University of Groningen

A new Bayesian multivariate exponentially weighted moving average control chart for phase II monitoring of multivariate multiple linear profiles

Yazdi, Ahmad Ahmadi; Hamadani, Ali Zeinal; Amiri, Amirhossein; Grzegorzczak, Marco

Published in:
Quality and Reliability Engineering International

DOI:
[10.1002/qre.2496](https://doi.org/10.1002/qre.2496)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Yazdi, A. A., Hamadani, A. Z., Amiri, A., & Grzegorzczak, M. (2019). A new Bayesian multivariate exponentially weighted moving average control chart for phase II monitoring of multivariate multiple linear profiles. *Quality and Reliability Engineering International*, 35(7), 2152-2177.
<https://doi.org/10.1002/qre.2496>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

RESEARCH ARTICLE

A new Bayesian multivariate exponentially weighted moving average control chart for phase II monitoring of multivariate multiple linear profiles

Ahmad Ahmadi Yazdi¹ | Ali Zeinal Hamadani¹ | Amirhossein Amiri² | Marco Grzegorzczak³

¹ Department of Industrial and Systems Engineering, Isfahan University of Technology (IUT), Isfahan, Iran

² Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, Iran

³ Department of Mathematics, Bernoulli Institute (BI), Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands

Correspondence

Ahmad Ahmadi Yazdi, Department of Industrial and Systems Engineering, Isfahan University of Technology (IUT), Isfahan 8415683111, Iran.
Email: ahmad_ahmadi_yazdi@yahoo.com

Funding information

Iran National Science Foundation (INSF)

Abstract

Recently, statistical profile monitoring methods have become efficient tools for monitoring the quality of a product (or a production process) using control charts. The key idea is to describe the relationship between a response variable and a set of explanatory variables in the form of a statistical regression model, which called profile. Traditionally, those control charts are constructed with standard “frequentistic” regression models. Recently, it has been proposed to apply Bayesian regression models instead, and it has been empirically demonstrated that Bayesian regression models have the potential to perform significantly better. In this paper, we introduce a novel Bayesian multivariate exponentially weighted moving average control chart for monitoring multivariate multiple linear profiles in phase II. The key idea is to use the data from historical data sets to generate informative prior distributions for the regression models in phase II. The results of our empirical simulation studies show that the Bayesian multivariate multiple linear regression model is superior to its classical “frequentistic” counterpart in terms of the average run length. Our empirical findings are in agreement with findings reported in recently published articles. To shed more light onto the merit of the proposed Bayesian method, we carry out a sensitivity analysis, in which we investigate how the amount of phase I data influences the results. We also demonstrate the applicability and superiority of the proposed Bayesian method by a real-world application.

KEYWORDS

Bayesian modelling, multivariate multiple linear regression, phase II, profile monitoring, statistical process monitoring

1 | INTRODUCTION

In most applications, the quality of a product (or production process) is characterized either by the univariate distribution of a single quality characteristic or by the multivariate distribution of several quality characteristics. But there are also other important applications, where the quality can better be quantified in terms of a functional relationship between a response variable and one or more explanatory variables. The latter is then usually referred to as “profile.”

The functional form of the relationship between the response variable and the explanatory variables is application specific, and statistical regression approaches have been used to model the relationship, eg, simple and multiple linear

regression models, polynomial regression models, and nonlinear regression models. Some of models and applications of profile monitoring have been introduced by many researchers; see, eg, Kang and Albin,¹ Woodall et al,² Wang and Tsung,³ Montgomery,⁴ Woodall,⁵ and Amiri et al.^{6,7}

Profile monitoring usually consists of two phases (phase I and phase II). The major objective of phase I is to identify out-of-control samples and to learn the true relationship between response and explanatory variables. The primary focus of phase II is to quickly detect changes in the process parameters. In this paper, we propose to use the “historical” data from phase I to build informative prior distribution for the Bayesian models employed in phase II.

Most studies that have been published in the literature focus either on phase I or on phase II. Kim et al,⁸ Zhang et al,⁹ Saghaei et al,¹⁰ Mahmoud et al,¹¹ Noorossana and Zerehsaz,¹² De Magalhães and Von Doellinger,¹³ Kamranrad and Amiri,¹⁴ Kazemzadeh et al,¹⁵ Khedmati and Niaki,¹⁶ Chiang et al,¹⁷ and Zhang et al¹⁸ investigated linear profile monitoring in phase II, whereas Mahmoud and Woodall,¹⁹ Mahmoud et al,²⁰ Noghondarian and Ghobadi,²¹ Amiri and Moein,²² Yeh and Zerehsaz,²³ Ghobadi et al,²⁴ Riaz and Touqeer,²⁵ and Taghipour et al²⁶ focused on phase I linear profile monitoring. Methods for monitoring nonlinear profiles have been proposed by Vaghefi et al,²⁷ Ding et al,²⁸ Jin and Shi,²⁹ Jensen and Birch,³⁰ Walker and Wright,³¹ and Steiner et al.³² For a more thorough review about profile monitoring approaches in phases I and II, we refer to the review paper by Woodall⁵ and the textbook edited by Noorossana et al.³³ An overview on recent profile monitoring papers (2008–2018) based on a conceptual classification scheme has been done by Maleki et al.³⁴

Multivariate simple linear (MSL) profiles are based on a linear regression model between one single explanatory variable and a set of response variables. Noorossana et al³⁵ proposed three methods for monitoring MSL profiles in phase II and applied them to the case of a calibration process from NASA Langley Research Center (LaRC). This type of profile has also been studied by Adibi et al,³⁶ Zou et al,³⁷ and Ayoubi et al.³⁸

The main focus of this paper is on multivariate multiple linear (MML) regression profiles. In an MML profile, there is a set of explanatory variables, and the explanatory variables have effects on more than one single response variable. Noorossana et al³⁹ proposed four methods for monitoring the parameters of this type of profiles in phase I. Eyvazian et al⁴⁰ developed four control chart schemes for monitoring the matrix of regression coefficients in phase II, namely, multivariate exponentially weighted moving average (MEWMA), modified MEWMA, MEWMA/ χ^2 , and likelihood ratio test (LRT). Amiri et al⁷ proposed a parameter reduction method for monitoring MML profiles in case the dimension of the vector of regression coefficients is large. Other aspects and applications of MML profiles have been studied by Amiri et al,⁴¹ Ayoubi et al,⁴² and Ghashghaei and Amiri.⁴³ In our study, we focus on the MEWMA control chart in phase II, and we show that its Bayesian variant leads to improved results.

Recently, Bayesian approaches have become very popular methods in statistical process monitoring. The main advantage of the Bayesian models is that researchers can make use of data from phase I to generate informative prior distributions for phase II. This way, preknowledge can be systematically extracted and incorporated into the phase II models, so as to improve their accuracies and performances. In statistical process monitoring applications, there is usually much expert knowledge, and large amounts of historical data are usually available.

That are ideal circumstances for exploiting the full potential of Bayesian statistical models. As modern manufacturing processes are usually associated with high production costs, the early detection of process parameter variations through optimized monitoring methods is of great importance for the manufacturing industry.

In a nutshell, Bayesian models allow scientists to formulate prior beliefs about the values of unknown model parameters in terms of statistical probability distributions. A prior distribution, $P(\theta)$, which reflects the user's prior belief about a set of unknown parameters θ , is updated in light of new data D . This data-based update yields the posterior probability distribution of the parameters, $P(\theta|D)$. The mathematical relationship is as follows:

$$P(\theta|D) = \frac{L(D|\theta)P(\theta)}{P(D)}, \quad (1)$$

where D is the data and θ is the set of unknown parameters. $L(D|\theta)$ is the likelihood function (probability of the data given the parameters), and $P(\theta)$ stands for the prior distribution of the parameters, where the latter should reflect the user's prior knowledge about θ . $P(D)$ in the denominator of Equation 1 is a normalization constant that does not depend on the parameters θ . Bayesian models search for those parameters that fit the new data well (in terms of the likelihood) but at the same time are also likely according to the prior belief (in terms of the prior distribution). A prior is called “conjugate” when the posterior distribution is of the same family of distribution as the prior distribution. For more mathematical details on Bayesian models and Bayesian modelling in general, we refer to the textbook by Hoff.⁴⁴

While the performances of classical “frequentistic” statistical models for profile monitoring have been studied a lot in the literature, Bayesian statistical models have only recently been studied in a few works. Bayesian models have, for example, been studied in Ali and Riaz,⁴⁵ Tan and Shi,⁴⁶ and Marcellus.⁴⁷

Very recently, Abbas et al⁴⁸ have proposed a new Bayesian approach for monitoring profiles on the basis of a simple linear regression model, and they have developed three new phase II univariate Bayesian exponentially weighted moving average (EWMA) control charts. Bayesian methods have also been applied to other types of monitoring schemes, such as cumulative sum (CUSUM) and doubled EWMA (DEWMA) control charts, for simple linear profiles; see, eg, Abbas et al.^{49,50} Recently Riaz et al⁵¹ have developed a Bayesian EWMA control chart using conjugate and nonconjugate priors and considering a loss function for detecting small to moderate shifts in the process parameters.

To the best of our knowledge, so far no Bayesian control chart method for monitoring MML profiles has been proposed in the literature. To fill this gap, we have developed a new Bayesian MEWMA control chart for phase II monitoring of MML profiles. Since Abbas et al⁴⁸ have shown that conjugate priors yield better detection performances for EWMA control charts, we here work with conjugate priors. The Gaussian distribution is here the conjugate prior, and so it appears to be the natural choice. The resulting model is a standard Bayesian model, discussed in many textbooks (see, eg, Hoff⁴⁴). In addition, Abbas et al^{48,49} and Riaz et al⁵¹ also applied Gaussian distribution (conjugate priors) as the prior of the regression coefficients for the simple linear profiles. Their study motivated us to use conjugate priors for multivariate profiles. We would also not see how we could justify to go for a nonconjugate prior instead.

