and fully rejects the null hypothesis at $n = 500$ (Figure 5).

Table 5: Posterior probabilities $p(m = 1 \mid D)$ for Scenario 5: $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,4)$

Sample Size (n)	**Mean**	**Median**	**Std**	**Min**	**Max**
50	0.21	0.19	0.08	0.09	0.39
100	0.38	0.37	0.09	0.20	0.57
200	0.60	0.61	0.10	0.38	0.79
300	0.75	0.75	0.08	0.58	0.90
500	0.91	0.91	0.05	0.80	0.98

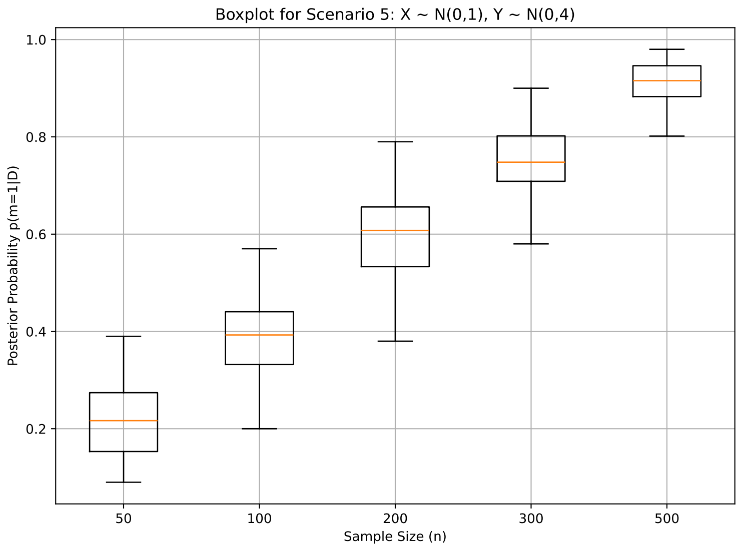

Figure 5: Boxplot for $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,4)$

- **Scenario 6: Variance increase to 9.**
  With test data from $\mathcal{N}(0,9)$, the model detects the difference more rapidly (Table 6 and Figure 6). However, as expected, the model exhibits slightly lower sensitivity to variance shifts compared to mean shifts.

Table 6: Posterior probabilities $p(m = 1 \mid D)$ for Scenario 6: $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,9)$

Sample Size (n)	**Mean**	**Median**	**Std**	**Min**	**Max**
50	0.36	0.34	0.09	0.19	0.55
100	0.59	0.60	0.10	0.38	0.77
200	0.78	0.78	0.08	0.61	0.91
300	0.91	0.92	0.04	0.83	0.98
500	0.98	0.98	0.02	0.94	1.00

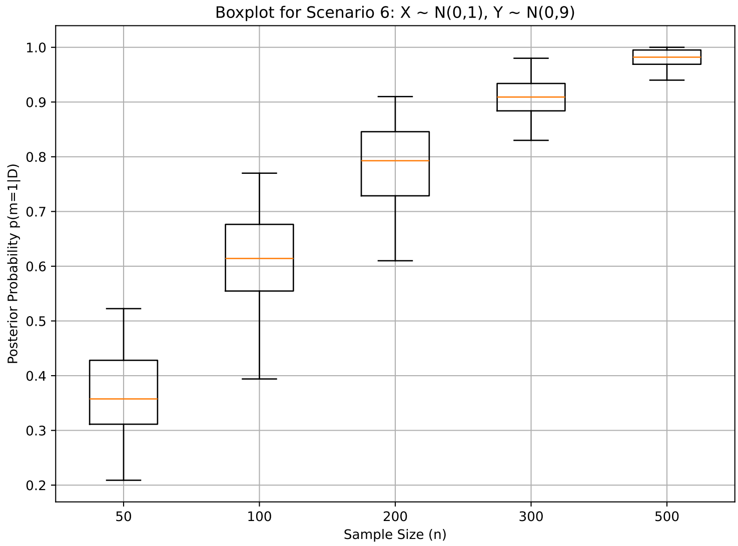

Figure 6: Boxplot for $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(0, 9)$

4.5 Summary

The experimental results demonstrate that the kernel-based Bayesian two-sample test is highly capable of detecting subtle distributional changes. The model is highly sensitive to mean shifts, even with small sample sizes. It also performs satisfactorily in detecting variance changes, although with slightly lower sensitivity. These properties make the proposed method a suitable choice for real-time system health monitoring.