The hyper-parameters of those prior distributions must be selected such that the available prior knowledge is properly incorporated and fully taken into account. For simple linear profiles, there are some methods on how to translate expert opinions into hyper-parameters, eg, the methods by Garthwaite et al⁵² and O'Hagan et al.⁵³ However, we could not find any suitable method for specifying the hyper-parameters of MML regression models in the literature. In particular, there are also cases, in which no reliable expert's knowledge is available. In this paper, we therefore focus on generating prior knowledge from historical data sets (HDSs). With respect to our profile monitoring applications, we propose to extract prior information from phase I data and to incorporate this prior knowledge into the phase II models.

This paper is organized as follows: In Section 2, the Bayesian MML regression model is discussed. The classical estimates of the MML profile parameters are briefly reviewed in Section 3. Section 4 presents the posterior estimates of MML profile parameters using conjugate priors. The proposed Bayesian MEWMA control chart for monitoring regression parameters is presented in Section 5. In Section 6, a comparison study between the performance of the Bayesian and the classical control chart schemes is carried out. Section 7 presents a detailed discussion on the Bayesian control chart. In Section 8, we compare MEWMA and Bayesian MEWMA methods on a real-world example. We summarize and conclude in Section 9.

2 | THE MULTIVARIATE MULTIPLE LINEAR REGRESSION MODEL

Suppose that we have data sets (random samples) $k = 1, 2, \dots$ and that each data set k consists of n observations of the form $(x_{1i}, x_{2i}, \dots, x_{qi}, y_{1ik}, y_{2ik}, \dots, y_{pik}), i = 1, 2, \dots, n$, where q and p denote the number of explanatory and response variables, respectively. Further assume that under in-control condition, the relationship between responses and explanatory variables can be described by an MML regression model:

$$\mathbf{Y}_k = \mathbf{X}\mathbf{B} + \mathbf{E}_k, \quad (2)$$

or equivalently

$$\begin{bmatrix} y_{11k} & y_{12k} & \dots & y_{1pk} \\ y_{21k} & y_{22k} & \dots & y_{2pk} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1k} & y_{n2k} & \dots & y_{npk} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1q} \\ 1 & x_{21} & \dots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nq} \end{bmatrix} \begin{bmatrix} \beta_{01} & \beta_{02} & \dots & \beta_{0p} \\ \beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \dots & \beta_{qp} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11k} & \varepsilon_{12k} & \dots & \varepsilon_{1pk} \\ \varepsilon_{21k} & \varepsilon_{22k} & \dots & \varepsilon_{2pk} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1k} & \varepsilon_{n2k} & \dots & \varepsilon_{npk} \end{bmatrix}, \quad (3)$$

where $\mathbf{Y}_k$ is an $n \times p$ matrix of response variables for the k th sample, $\mathbf{X}$ is an $n \times (q+1)$ design matrix of the values of the explanatory variables, which is considered to be the same for each data set k , $\mathbf{B}$ is the matrix of the regression coefficients, and $\mathbf{E}_k$ is an $n \times p$ matrix of error terms (noise variables). It is assumed that the error vector for each observation follows a p -variate Gaussian distribution: $\varepsilon_{ik} = (\varepsilon_{i1k}, \varepsilon_{i2k}, \dots, \varepsilon_{ipk}) \sim N(\mathbf{0}, \mathbf{\Sigma}_{\varepsilon k})$, (4)

where $\Sigma_{\epsilon k}$ denotes a $p \times p$ covariance matrix:

$$\Sigma_{\epsilon k} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}, \quad (5)$$

The element σ_{hj} is the covariance between the errors of the h th and the j th noise variable (and response variable).

3 | CLASSICAL ESTIMATES OF MML PROFILE PARAMETERS

The ordinary least squares (OLS) estimator of the matrix $\mathbf{B}$ for the k th sample is⁵⁴

$$\hat{\mathbf{B}}_k = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_k. \quad (6)$$

And the covariance matrix ($\Sigma_{\epsilon k}$) of the k th sample can be estimated by⁵⁴

$$\hat{\Sigma}_{\epsilon k} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{y}_{ik} - \bar{\mathbf{y}}_k)(\mathbf{y}_{ik} - \bar{\mathbf{y}}_k)^T, \quad (7)$$

where $\mathbf{y}_{ik}$ is the i th vector of response observations of the k th sample and $\bar{\mathbf{y}}_k$ is the sample mean $\bar{\mathbf{y}}_k = \frac{1}{n} \sum_{i=1}^n \mathbf{y}_{ik}$. Other estimators for $\Sigma_{\epsilon k}$ have been proposed in the literature; see, eg, Holmes and Mergen,⁵⁵ Williams et al,⁵⁶ and Rousseeuw.⁵⁷

4 | THE BAYESIAN MULTIVARIATE MULTIPLE LINEAR REGRESSION MODEL

From Equations 2 and 4, it follows that the likelihood of the MML model fulfills

$$p(\mathbf{E}_k | \Sigma_{\epsilon k}) \propto |\Sigma_{\epsilon k}|^{-\frac{n}{2}} \exp \left(\text{tr} \left(-\frac{1}{2} \mathbf{E}_k^T \mathbf{E}_k \Sigma_{\epsilon k}^{-1} \right) \right), \quad (8)$$

$$p(\mathbf{Y}_k | \mathbf{X}, \mathbf{B}, \Sigma_{\epsilon k}) \propto |\Sigma_{\epsilon k}|^{-\frac{n}{2}} \exp \left(\text{tr} \left(-\frac{1}{2} (\mathbf{Y}_k - \mathbf{X}\mathbf{B})^T (\mathbf{Y}_k - \mathbf{X}\mathbf{B}) \Sigma_{\epsilon k}^{-1} \right) \right), \quad (9)$$

where $\text{tr}(\cdot)$ denotes the trace operator. The sum of squares can be decomposed:

$$p(\mathbf{Y}_k | \mathbf{X}, \mathbf{B}, \Sigma_{\epsilon k}) \propto |\Sigma_{\epsilon k}|^{-\frac{n}{2}} \exp \left(\text{tr} \left(-\frac{1}{2} \left(\mathbf{S} + (\mathbf{B} - \hat{\mathbf{B}}_k)^T \mathbf{X}^T \mathbf{X} (\mathbf{B} - \hat{\mathbf{B}}_k) \right) \Sigma_{\epsilon k}^{-1} \right) \right), \quad (10)$$

where $\mathbf{S} = (\mathbf{Y}_k - \mathbf{X}\hat{\mathbf{B}}_k)^T (\mathbf{Y}_k - \mathbf{X}\hat{\mathbf{B}}_k)$ and $\hat{\mathbf{B}}_k$ is obtained by Equation 6. Equation 10 can be rewritten as

$$p(\mathbf{Y}_k | \mathbf{X}, \mathbf{B}, \Sigma_{\epsilon k}) \propto |\Sigma_{\epsilon k}|^{-\frac{n-(q+1)}{2}} \exp \left(\text{tr} \left(-\frac{1}{2} \mathbf{S} \Sigma_{\epsilon k}^{-1} \right) \right) |\Sigma_{\epsilon k}|^{-\frac{(q+1)}{2}} \times \exp \left(\text{tr} \left(-\frac{1}{2} (\mathbf{B} - \hat{\mathbf{B}}_k)^T \mathbf{X}^T \mathbf{X} (\mathbf{B} - \hat{\mathbf{B}}_k) \Sigma_{\epsilon k}^{-1} \right) \right). \quad (11)$$

The standard conjugate prior selection for the MML profile model is⁵⁸

$$p(\Sigma_{\epsilon k}, \mathbf{B}) = p(\Sigma_{\epsilon k}) p(\mathbf{B} | \Sigma_{\epsilon k}), \quad (12)$$

where

$$p(\Sigma_{\epsilon k}) \sim W^{-1}(\mathbf{V}_0, \nu_0), \quad (13)$$

and

$$p(\mathbf{B}|\Sigma_{\varepsilon k}) \sim MN_{(q+1) \times p}(\mathbf{B}_0, \Lambda_0^{-1}, \Sigma_{\varepsilon k}) \cong N_{((q+1)p) \times 1}(\beta_0, \Sigma_{\varepsilon k} \otimes \Lambda_0^{-1}), \quad (14)$$

where $W^{-1}(\cdot)$ denotes the inverse Wishart distribution, $\mathbf{V}_0$ and ν_0 are the scale matrix and the degrees of freedom of the Wishart distribution, respectively, and $MN(\cdot)$ is used as symbol for the matrix Gaussian distribution. $\mathbf{B}_0$ is the mean of $MN(\cdot)$, β_0 is the mean of the multivariate Gaussian distribution ($N(\cdot)$), Λ_0^{-1} is a symmetric $(q+1) \times (q+1)$ matrix, which we need to compute the covariance matrix of $N(\cdot)$, and $\otimes$ is Kronecker product. We note that we use the subscript “0” to indicate the hyper-parameters of the prior distributions. For notational convenience, the regression coefficient matrix $\mathbf{B}$ has been converted to a vector in Equation 14; symbolically, we have $\beta = \text{vec}(\mathbf{B})$, where

$$\beta = (\beta_{01}, \beta_{11}, \dots, \beta_{q1}, \beta_{02}, \beta_{12}, \dots, \beta_{q2}, \dots, \beta_{0p}, \beta_{1p}, \dots, \beta_{qp})^T. \quad (15)$$

The posterior distribution is proportional to likelihood times the prior distributions. Leaving all factors out that do not depend on the parameters, we obtain the relationship

$$\begin{aligned} p(\beta, \Sigma_{\varepsilon k} | \mathbf{Y}_k, \mathbf{X}) &\propto |\Sigma_{\varepsilon k}|^{-\frac{(\nu_0 + p + 1)}{2}} \exp\left(-\frac{1}{2} \text{tr}(\mathbf{V}_0 \Sigma_{\varepsilon k}^{-1})\right) \\ &\times |\Sigma_{\varepsilon k}|^{-\frac{(q+1)}{2}} \exp\left(-\frac{1}{2} \text{tr}\left((\mathbf{B} - \mathbf{B}_0)^T \Lambda_0 (\mathbf{B} - \mathbf{B}_0) \Sigma_{\varepsilon k}^{-1}\right)\right) \\ &\times |\Sigma_{\varepsilon k}|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left((\mathbf{Y}_k - \mathbf{X}\mathbf{B})^T (\mathbf{Y}_k - \mathbf{X}\mathbf{B}) \Sigma_{\varepsilon k}^{-1}\right)\right). \end{aligned} \quad (16)$$

Equation 16 can be brought into a more common form:

$$\begin{aligned} p(\beta, \Sigma_{\varepsilon k} | \mathbf{Y}_k, \mathbf{X}) &\propto |\Sigma_{\varepsilon k}|^{-\frac{(\nu_0 + p + n + 1)}{2}} \\ &\exp\left(-\frac{1}{2} \text{tr}\left(\left(\mathbf{V}_0 + (\mathbf{Y}_k - \mathbf{X}\mathbf{B}_n)^T (\mathbf{Y}_k - \mathbf{X}\mathbf{B}_n) + (\mathbf{B}_n - \mathbf{B}_0)^T \Lambda_0 (\mathbf{B}_n - \mathbf{B}_0)\right) \Sigma_{\varepsilon k}^{-1}\right)\right) \\ &|\Sigma_{\varepsilon k}|^{-\frac{(q+1)}{2}} \exp\left(-\frac{1}{2} \text{tr}\left((\mathbf{B} - \mathbf{B}_n)^T \Lambda_0 (\mathbf{B} - \mathbf{B}_n) \Sigma_{\varepsilon k}^{-1}\right)\right). \end{aligned} \quad (17)$$

And it follows from the shape of the latter joint density that

$$p(\Sigma_{\varepsilon k} | \mathbf{Y}_k, \mathbf{X}, \mathbf{B}_n) \sim W^{-1}(\mathbf{V}_n, \nu_n), \quad (18)$$

and

$$p(\beta | \mathbf{Y}_k, \mathbf{X}, \Sigma_{\varepsilon k}) \sim MN_{(q+1)p}(\mathbf{B}_n, \Lambda_n^{-1}, \Sigma_{\varepsilon k}) \cong N_{((q+1)p) \times 1}(\beta_n, \Sigma_{\varepsilon k} \otimes \Lambda_n^{-1}). \quad (19)$$

The parameters of the later “full conditional distributions” are given by⁵⁸

$$\mathbf{V}_n = \mathbf{V}_0 + (\mathbf{Y}_k - \mathbf{X}\mathbf{B}_n)^T (\mathbf{Y}_k - \mathbf{X}\mathbf{B}_n) + (\mathbf{B}_n - \mathbf{B}_0)^T \Lambda_0 (\mathbf{B}_n - \mathbf{B}_0), \quad (20)$$

$$\nu_n = \nu_0 + n, \quad (21)$$

$$\mathbf{B}_n = (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} (\mathbf{X}^T \mathbf{Y}_k + \Lambda_0 \mathbf{B}_0), \quad (22)$$

$$\Lambda_n = \mathbf{X}^T \mathbf{X} + \Lambda_0, \quad (23)$$

where $\mathbf{V}_n$ and ν_n are the scale matrix and the degrees of freedom of the Wishart distribution, respectively. $\mathbf{B}_n$ and β_n are the mean of $MN(\cdot)$ and $N(\cdot)$, respectively, and Λ_n^{-1} is a symmetric $(q+1) \times (q+1)$ matrix for computing the covariance

matrix of $N(\cdot)$. We note that we use the subscript “ n ” to indicate the parameters of the posterior distributions, obtained from a sample of size n .

5 | BAYESIAN MEWMA CONTROL CHART STRUCTURE

In this section, we propose an extension of the classical MEWMA control chart proposed by Eyvazian et al.⁴⁰ Since this method is developed for phase II monitoring, it can be assumed that the true target values of the regression coefficients vector (β) and the covariance matrix of error terms (Σ_ϵ) are known from phase I. Assume that the k th sample has been observed from the process. We use the posterior distribution mean of the regression coefficients as estimator for the unknown regression coefficients. For the k th sample, we then have

$$\hat{\mathbf{B}}_k^b = E(\beta | \mathbf{Y}_k, \mathbf{X}, \Sigma_{\epsilon k}) = (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} (\mathbf{X}^T \mathbf{Y}_k + \Lambda_0 \mathbf{B}_0). \quad (24)$$

We can rewrite $\hat{\mathbf{B}}_k^b$ as a $((q+1)p) \times 1$ dimensional random vector, denoted $\hat{\beta}_k^b$:

$$\hat{\beta}_k^b = (\hat{\beta}_{k01}^b, \hat{\beta}_{k11}^b, \dots, \hat{\beta}_{kq1}^b, \hat{\beta}_{k02}^b, \hat{\beta}_{k12}^b, \dots, \hat{\beta}_{kq2}^b, \dots, \hat{\beta}_{kop}^b, \hat{\beta}_{k1p}^b, \dots, \hat{\beta}_{kqp}^b)^T. \quad (25)$$

When the process is in control state, the covariance matrix of $\hat{\beta}_k^b$ is

$$\Sigma_{\hat{\beta}_k^b} = \Sigma_{\epsilon k} \otimes \Delta = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1p} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{p1} & \Sigma_{p2} & \cdots & \Sigma_{pp} \end{bmatrix}. \quad (26)$$

The matrix Σ_{hj} is a $(q+1) \times (q+1)$ symmetric matrix equal to $\sigma_{hj} \Delta$, where $\Delta = (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1}$ (see Appendix A for a proof). In our simulation study, we follow Eyvazian et al.⁴⁰ and assume the covariance matrix Σ_ϵ to be known from phase I, so that it stays fixed and does not have to be resampled from Equation 18. We propose the following Bayesian MEWMA (B_MEWMA) statistic:

$$T_{z_k^b}^2 = (z_k^b)^T \Sigma_{z_k^b}^{-1} z_k^b, \quad (27)$$

where

$$z_k^b = \lambda (\hat{\beta}_k^b - \beta) + (1 - \lambda) z_{k-1}^b, \quad (28)$$

$$\Sigma_{z_k^b} = \frac{\lambda}{2 - \lambda} \Sigma_{\hat{\beta}_k^b}. \quad (29)$$

The smoothing constant (λ) can take values between 0 and 1, and z_0^b is a $((q+1)p) \times 1$ zero vector. The upper control limit (UCL) for the Bayesian MEWMA statistic can be obtained such that a desired in-control average run length (ARL) is achieved by simulation runs, as, for example, explained in Noorossana et al.³⁵ and Eyvazian et al.⁴⁰ The procedural flowchart of the proposed Bayesian MEWMA control chart is shown in Figure 1.

5.1 | Selecting the hyper-parameters

Since the posterior mean depends on the hyper-parameters of the prior distributions, it is crucial to select adequate values for them. Often it is recommended to make use of expert knowledge and to translate this knowledge into informative prior distributions. For example, Abbas et al.⁴⁸ used the elicitation technique proposed by Garthwaite et al.⁵² for specifying the prior hyper-parameters. For many real-world applications, there is no genuine expert's knowledge available or it is very challenging (if not impossible) to translate the expert opinion objectively into an informative prior. In this paper, we therefore follow another approach, which we feel is more generally applicable. We propose to make use

FIGURE 1 Illustrative flowchart for the Bayesian MEWMA (B_MEWMA) control chart, proposed in this paper. The flowchart gives an overview over the individual steps. ARL, average run length; HDS, historical data set (eg, from phase I); MML, multivariate multiple linear regression; RL, run length; UCL, upper control limit. Detailed descriptions of the individual steps are provided in the main text

of HDS instead. For profile monitoring applications, it is realistic to assume that there are data from phase I available and that the phase I data contain valuable information about the model parameters. In our simulation studies, we assume that there are m in-control samples each of size n in phase I, which we merge into one large data set. We then use the Bayesian MML model with noninformative priors, to compute the mean of the posterior distribution of the parameters. More precisely, when analyzing m HDSs with n observations each, we choose the uninformative hyper-parameters: $\mathbf{B}_0 = \mathbf{0}$, $\Lambda_0^{-1} = \mathbf{I}$. In light of the historical data, those parameters are then updated to the posterior pendants, $\mathbf{B}_n$ and $(\Lambda_n^m)^{-1}$, cf Equations 22 and 23. Here, Λ_n^m is the posterior parameter derived from m HDSs with n observations each. Later in phase II, we do not use uninformative prior parameter anymore, but we choose the informative parameters $\mathbf{B}_n$ and Λ_n^m that we have earlier inferred from historical data.

The latter expressions are no longer uninformative but updated in light of HDSs from phase I. So to speak, the updated parameters of the posterior distributions from phase I data become informative prior hyper-parameters in phase II.

To be consistent with the study of Eyvazian et al,⁴⁰ we assume the covariance matrix Σ_ϵ to be known, so that we do not have to specify the hyper-parameters of the inverted Wishart distribution in Equation 13. We note that the posterior

mean, which we use as estimator for the regression coefficients and to compute the new B_MEWMA statistic, is not affected by Σ_ε . Finally, we use the $\mathbf{B}_n, (\Lambda_n^m)^{-1}$ as hyper-parameters for MML regression model in phase II. Figure 2 shows a flowchart illustration of the proposed method.