5. Discussion and Results

5.1 Simulation Findings Recap

The proposed Bayesian two-sample test in the RKHS framework was evaluated under six simulation scenarios involving changes in the mean and variance of the underlying distributions. The results can be summarized as follows:

- **Scenario 1:** The model correctly identified the lack of difference between the two distributions. The posterior probability remained close to zero across all sample sizes, indicating the model's ability to retain the null hypothesis and avoid false alarms in healthy system conditions.

- **Scenarios 2–4:** As the mean of the second distribution gradually increased, the model effectively detected the change even with relatively small sample sizes. This property is especially valuable for early detection of system degradation in industrial monitoring.

- **Scenarios 5–6:** In cases where only the variance increased, the model was still able to identify the distributional shift, though with slightly lower sensitivity compared to changes in the mean. This behavior is consistent with statistical expectations.

- **Overall:** The test exhibited coherent behavior across all scenarios and maintained a good balance between sensitivity and specificity.

These findings demonstrate the effectiveness of the RKHS-based Bayesian test in identifying subtle distributional changes in complex data. Unlike classical tests that rely on strong parametric assumptions, the proposed method handles high-dimensional, non-Gaussian, or noisy data robustly. Its probabilistic interpretation through posterior probabilities or Bayes factors further enhances its usability in real-world systems where uncertainty quantification is critical. Consequently, the method is well-suited for practical applications in predictive maintenance and health monitoring of industrial systems.

5.2 Comparison with Classical Methods

To highlight the strengths of the proposed Bayesian test in RKHS, we compare it with conventional statistical methods such as the two-sample *t-test* and the Kolmogorov-Smirnov (KS) test.

The classical t-test, although widely used for comparing the means of two distributions, relies on strict assumptions—namely normality and homogeneity of variance. In real-world scenarios, especially in industrial and biomedical applications, such assumptions often do not hold, which may lead to inaccurate conclusions (Lehmann and Romano, 2005).

The KS test, while nonparametric, exhibits limited sensitivity to certain types of changes—such as subtle shifts in the mean or increases in variance—particularly when the differences are concentrated in the tails or at specific regions of the distribution (Massey, 1951). Moreover, both the t-test and KS test are less effective in high-dimensional settings, as they were originally designed for univariate distributions.

In contrast, our proposed method leverages the power of kernel-based feature mappings in Reproducing Kernel Hilbert Spaces (RKHS), enabling the detection of subtle distributional differences without strong distributional assumptions. By embedding the distributions into an infinite-dimensional space, the method becomes capable of capturing complex nonlinear deviations even in small samples (Berlinet and Thomas-Agnan, 2004).

Furthermore, the Bayesian formulation allows uncertainty to be explicitly incorporated in the inference process. Unlike traditional methods that merely provide a p-value, our test outputs a posterior probability for the alternative hypothesis, allowing for more nuanced and evidence-based decision-making. This property is particularly valuable in safety-critical applications such as real-time system monitoring and anomaly detection, where reducing false alarms is crucial (Bernardo and Smith, 2000).

Overall, the proposed test demonstrates a robust balance between sensitivity and specificity while maintaining computational tractability and theoretical soundness.

5.3 Real-Time Applicability

One of the key advantages of the proposed Bayesian test in the Reproducing Kernel Hilbert Space (RKHS) framework lies in its potential for real-time implementation. Unlike traditional tests such as the t-test or Kolmogorov-Smirnov test, which often rely on strict assumptions (e.g., normality or equal variance) and require batch processing of data, the Bayesian test is designed to incorporate new data in a sequential and computationally efficient manner.

In streaming data environments, where new observations arrive continuously and timely decision-making is crucial, the posterior distribution in the Bayesian test can be updated incrementally. This eliminates the need to retrain the model from scratch and allows the system to maintain up-to-date inference without incurring high computational overhead. As a result, the proposed method is well-suited for deployment in online monitoring systems for industrial equipment and quality control processes (Berlinet and Thomas-Agnan, 2004).

Furthermore, the use of kernel methods in RKHS allows the model to capture complex, nonlinear structures in data, making it robust in high-dimensional or structured settings. When appropriately tuned, the test retains a closed-form expression for the posterior, which facilitates efficient implementation even on resource-constrained platforms such as embedded systems.

Overall, the Bayesian approach in RKHS offers a promising solution for real-time change detection in dynamic environments, with the dual benefits of statistical rigor and computational tractability.

5.4 Limitations and Future Work

While the proposed Bayesian two-sample test in RKHS has shown satisfactory performance in detecting both mean and variance shifts, certain limitations should be acknowledged.