6 | COMPARISON STUDY (CLASSICAL VS BAYESIAN MEWMA CONTROL CHART)

For a fair comparison between the classical and the proposed new Bayesian MEWMA (B_MEWMA) control chart, we first reuse a multivariate profile model from Eyvazian et al.⁴⁰ We have $p = 2$ response variables and $q = 2$ explanatory variables, and the in-control MML regression model is as follows:

$$\begin{cases} Y_1 = 3 + 2X_1 + X_2 + \varepsilon_1, \\ Y_2 = 2 + X_1 + X_2 + \varepsilon_2. \end{cases} \quad (30)$$

The two explanatory variables X_1 and X_2 have four different value combinations, namely, (2,1), (4,2), (6,3), and (8,2). Note that the number of observations (n) should be larger than the number of response variables (p), $n > p$, in the simulation study. The vector $(\varepsilon_1, \varepsilon_2)$ follows a bivariate Gaussian distribution with zero mean vector and a known covariance matrix of $\Sigma_\varepsilon = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$, where $\sigma_1^2 = \sigma_2^2 = 1$, and the correlation coefficient ρ equals to

FIGURE 2 Flowchart overview on how we propose to generate informative prior distributions. The key idea is to make use of historical data sets (HDSs), eg, from phase I. We analyze the HDS using uninformative priors with set, and we use the resulting posterior pendants as hyper-parameters for phase II. From each regression coefficient vector in phase II, a chart statistic is computed and monitored, so as to check whether the process is in-control or out-of-control

0.9. Note that we also performed simulations for other correlation coefficients ρ (eg, for $\rho = 0.5$ or 0.1), but we always observed exactly the same trends.

We set the UCL to achieve a predetermined in-control ARL, and set the value of the smoothing parameter λ equals to 0.2, where the latter is a very common choice in the quality control literature.

To avoid that our results are influenced and biased by the values of the noise variables, we generate 5000 data sets and report the ARLs, averaged across 5000 data sets. For our simulation study, we need to generate in-control and out-of-control samples. When generating data, we first sampled the rows of the matrix of the error terms from multivariate Gaussian distributions with zero mean vector and known covariance matrix Σ_ϵ . Subsequently, we sampled the individual observations from the true MML regression model. As mentioned in Section 5.1, to be consistent with the study of Eyvazian et al,⁴⁰ we thereby assumed the covariance matrix Σ_ϵ to be known. This data generation approach is typically used in the literature of Bayesian profile monitoring; see, eg, Abbas et al.⁴⁸⁻⁵⁰ The proposed Bayesian framework is not used for data generation. Only after having generated data, we use the Bayesian model for estimating the model parameters.

The UCL of the newly proposed B_MEWMA control chart depends on the size of the HDS (m). Table 1 shows the UCL for different selections of m . The UCL of the classical approach must be set to 17.55 to reach an in-control ARL of approximately 200. It can be seen from Table 1 that an increase of m leads to a decrease of both: the standard deviation (SD) of ARL and the value of UCL. For $m \geq 30$, the UCL of B_MEWMA is smaller than the UCL of classical MEWMA (C_MEWMA). This means that the in-control area of the B_MEWMA control chart becomes narrower as m increases. This is a first indication that the Bayesian MEWMA model has the potential to give more accurate estimates than the classical one.

The advantage of the newly proposed Bayesian control chart depends on the availability of historical data. However, according to the literature, every statistical process monitoring (SPM) method including profile monitoring involves two phases, namely, phase I and phase II. We would therefore argue that historical data (from phase I) can typically be assumed to be available. In our work, we show that those data should not be ignored but can be used to define an improved Bayesian control chart. In the classical approaches (like C_MEWMA), the phase I data are only used for estimating the target values of the process parameters for later use in phase II. We show that the Bayesian framework allows us to use the phase I data also when estimating the profile parameters for a new phase II sample. This does not come with further expenses, as no new in-control samples have to be taken.

We note that C_MEWMA control charts also require phase I data for estimating the target values of the process parameters. Without phase I data, the target values for phase II would not be known.

The out-of-control performance of the B_MEWMA control chart compared with C_MEWMA is evaluated in terms of the out-of-control average run length (ARL₁) criterion under two types of shifts in the regression parameters:

- Shifts in one single parameter, for which we distinguish the following three different shift types: (1a) from β_{01} to $\beta_{01} + \delta_1 \sigma_1$, or (1b) from β_{11} to $\beta_{11} + \phi_1 \sigma_1$ or (1c) from σ_1 to $\gamma_1 \sigma_1$.
- Simultaneous shifts of parameters, for which we distinguish the three different shift types: (2a) from (β_{01}, β_{02}) to $(\beta_{01} + \delta_1 \sigma_1, \beta_{02} + \delta_2 \sigma_2)$, or (2b) from (β_{11}, β_{12}) to $(\beta_{11} + \phi_1 \sigma_1, \beta_{12} + \phi_2 \sigma_2)$ or (2c) from (σ_1, σ_2) to $(\gamma_1 \sigma_1, \gamma_2 \sigma_2)$.

Again, to ensure a fair comparison, the types and magnitudes of the shifts, which are considered in this paper, are exactly the same as those studied by Eyvazian et al.⁴⁰ The simulated ARL₁ values for shifts in one single regression

TABLE 1 The values of the upper control limits (UCLs) and the standard deviations of the average run lengths (ARLs) for the proposed B_MEWMA control chart based on historical data sets (HDSs) of different sizes m based on known parameters

m	UCL	In-Control ARL
15	19.15	200 $\pm$ 3.36
30	17.48	200 $\pm$ 2.97
100	17.27	200 $\pm$ 2.61
500	17.13	200 $\pm$ 2.49
2000	16.94	200 $\pm$ 1.79

Note. Results are averaged across 5000 independent data instantiations.

parameter have been summarized in Tables 2–4. The performance of the B_MEWMA method for small values of m ($m = 15$) is slightly worse than the performance of C_MEWMA, but for $m \geq 30$ B_MEWMA, this trend changes. B_MEWMA then becomes slightly superior to C_MEWMA. The empirical results also show the ARL_1 decreases as the size m of the HDS increases. For large m , the B_MEWMA scheme detects earlier that the parameters have been shifted and the process is no longer in-control. Although the ARL_1 for $m \geq 30$ is consistently smaller for B_MEWMA method, the improvement is relatively small and so not necessarily of great practical relevance.

According to Eyvazian et al,⁴⁰ detecting simultaneous shifts with the C_MEWMA control chart is not as easy as the detection of single parameter shifts. Tables 5–7 show the simulated ARL_1 values for simultaneous shifts, cf scenario (2) above. For lack of space, we here report the results for $m = 15$, $m = 30$, and $m = 500$ only, but we note that we found very similar trends for all other values of m .

It can be seen that the performance of the proposed B_MEWMA method is now significantly better than the performance of the C_MEWMA method for $m \geq 30$. However, similar to the previous results, it is observed that the detection performance of B_MEWMA method is worse than C_MEWMA for small values of m ($m = 15$) under simultaneous shifts. For $m \geq 30$, B_MEWMA detects this type of shift much quicker and thus reaches significantly lower ARL_1 , especially for small to moderate shifts. For example, when $\delta_1 = 0.4$ and $\delta_2 = 0.2$, the ARL_1 value of the proposed method based on 30 samples of HDS is equal to 13.09, while the corresponding value for the classical approach is equal to 22.36. This amount of improvement with respect to the detection time of out-of-control samples can lead to substantial reductions of the production costs for manufacturing companies. Another very interesting finding is the excellent

TABLE 2 The ARL_1 values for shifts from β_{01} to $\beta_{01} + \delta_1 \sigma_1$

δ_1	B_MEWMA					C_MEWMA
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = 2000$	
0.2	20.19	17.31	16.32	16.47	16.15	17.33
0.4	7.17	5.41	5.30	5.21	4.98	5.42
0.6	5.19	3.33	3.31	3.18	3.10	3.36
0.8	2.49	2.41	2.37	2.26	2.16	2.45
1	2.16	2.04	2.00	1.97	1.95	2.06
1.2	2.08	1.85	1.77	1.58	1.53	1.88
1.4	1.73	1.55	1.53	1.47	1.36	1.58
1.6	1.44	1.24	1.22	1.12	1.11	1.24
1.8	1.10	1.06	1.05	1.03	1.02	1.06
2	1.09	1.00	1.00	1.00	1.00	1.01

TABLE 3 The ARL_1 values for shifts from β_{11} to $\beta_{11} + \phi_1 \sigma_1$

ϕ_1	B_MEWMA					C_MEWMA
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = 2000$	
0.025	37.88	35.63	35.23	35.08	34.85	36.67
0.05	10.71	9.56	9.20	8.71	8.52	9.58
0.075	7.96	5.25	5.15	5.03	4.73	5.29
0.1	3.79	3.61	3.54	3.36	3.02	3.64
0.125	3.01	2.85	2.67	2.54	2.53	2.83
0.15	2.40	2.35	2.15	2.11	1.94	2.35
0.175	2.13	2.08	2.04	1.99	1.74	2.10
0.2	2.02	1.93	1.75	1.738	1.50	1.96
0.225	1.83	1.82	1.76	1.72	1.68	1.81
0.25	1.65	1.60	1.57	1.55	1.51	1.63

TABLE 4 The ARL_1 values for shifts from σ_1 to $\gamma_1\sigma_1$

γ_1	B_MEWMA					C_MEWMA
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = 2000$	
1.2	45.63	42.17	41.93	41.23	39.34	42.76
1.4	15.78	14.43	14.27	14.08	13.23	15.03
1.6	7.87	7.47	7.33	6.86	6.26	7.61
1.8	5.39	4.69	4.53	4.33	4.14	4.88
2.0	3.61	3.20	3.03	2.85	2.18	3.54
2.2	3.03	2.63	2.50	2.33	1.86	2.75
2.4	2.33	2.24	2.14	2.02	1.65	2.27
2.6	2.19	1.83	1.81	1.68	1.41	1.94
2.8	1.84	1.63	1.50	1.47	1.34	1.70
3.0	1.65	1.53	1.31	1.26	1.20	1.57

performance of the Bayesian control chart for simultaneous shifts in the SDs. The latter especially applies to small shifts. For example, when $(\gamma_1, \gamma_2) = (1.2, 1.2)$ and 30 samples are available in HDS, the ARL_1 value of the C_MEWMA chart is equal to 34.94, while this value decreases (improves) to 31.17 when the proposed Bayesian control chart (B_MEWMA) is used. It can also be seen from Table 6 that the performance of the classical and Bayesian method both converge to 1 and thus get similar for large shifts in SDs. Moreover, and similar to our earlier results, increasing the value of m consistently leads to better detection performance of the B_MEWMA method. That is, the performance of the proposed B_MEWMA control chart improves as the size of the HDS (from phase I) increases.

In order to assess the effect of the sample size n on the performances of the two control charts, we also perform simulation studies for the sample size $n = 10$. It is obvious that the values of UCL will change when increasing n . The UCL of the C_MEWMA control chart for $n = 10$ is equal to 17.50. In addition, the UCLs of the B_MEWMA control chart for $n = 10$ are summarized in Table 8. Table 8 shows that the UCLs are smaller than those obtained for $n = 4$. However, it can also be seen that increasing n has only a minor effect not on the in-control performance of both control charts. However, we expect that the detection performance of both methods improves when there are more samples in each profile.

We here focus our attention on shifts in one parameters. The results are summarized in Tables 9–11. The types and the sizes of the shifts are the same as those studied earlier for the sample size $n = 4$. From the simulation results, it becomes obvious that the trends do not change; ie, the trends for $n = 10$ do not differ from the trends that we have observed for $n = 4$.

The results in Tables 9–11 show the following: When the value of n is set to 10, the detection performance of both control charts clearly improves for shifts in β_{01} and β_{11} . For example, the ARL_1 value of C_MEWMA method for $\delta_1 = 0.2$ is equal to 17.33 when $n = 4$, while the corresponding value is decreased to 7.66 for $n = 10$. Only for shifts in σ_1 the impact of n on the performance of control charts does not appear to be significant. For simultaneous shifts and $n = 10$, we observed the same trends (empirical results not shown).

The above simulations results, for both control charts (C_MEWMA and B_MEWMA), were obtained under the assumption that the true profile parameters (eg, $\mathbf{B}$ and Σ_ε) are known. This has been a common assumption in most of the related studies; see, for example, Abbas et al,^{48–50} Eyvazian et al,⁴⁰ Kamranrad and Amiri,¹⁴ Kazemzadeh et al,¹⁵ Khedmati and Niaki,¹⁶ Chiang et al,¹⁷ and Zhang et al.¹⁸

Following those earlier works, we obtained only one UCL value for the classical approach (see Table 1). We then considered different values of m to monitor this effect on the detection performance of the Bayesian control chart.

However, a more practical assumption could be that the profile parameters have to be estimated in phase I from m in-control samples. Then, the profile parameters in phase II are no longer equal to their true values but have to be

replaced by those estimates. For example, the true vector β in Equation 28 would have to be replaced by $\hat{\beta} = \frac{\sum_{j=1}^m \hat{\beta}_j}{m}$. In this situation, the UCL values of the classical control chart also depend on the values of m . On the basis of this new assumption, we also run a simulation study. Table 12 shows the UCL values for the B_MEWMA and C_MEWMA

TABLE 5 The ARL_1 values for simultaneous shifts from (β_{01}, β_{02}) to $(\beta_{01} + \delta_1 \sigma_1, \beta_{02} + \delta_2 \sigma_2)$

		$m = 15$				
		δ_2				
δ_1		0.2	0.4	0.6	0.8	1
0.2	C_MEWMA	63.35	22.55	10.59	6.25	4.58
	B_MEWMA	89.94	28.11	13.34	8.93	6.53
0.4	C_MEWMA	22.36	16.83	10.19	6.32	4.52
	B_MEWMA	37.83	19.94	11.66	7.94	5.56
0.6	C_MEWMA	10.19	9.05	8.31	6.23	4.82
	B_MEWMA	13.35	11.80	9.63	7.12	5.33
0.8	C_MEWMA	6.45	6.33	6.28	5.45	4.58
	B_MEWMA	7.17	8.54	7.10	5.95	5.24
1	C_MEWMA	4.53	4.81	4.56	4.50	4.02
	B_MEWMA	7.03	6.67	5.28	4.76	4.18
		$m = 30$				
		δ_2				
δ_1		0.2	0.4	0.6	0.8	1
0.2	C_MEWMA	63.35	22.55	10.59	6.25	4.58
	B_MEWMA	62.27	20.57	9.25	6.24	4.56
0.4	C_MEWMA	22.36	16.83	10.19	6.32	4.52
	B_MEWMA	13.09	12.40	8.58	6.23	4.46
0.6	C_MEWMA	10.19	9.05	8.31	6.23	4.82
	B_MEWMA	8.84	6.28	5.03	4.80	3.85
0.8	C_MEWMA	6.45	6.33	6.28	5.45	4.58
	B_MEWMA	5.00	4.45	4.41	3.22	2.98
1	C_MEWMA	4.53	4.81	4.56	4.50	4.02
	B_MEWMA	3.90	3.18	3.01	2.40	2.39
		$m = 500$				
		δ_2				
δ_1		0.2	0.4	0.6	0.8	1
0.2	C_MEWMA	63.35	22.55	10.59	6.25	4.58
	B_MEWMA	60.15	18.08	8.91	6.19	4.42
0.4	C_MEWMA	22.36	16.83	10.19	6.32	4.52
	B_MEWMA	12.61	11.59	7.57	5.12	4.15
0.6	C_MEWMA	10.19	9.05	8.31	6.23	4.82
	B_MEWMA	7.51	5.0	4.96	4.175	3.19
0.8	C_MEWMA	6.45	6.33	6.28	5.45	4.58
	B_MEWMA	4.75	4.38	4.16	3.06	2.59
1	C_MEWMA	4.53	4.81	4.56	4.50	4.02
	B_MEWMA	3.54	2.83	2.68	2.38	2.11

control charts on the basis of different values of m to achieve the in-control ARL of 200. The UCL values obviously differ from the UCL values reported in Table 1. It can be seen from Table 12 that the UCLs of the Bayesian approach are smaller than those of the classical method for $m \geq 30$. This is because the Bayesian estimator has a lower variance. In addition, it can be seen that the values of UCLs decrease as m increases.

We again compare the out-of-control performance of the two control charts B_MEWMA and C_MEWMA under shifts in single parameter (including shifts in β_{01} , β_{11} , and σ_1) in terms of the ARL_1 criterion. The magnitudes and types

TABLE 6 The ARL_1 values for simultaneous shifts from (β_{11}, β_{12}) to $(\beta_{11} + \phi_1 \sigma_1, \beta_{12} + \phi_2 \sigma_2)$

		$m = 15$				
		ϕ_2				
ϕ_1		0.025	0.05	0.075	0.1	0.125
	C_MEWMA	106.41	45.05	23.81	11.48	7.33
	B_MEWMA	129.29	53.31	27.62	13.55	8.96
0.05	C_MEWMA	47.01	37.42	21.02	12.23	8.36
	B_MEWMA	58.30	42.19	23.54	12.92	8.75
0.075	C_MEWMA	20.57	20.17	16.13	11.52	8.34
	B_MEWMA	23.06	22.42	17.53	12.02	8.59
0.1	C_MEWMA	11.76	12.21	11.48	9.72	7.75
	B_MEWMA	12.98	12.63	12.26	10.14	7.83
0.125	C_MEWMA	7.94	8.83	8.34	7.69	6.63
	B_MEWMA	8.72	8.69	8.42	8.00	6.75
		$m = 30$				
		ϕ_2				
ϕ_1		0.025	0.05	0.075	0.1	0.125
0.025	C_MEWMA	106.41	45.05	23.81	11.48	7.33
	B_MEWMA	102.32	44.37	20.69	11.38	7.15
0.05	C_MEWMA	47.01	37.42	21.02	12.23	8.36
	B_MEWMA	26.67	26.07	16.8	9.20	6.83
0.075	C_MEWMA	20.57	20.17	16.13	11.52	8.34
	B_MEWMA	16.44	11.26	8.68	8.21	6.26
0.1	C_MEWMA	11.76	12.21	11.48	9.72	7.75
	B_MEWMA	8.59	7.41	7.41	5.01	4.65
0.125	C_MEWMA	7.94	8.83	8.34	7.69	6.63
	B_MEWMA	6.26	5.01	4.64	3.59	3.57
		$m = 500$				
		ϕ_2				
ϕ_1		0.025	0.05	0.075	0.1	0.125
	C_MEWMA	106.41	45.05	23.81	11.48	7.33
	B_MEWMA	97.14	40.58	16.14	9.18	6.09
0.05	C_MEWMA	47.01	37.42	21.02	12.23	8.36
	B_MEWMA	24.64	21.52	14.69	7.03	5.06
0.075	C_MEWMA	20.57	20.17	16.13	11.52	8.34
	B_MEWMA	14.64	9.18	7.65	5.05	4.68
0.1	C_MEWMA	11.76	12.21	11.48	9.72	7.75
	B_MEWMA	7.50	6.33	6.242	4.23	3.19
0.125	C_MEWMA	7.94	8.83	8.34	7.69	6.63
	B_MEWMA	5.25	4.78	4.10	3.14	2.43

of the shifts are identical to those that we have used in the earlier simulation studies (see Tables 2–4). The simulated ARL_1 values are summarized in Tables 13–15. The new results are consistent with our pervious simulation results. Consistently, the detection performance for the Bayesian control chart is superior to the detection performance of the classical chart. Particularly for small shifts, the detection performance is substantially improved. Also, it can be seen that the performances of both control charts improve as m increases. This is not surprising, as m is the number of phase I in-control samples. Similar to the earlier results, the classical control chart performs better than the Bayesian method for $m = 15$. For example, according to Table 13, for $\delta_1 = 0.2$ and $m = 15$, the ARL_1 value of the B_MEWMA method is

TABLE 7 The ARL_1 values for simultaneous shifts from (σ_1, σ_2) to $(\gamma_1 \sigma_1, \gamma_2 \sigma_2)$