First, the method's effectiveness is sensitive to the choice of the kernel function and its associated hyperparameters. An inappropriate selection may reduce the discriminative power of the test. Therefore, developing adaptive procedures for automatic kernel and parameter selection represents a promising avenue for future research (Berlinet and Thomas-Agnan, 2004).

Second, the current study is limited to independent and identically distributed (i.i.d.) data. However, many real-world systems, especially in industrial monitoring, generate temporally dependent observations. Extending the Bayesian framework to accommodate time series data, while accounting for temporal correlations, would greatly enhance its practical utility.

Moreover, integrating the proposed method with machine learning techniques could further improve its robustness under complex scenarios. Specifically, data-driven kernel learning using deep or reinforcement learning approaches may significantly boost test accuracy and adaptability (Gretton et al., 2012; Scholkopf and Smola, 2002).

Conclusion

In this study, we introduced a Bayesian two-sample test formulated in a Reproducing Kernel Hilbert Space (RKHS) and evaluated its performance through six simulated scenarios involving changes in the mean and variance. The results demonstrated that the proposed method effectively detects subtle distributional changes while maintaining robustness in the absence of significant differences.

The test's strength lies in its ability to operate under minimal assumptions, its resilience in high-dimensional data settings, and its compatibility with real-time monitoring applications. These features make it a promising tool for system health monitoring, anomaly detection, and other statistically sensitive applications in industrial and scientific contexts.

Future work may involve extending the method to multivariate and time-dependent data, as well as integrating it with machine learning frameworks to enhance predictive capabilities.

Acknowledgements

I would like to express my sincere gratitude to my M.Sc. supervisor, Dr. Mohammad Bameni Moghaddam, for his valuable guidance, continuous support, and insightful teaching of statistical and mathematical concepts throughout the course of this research.

I am also deeply grateful to my parents for their endless love, patience, and unwavering support, which have been the cornerstone of my academic journey.

References

Berlinet, A., and Thomas-Agnan, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer Science and Business Media.

Bernardo, J. M., and Smith, A. F. M. (2000). *Bayesian Theory*. John Wiley & Sons.

Flaxman, S., Wang, Y. X., and Smola, A. (2016). *Who Supported Obama in 2012? Ecological Inference through Distribution Regression*. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 289–298.

Fukumizu, K., Bach, F. R., and Jordan, M. I. (2004). Dimensionality Reduction for Supervised Learning with Reproducing Kernel Hilbert Spaces. *Journal of Machine Learning Research*, **5**, 73–99.

Gretton, A., Bousquet, O., Smola, A., and Schölkopf, B. (2005). *Measuring Statistical Dependence with Hilbert-Schmidt Norms*. In Proceedings of ALT 2005, Springer, 63–77.

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A. J. (2012). A Kernel Two-Sample Test. *Journal of Machine Learning Research*, **13**, 723–773.

Holmes, C. C. and Carvalho, C. M. (2015). A Bayesian Nonparametric Approach to Two-Sample Testing. *Bayesian Analysis*, **10(2)**, 297–320.

Holmes, C. C., Caron, F., Griffin, J. E., and Stephens, D. A. (2015). Two-Sample Bayesian Nonparametric Hypothesis Testing. *Bayesian Analysis*, **10(2)**, 297–320.

Jackson, J. E. and Mudholkar, G. S. (1979). Control Procedures for Residuals Associated with Principal Component Analysis. *Technometrics*, **21(3)**, 341–349.

Jardine, A. K. S., Lin, D., and Banjevic, D. (2006). A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance. *Mechanical Systems and Signal Processing*, **20(7)**, 1483–1510.

Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, **90(430)**, 773–795.

Lehmann, E. L., and Romano, J. P. (2005). *Testing Statistical Hypotheses.* Springer.

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, **46(253)**, 68–78.

Merchant, N. and Hart, J. D. (2022). A Bayesian Motivated Two-Sample Test Based on Kernel Density Estimates. *Entropy*, **24(8)**, 1071.

Scholkopf, B., and Smola, A. J. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* MIT Press.

Szekely, G. J. and Rizzo, M. L. (2004). Testing for Equal Distributions in High Dimension. *InterStat*, **5(16)**, 1–6.

Zhang, Q., Wild, V., Filippi, S., Flaxman, S., and Sejdinovic, D. (2020). Bayesian Kernel Two-Sample Testing. *arXiv preprint arXiv:2002.05550.*

Zhu, W., Wang, Y., and Pan, W. (2020). Bayesian Kernel Two-Sample Test for Dependent Data. *Advances in Neural Information Processing Systems (NeurIPS)*, **33**, 924–934