		$m = 15$				
		γ_2				
γ_1		1.2	1.4	1.6	1.8	2
1.2	C_MEWMA	34.94	9.76	4.38	2.45	1.64
	B_MEWMA	48.17	17.24	6.19	3.84	2.91
1.4	C_MEWMA	1.49	1.32	1.19	1.10	1.06
	B_MEWMA	3.64	2.18	2.02	1.52	1.13
1.6	C_MEWMA	1.04	1.02	1.01	1.01	1.00
	B_MEWMA	1.11	1.08	1.06	1.03	1.01
1.8	C_MEWMA	1.00	1.00	1.00	1.00	1.00
	B_MEWMA	1.09	1.05	1.02	1.01	1.00
2	C_MEWMA	1.00	1.00	1.00	1.00	1.00
	B_MEWMA	1.01	1.00	1.00	1.00	1.00
		$m = 30$				
		γ_2				
γ_1		1.2	1.4	1.6	1.8	2
1.2	C_MEWMA	34.94	9.76	4.38	2.45	1.64
	B_MEWMA	31.17	8.65	3.19	2.27	1.43
1.4	C_MEWMA	1.49	1.32	1.19	1.10	1.06
	B_MEWMA	1.30	1.26	1.12	1.08	1.00
1.6	C_MEWMA	1.04	1.02	1.01	1.01	1.00
	B_MEWMA	1.03	1.02	1.00	1.00	1.00
1.8	C_MEWMA	1.00	1.00	1.00	1.00	1.00
	B_MEWMA	1.00	1.00	1.00	1.00	1.00
2	C_MEWMA	1.00	1.00	1.00	1.00	1.00
	B_MEWMA	1.00	1.00	1.00	1.00	1.00
		$m = 500$				
		γ_2				
γ_1		1.2	1.4	1.6	1.8	2
1.2	C_MEWMA	34.94	9.76	4.38	2.45	1.64
	B_MEWMA	26.73	7.41	2.95	2.17	1.22
1.4	C_MEWMA	1.49	1.32	1.19	1.10	1.06
	B_MEWMA	1.25	1.18	1.07	1.02	1.00
1.6	C_MEWMA	1.04	1.02	1.01	1.01	1.00
	B_MEWMA	1.01	1.00	1.00	1.00	1.00
1.8	C_MEWMA	1.00	1.00	1.00	1.00	1.00
	B_MEWMA	1.00	1.00	1.00	1.00	1.00
2	C_MEWMA	1.00	1.00	1.00	1.00	1.00
	B_MEWMA	1.00	1.00	1.00	1.00	1.00

133.77, while the corresponding value of C_MEWMA is equal to 78.00. However, the Bayesian method becomes the superior method for all higher m values ($m \geq 30$), especially for small shifts. For example, according to Table 14, for $\phi_1 = 0.05$ and $m = 30$, the ARL_1 value of the B_MEWMA method is 12.44, while the corresponding value of C_MEWMA is 14.04. This shows the quicker detection performance of the Bayesian method. For simultaneous shifts, we observed exactly the same trends (empirical results not shown).

TABLE 8 The upper control limits (UCLs) of B_MEWMA control chart when $n = 10$

m	UCL
15	18.09
30	17.40
100	17.24
500	17.11
2000	16.92

TABLE 9 The ARL_1 values for shifts from β_{01} to $\beta_{01} + \delta_1 \sigma_1$ when $n = 10$

δ_1	B_MEWMA					C_MEWMA
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = 2000$	
0.2	8.53	7.58	7.52	7.11	7.02	7.66
0.4	3.29	3.13	3.09	2.86	2.73	3.16
0.6	2.18	2.07	2.00	1.92	1.86	2.11
0.8	1.83	1.70	1.65	1.59	1.49	1.77
1	1.35	1.25	1.21	1.18	1.15	1.27
1.2	1.03	1.00	1.00	1.00	1.00	1.02
1.4	1.00	1.00	1.00	1.00	1.00	1.00
1.6	1.00	1.00	1.00	1.00	1.00	1.00
1.8	1.00	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	1.00	1.00	1.00

TABLE 10 The ARL_1 values for shifts from β_{11} to $\beta_{11} + \phi_1 \sigma_1$ when $n = 10$

ϕ_1	B_MEWMA					C_MEWMA
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = 2000$	
0.025	12.51	11.33	11.17	11.05	10.88	11.39
0.05	4.31	4.12	4.09	4.01	3.84	4.16
0.075	2.72	2.57	2.52	2.45	2.30	2.63
0.1	2.10	2.01	1.96	1.92	2.03	2.06
0.125	1.86	1.80	1.66	1.60	1.56	1.81
0.15	1.50	1.36	1.31	1.27	1.22	1.40
0.175	1.12	1.06	1.02	1.00	1.00	1.09
0.2	1.02	1.00	1.00	1.00	1.00	1.00
0.225	1.00	1.00	1.00	1.00	1.00	1.00
0.25	1.00	1.00	1.00	1.00	1.00	1.00

7 | DISCUSSION

The main motivation for using Bayesian estimators is to obtain more accurate estimates of the model. In this section, we elucidate how the newly proposed Bayesian approach differs from the frequentist's approach.

For $m > 30$, the Bayesian approach yields smaller UCL values than the classical approach. This shows that there is less variance in the Bayesian statistic, proposed in Equation 29. The smaller UCL values can lead to quicker detections of out-of-control samples.

The smaller ARL_1 values of the Bayesian method (see Tables 2–8) indicate that the proposed Bayesian control chart has a better detection performance. We are convinced that the simulated performance metric (ARL_1) is a strong tool to

TABLE 11 The ARL_1 values for shifts from σ_1 to $\gamma_1\sigma_1$ when $n = 10$

γ_1	B_MEWMA					C_MEWMA
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = 2000$	
1.2	51.92	42.27	41.16	40.28	37.19	42.54
1.4	16.31	14.36	14.10	14.02	12.62	14.42
1.6	8.42	7.34	7.27	6.53	6.17	7.58
1.8	5.17	4.55	4.48	4.12	4.10	4.83
2.0	3.79	3.13	3.00	2.77	2.11	3.50
2.2	2.96	2.52	2.42	2.29	1.82	2.71
2.4	2.36	2.17	2.02	1.96	1.60	2.24
2.6	2.00	1.82	1.71	1.57	1.35	1.92
2.8	1.76	1.54	1.45	1.40	1.32	1.70
3.0	1.58	1.50	1.27	1.22	1.16	1.53

TABLE 12 The values of the upper control limits (UCLs) for the Bayesian and classical control chart when the process parameters are estimated on the basis of historical data set of size m

m	UCL_{Bayesian}	$UCL_{\text{Classical}}$
15	29.70	27.95
30	20.65	22.02
100	17.64	17.89
500	17.38	17.76
2000 (∞)	16.94	17.55

TABLE 13 The ARL_1 values for shifts from β_{01} to $\beta_{01} + \delta_1\sigma_1$ when the process parameters are estimated on the basis of historical data set of size m

δ_1	B_MEWMA					C_MEWMA				
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = \infty$	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = \infty$
0.2	133.77	26.10	17.39	16.99	16.15	78.00	31.06	18.07	17.66	17.33
0.4	10.39	6.47	5.50	5.41	4.98	9.06	6.85	5.61	5.57	5.42
0.6	5.02	3.75	3.35	3.31	3.10	4.63	3.89	3.42	3.39	3.36
0.8	3.44	2.70	2.46	2.45	2.16	3.24	2.82	2.48	2.47	2.45
1	2.67	2.17	2.05	2.04	1.95	2.55	2.24	2.06	2.06	2.06
1.2	2.18	1.97	1.86	1.85	1.53	2.11	1.99	1.91	1.89	1.88
1.4	2.00	1.79	1.57	1.56	1.36	1.97	1.83	1.61	1.58	1.58
1.6	1.93	1.46	1.25	1.22	1.11	1.87	1.56	1.26	1.24	1.24
1.8	1.72	1.15	1.05	1.04	1.02	1.59	1.23	1.06	1.05	1.06
2	1.38	1.02	1.00	1.00	1.00	1.27	1.04	1.00	1.00	1.00

empirically show that the Bayesian method is more effective than the classical one. Simulation studies, making use of popular performance metrics, are a widely applied method to cross-compare competing methods. Comparable simulation performance metrics are typically used when one cannot analytically prove or show that a new method performs better than the competing method. Several authors, such as Abbas et al.⁴⁸⁻⁵⁰ and Riaz et al.,⁵¹ have recently shown in

TABLE 14 The ARL_1 values for shifts from β_{11} to $\beta_{11} + \phi_1 \sigma_1$ when the process parameters are estimated on the basis of historical data set of size m

ϕ_1	B_MEWMA					C_MEWMA				
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = \infty$	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = \infty$
0.025	182.42	65.93	37.33	36.24	34.85	146.22	100.09	38.56	37.78	36.67
0.05	16.29	12.44	9.80	9.61	8.52	15.12	14.04	10.10	9.73	9.58
0.075	9.24	6.17	5.32	5.22	4.73	8.61	6.53	5.45	5.34	5.29
0.1	5.60	4.13	3.67	3.65	3.02	5.27	4.35	3.75	3.72	3.64
0.125	4.06	3.17	2.88	2.85	2.53	3.91	3.33	2.92	2.89	2.83
0.15	3.26	2.62	2.40	2.37	1.94	3.16	2.73	2.43	2.39	2.35
0.175	2.76	2.23	2.11	2.10	1.74	2.65	2.32	2.14	2.12	2.10
0.2	2.36	2.03	1.94	1.93	1.50	2.28	2.07	1.96	1.96	1.96
0.225	2.11	1.94	1.82	1.78	1.68	2.07	1.98	1.85	1.83	1.81
0.25	2.01	1.82	1.63	1.60	1.51	2.00	1.88	1.65	1.63	1.63

TABLE 15 The ARL_1 values for shifts from σ_1 to $\gamma_1 \sigma_1$ when the process parameters are estimated on the basis of historical data set of size m

γ_1	B_MEWMA					C_MEWMA				
	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = \infty$	$m = 15$	$m = 30$	$m = 100$	$m = 500$	$m = \infty$
1.2	162.88	80.53	43.76	42.22	39.34	134.53	113.41	47.06	43.51	42.76
1.4	61.22	21.38	14.79	14.76	13.23	53.88	24.83	15.57	15.41	15.03
1.6	20.35	9.83	7.65	7.64	6.26	18.48	11.42	7.86	7.68	7.61
1.8	10.50	6.19	4.95	4.87	4.14	9.49	6.75	5.07	4.95	4.88
2.0	6.84	4.34	3.62	3.49	2.18	6.15	4.58	3.65	3.63	3.54
2.2	4.77	3.23	2.74	2.75	1.86	4.47	3.41	2.81	2.79	2.75
2.4	3.66	2.57	2.29	2.25	1.65	3.49	2.78	2.31	2.27	2.27
2.6	2.96	2.20	1.92	1.88	1.41	2.84	2.36	2.00	1.98	1.94
2.8	2.50	1.87	1.68	1.69	1.34	2.38	1.99	1.75	1.73	1.70
3.0	2.20	1.68	1.57	1.53	1.20	2.09	1.77	1.57	1.57	1.57

related works that Bayesian control charts (eg, for monitoring simple linear profile) have a superior performance. They have shown this empirically by reporting simulation study results in their papers. In those papers, we could not find any theoretical reasoning or explanations where the improved performance mathematically comes from. Although we feel that we cannot provide a proper theoretical explanation, we try to shed some light onto this issue.

Let us assume that m in-control samples are available (=historical data) and that we select the noninformative priors:

$\Lambda_0^{\text{HDS}} = \mathbf{I}$ and $\mathbf{B}_0^{\text{HDS}} = \mathbf{0}$. We build a large design matrix, where the matrix $\mathbf{X}$ appears m times $\begin{bmatrix} \mathbf{X} \\ \vdots \\ \mathbf{X} \end{bmatrix}$ and also the cor-

responding response of the form of $\begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_m \end{bmatrix}$. The resulting posterior parameters (updates of Λ_0^{HDS} and $\mathbf{B}_0^{\text{HDS}}$) are then

$$\Lambda_n^{\text{HDS}} = [\mathbf{X}^T, \dots, \mathbf{X}^T] [\mathbf{X}, \dots, \mathbf{X}]^T + \mathbf{I}, \quad (31)$$

$$\mathbf{B}_n^{\text{HDS}} = \left(\begin{bmatrix} \mathbf{X}^T, \dots, \mathbf{X}^T \end{bmatrix} \begin{bmatrix} \mathbf{X}, \dots, \mathbf{X} \end{bmatrix}^T + \mathbf{I} \right)^{-1} \begin{bmatrix} \mathbf{X}^T, \dots, \mathbf{X}^T \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1^T, \dots, \mathbf{Y}_m^T \end{bmatrix}^T. \quad (32)$$

These equations can be rewritten as follows:

$$\mathbf{\Lambda}_n^{\text{HDS}} = m\mathbf{X}^T \mathbf{X} + \mathbf{I}, \quad (33)$$

$$\mathbf{B}_n^{\text{HDS}} = (m\mathbf{X}^T \mathbf{X} + \mathbf{I})^{-1} [\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m]. \quad (34)$$

Therefore, we obtain

$$\mathbf{\Lambda}_n^{\text{HDS}} \mathbf{B}_n^{\text{HDS}} = [\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m]. \quad (35)$$

Only in phase I, we use noninformative priors for the profile parameters. In phase I, the resulting posterior expectation for regression coefficients is thus very similar to the frequentist's counterpart. In phase I, the Bayesian estimator for the unknown regression coefficients using noninformative priors is given by Equation 34, while the classical estimator is obtained by Equation 6. The major innovative idea is to use the posterior parameters (obtained from phase I) as informative priors when applying the same Bayesian model to a new sample in phase II. The uninformative hyper-parameters $\mathbf{B}_0 = \mathbf{0}$ and $\mathbf{\Lambda}_0 = \mathbf{I}$ are then replaced by the informative hyper-parameters. As proposed, we now use the posterior parameters as prior hyper-parameters for the newly data sample in phase II ($\mathbf{Y}_{\text{new}}$). This yields

$$\begin{aligned} \hat{\boldsymbol{\beta}}^b &= \text{vec}(\hat{\mathbf{B}}^b) = (\mathbf{X}^T \mathbf{X} + \mathbf{\Lambda}_n^{\text{HDS}})^{-1} (\mathbf{X}^T \mathbf{Y}_{\text{new}} + \mathbf{\Lambda}_n^{\text{HDS}} \mathbf{B}_n^{\text{HDS}}) \hat{\boldsymbol{\beta}}^b \\ &= (\mathbf{X}^T \mathbf{X} + m\mathbf{X}^T \mathbf{X} + \mathbf{I})^{-1} (\mathbf{X}^T \mathbf{Y}_{\text{new}} + (\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m)) \hat{\boldsymbol{\beta}}^b \\ &= ((m+1)\mathbf{X}^T \mathbf{X} + \mathbf{I})^{-1} (\mathbf{X}^T \mathbf{Y}_{\text{new}} + (\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m)) \hat{\boldsymbol{\beta}}^b \\ &= ((m+1)\mathbf{X}^T \mathbf{X} + \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y}_{\text{new}} + ((m+1)\mathbf{X}^T \mathbf{X} + \mathbf{I})^{-1} (\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m) \end{aligned} \quad (36)$$

As a result, we have

$$\hat{\boldsymbol{\beta}}^b = \frac{1}{m+1} \left(\mathbf{X}^T \mathbf{X} + \frac{1}{m+1} \mathbf{I} \right)^{-1} \mathbf{X}^T \mathbf{Y}_{\text{new}} + \frac{1}{m+1} \left(\mathbf{X}^T \mathbf{X} + \frac{1}{m+1} \mathbf{I} \right)^{-1} (\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m), \quad (37)$$

and finally,

$$\hat{\boldsymbol{\beta}}^b = \frac{1}{m+1} \left(\mathbf{X}^T \mathbf{X} + \frac{1}{m+1} \mathbf{I} \right)^{-1} \mathbf{X}^T \mathbf{Y}_{\text{new}} + \frac{m}{m+1} \left(\mathbf{X}^T \mathbf{X} + \frac{1}{m+1} \mathbf{I} \right)^{-1} \frac{1}{m} (\mathbf{X}^T \mathbf{Y}_1 + \dots + \mathbf{X}^T \mathbf{Y}_m) \quad (38)$$

From the last line, it can be seen that the Bayesian estimator (posterior expectation) is effectively a mean:

$$\hat{\boldsymbol{\beta}}^b = \frac{1}{m+1} \sum_{i=0}^m \left(\mathbf{X}^T \mathbf{X} + \frac{1}{m+1} \mathbf{I} \right)^{-1} \mathbf{X}^T \mathbf{Y}_i \text{ where } \mathbf{Y}_0 = \mathbf{Y}_{\text{new}} \quad (39)$$

When distinguishing between the historical and the new data, we have a weighted average. The weight $\frac{1}{m+1}$ goes to the new sample ($\mathbf{Y}_{\text{new}}$). The weight $\frac{m}{m+1}$ goes to the (average) historical data, $\mathbf{Y}_1, \dots, \mathbf{Y}_m$. So the newly proposed Bayesian estimator is of the form

$$\hat{\boldsymbol{\beta}}^b = p \hat{\boldsymbol{\beta}}^{\text{HDS}} + (1-p) \hat{\boldsymbol{\beta}}^{\text{new data}}, \quad (40)$$

where $\hat{\boldsymbol{\beta}}^{\text{HDS}}$ denotes the vector of estimated regression coefficients on the basis of HDS data and $\hat{\boldsymbol{\beta}}^{\text{new data}}$ denotes the vector of estimated regression coefficients on the basis of the new sample taken in phase II. We now know that the

mixture weight is $p = \frac{m}{m+1}$. For simplification, we now assume that $\lambda = 1$ in the chart statistic in Equation 28. By plugging Equation 40 into the chart statistic, we get

$$T_{z^b}^2 = \left(\beta - \left(p\hat{\beta}^{\text{HDS}} + (1-p)\hat{\beta}^{\text{new data}} \right) \right)^T \Sigma_{z^b}^{-1} \left(\beta - \left(p\hat{\beta}^{\text{HDS}} + (1-p)\hat{\beta}^{\text{new data}} \right) \right). \quad (41)$$

Multiplying this out gives a long expression with nine terms that can be summarized to

$$\begin{aligned} T_{z^b}^2 = & \beta^T \Sigma_{z^b}^{-1} \beta + p^2 \left(\hat{\beta}^{\text{HDS}} \right)^T \Sigma_{z^b}^{-1} \hat{\beta}^{\text{HDS}} + (1-p)^2 \left(\hat{\beta}^{\text{new data}} \right)^T \Sigma_{z^b}^{-1} \hat{\beta}^{\text{new data}} \\ & + 2p(1-p) \left(\hat{\beta}^{\text{HDS}} \right)^T \Sigma_{z^b}^{-1} \hat{\beta}^{\text{new data}} - 2p \left(\hat{\beta}^{\text{HDS}} \right)^T \Sigma_{z^b}^{-1} \beta - 2(1-p) \hat{\beta}^{\text{new data}} \Sigma_{z^b}^{-1} \beta \end{aligned} \quad (42)$$

It can be seen that Equation 42 possesses cross-product terms that do not appear in the classical frequentist's (non-Bayesian) chart statistic. Without historical data (prior knowledge), one only gets

$$T_z^2 = \left(\beta - \hat{\beta}^{\text{new data}} \right)^T \Sigma_{z^b}^{-1} \left(\beta - \hat{\beta}^{\text{new data}} \right) = \beta^T \Sigma_{z^b}^{-1} \beta + \left(\hat{\beta}^{\text{new data}} \right)^T \Sigma_{z^b}^{-1} \hat{\beta}^{\text{new data}} - 2\hat{\beta}^{\text{new data}} \Sigma_{z^b}^{-1} \beta. \quad (43)$$

When comparing Equation 42 with Equation 43, it becomes obvious that the Bayesian approach yields different cross-product terms and also assigns different weights to them. In particular, there is a cross-product, where the historical (in-control) estimator “meets” the new estimator. And depending on the process situation, the new estimator might either refer to a new in-control sample or refer to a new out-of-control sample. It is hard to show analytically that the difference between Equations 42 and 43 leads to an improvement. We therefore followed Abbas et al.⁴⁸⁻⁵⁰ and Riaz et al.⁵¹ and performed extensive simulation studies to empirically cross-compare the performances of the newly proposed method (B_MEWMA) and the state-of-the art method (C_MEWMA). This way we found empirical evidence that B_MEWMA (making use of Equation 42) detects out-of-control samples significantly earlier than C_MEWMA (which makes use of Equation 43).

8 | A CASE STUDY OF AERODYNAMIC FORCES

We now compare the two models on calibration data from the NASA LaRC. This example was introduced by Parker et al.⁵⁹ and also used by Eyvazian et al.⁴⁰ The goal is to model the relationship between $q = 6$ explanatory variables, including three aerodynamic forces and three moments, namely, the normal force (NF), the axial force (AF), the side force (SF), the pitching moment (PM), the rolling moment (RM), and the yawing moment (YM). The electrical measurements of these $p = 6$ forces and moments are the response variables, which are correlated. As each response variable depends on all six explanatory variables, the relationships between the response and explanatory variables can be described by an MML regression model.

Given the in-control situation, the relationships between the response and explanatory variables are as follows:

$$\begin{aligned} Y_1 &= -0.05 + 10x_1 - 0.01x_2 - 0.03x_3 + 0.26x_4 + 0x_5 + 0.03x_6 + \varepsilon_1, \\ Y_2 &= 0.48 + 0.24x_1 + 21.01x_2 - 0.09x_3 + 0.03x_4 - 0.12x_5 + 0.01x_6 + \varepsilon_2, \\ Y_3 &= 0.37 + 0.09x_1 + 0.01x_2 + 6.81x_3 + 0.04x_4 + 0.02x_5 - 0.03x_6 + \varepsilon_3, \\ Y_4 &= 0.04 + 0x_1 + 0x_2 + 0x_3 + 10.53x_4 - 0.47x_5 + 0.21x_6 + \varepsilon_4, \\ Y_5 &= 0.09 - 0.021x_1 + 0x_2 + 0.01x_3 + 0.02x_4 + 7x_5 - 0.34x_6 + \varepsilon_5, \\ Y_6 &= 0.09 + 0.04x_1 + 0x_2 - 0.01x_3 + 0.18x_4 - 0.34x_5 + 11.46x_6 + \varepsilon_6. \end{aligned} \quad (44)$$

TABLE 16 The values of explanatory variables used for aerodynamic forces calibration data²⁶

Observation	x_1	x_2	x_3	x_4	x_5	x_6
1	0.0	0.0	0.0	0.0	0.0	0.0
2	-68.5	19.2	-106.5	43.5	54.0	26.9
3	-62.1	19.2	-96.5	37.8	-49.7	-39.5
4	-61.7	-21.8	-97.0	39.6	46.9	38.7
5	-68.4	-19.3	-107.4	42.1	-54.2	-27.0
6	68.5	-19.3	106.6	-43.5	-53.8	-26.7
7	61.1	-22.2	94.9	-37.2	47.9	39.5
8	62.1	20.8	97.6	-39.8	-47.4	-38.7
9	0.0	0.0	0.0	0.0	0.0	0.0
10	-68.4	19.3	107.3	42.0	54.5	-27.2
11	-60.5	22.4	95.1	38.8	-48.9	40.3
12	-61.1	-22.3	95.0	37.2	47.6	-39.3
13	-68.5	-19.0	106.5	43.4	-54.2	27.1
14	68.7	-19.1	-107.8	-42.2	-53.4	26.5
15	61.5	-21.6	-96.7	-39.4	47.7	-39.2
16	61.6	22.4	-95.7	-37.5	-46.2	38.5

FIGURE 3 Aerodynamic forces calibration data: C_MEWMA vs the proposed B_MEWMA control chart for shifts from $\beta_{12} = 0.24$ to $\beta_{12} = 0.4$. On average, C_MEWMA needs one more sample to go beyond the upper control limit (dotted horizontal line)

FIGURE 4 Aerodynamic forces calibration data: C_MEWMA vs the proposed B_MEWMA control chart for simultaneous shifts from $\beta_{12} = 0.24$ and $\beta_{13} = 0.09$ to $\beta_{12} = 0.4$ and $\beta_{13} = 0.12$, respectively. Here, C_MEWMA requires on average two more samples, before it crosses the upper control limit (dotted horizontal line)

In addition, the known covariance matrix for the error terms is as follows:

$$\Sigma_{\varepsilon} = \begin{bmatrix} 99 & 14 & 17 & 22 & 18 & 15 \\ 14 & 94 & 20 & 24 & 18 & 15 \\ 17 & 20 & 91 & 27 & 11 & 22 \\ 22 & 24 & 27 & 104 & 20 & 21 \\ 18 & 18 & 11 & 20 & 101 & 19 \\ 15 & 15 & 22 & 21 & 19 & 90 \end{bmatrix}. \quad (45)$$

For this application, there are 16 different possible explanatory variable vectors; see Table 16 for an overview.

We applied both control charts (C_MEWMA and the proposed B_MEWMA chart) to monitor the performance. According to the earlier results in Section 6, we expect that the B_MEWMA control chart performs better than C_MEWMA when an HDS with at least 30 samples is available. To achieve in-control ARL of around 200, we have to set the UCLs of the B_MEWMA and the C_MEWMA control chart to 67.25 and 68, respectively.

We generate 5000 data sets, using the MML model in Equation 44, where each data set includes 30 samples, with the first nine samples being in-control and the remaining 21 samples being generated under sustained shifts. For each data set, we compute the chart statistic, and we compare the two methods in terms of the average chart statistic (averaged across the 5000 individual results). We monitor the chart statistics along the 30 samples.

Two types of shifts were considered for this example: first, a shift in the regression coefficient of the second regression profile from $\beta_{12} = 0.24$ to $\beta_{12} = 0.4$ and second, a simultaneous shift from $\beta_{12} = 0.24$ and $\beta_{13} = 0.09$ to $\beta_{12} = 0.4$ and $\beta_{13} = 0.12$, respectively.

For each sample, we compute the chart statistic, and we plot the control chart statistics (vertical axis) against the sample index (horizontal axis).

Figure 3 shows the performances of the classical (A) and the Bayesian (B) control chart for shifts in one single parameter. It can be seen that the performances of both methods are relatively similar in this case; however, the B_MEWMA control detects the shift at the 11th sample, while classical method detects the shift at the 12th sample, which shows that the Bayesian approach performs again slightly better than the competing approach.

In the case of simultaneous shifts, the superiority of the proposed B_MEWMA method is more substantial. Figure 4 shows that B_MEWMA control chart detects the out-of-control status at sample 12th, while the first signal of C_MEWMA control chart detects it at sample 14th.

9 | CONCLUSIONS AND FURTHER DISCUSSION

In this paper, we have proposed a new MEWMA control chart. The new control chart is based on a Bayesian MML regression model and can be used in phase II. This method is an extension of the MEWMA approach proposed by Eyvazian et al,⁴⁰ and we have referred to it as Bayesian MEWMA (B_MEWMA) control chart. Our work has been motivated by the study from Abbas et al,⁴⁸ in which a Bayesian EWMA approach was developed for monitoring simple linear profiles and turned out to lead to a significant improvement. We followed Abbas et al⁴⁸ and used conjugate priors for the unknown model parameters. Proper selections of the hyper-parameters are very important in Bayesian models to avoid biased posterior results and erroneous conclusions. Since HDSs (eg, from phase I) are usually available, we have proposed to exploit phase I data to determine suitable values for the hyper-parameters of the priors of the Bayesian models that are used in phase II.

Throughout the paper, we have compared the performance of the newly proposed B_MEWMA control chart with the classical MEWMA (C_MEWMA) control chart. For the comparison, we have used the out-of-control ARL criterion. For our simulation studies, we have distinguished two types of sustained shifts: (1) shifts in one single parameter and (2) simultaneous shifts in multiple parameters. According to the literature, the performance of classical MEWMA control chart in detection of simultaneous shifts is not as good as for shifts in one single parameter. Our empirical results in Section 6 have shown that the proposed new B_MEWMA control chart does not only lead to a slightly better performance for single parameter shifts but also lead to a substantial improvement for simultaneous shifts.

We have also applied both approaches (C_MEWMA and the proposed B_MEWMA chart) to a real-world calibration problem from the NASA LaRC (Section 8). The results were consistent with our earlier results, and they again confirm the superiority of the newly proposed B_MEWMA control chart in detecting out-of-control situations.

In this study, we have focused on the MML regression model, the MEWMA control chart, and the average run length (ARL_1) to detect out-of-control states. In our future research, we will study alternative control charts and develop novel Bayesian models for them.

ACKNOWLEDGEMENTS

The authors thank the respectful referees for detailed and thoughtful recommendations that led to significant improvement in the paper. This research was partially supported by a grant from Iran National Science Foundation (INSF).

ORCID

Ahmad Ahmadi Yazdi <https://orcid.org/0000-0002-4274-4804>

Amirhossein Amiri <https://orcid.org/0000-0002-2385-8910>

REFERENCES

1. Kang LA, Albin SL. On-line monitoring when the process yields a linear profile. *J Qual Technol*. 2000;32(4):418-426.
2. Woodall WH, Spitzner DJ, Montgomery DC, Gupta S. Using control charts to monitor process and product quality profiles. *J Qual Technol*. 2004;36(3):309-320.
3. Wang K, Tsung F. Using profile monitoring techniques for a data-rich environment with huge sample size. *Qual Reliab Eng Int*. 2005;21(7):677-688.
4. Montgomery DC. *Introduction to Statistical Quality Control*. Fifth ed. John Wiley and Sons; 2005.
5. Woodall WH. Current research on profile monitoring. *Production*. 2007;17(3):420-425.

6. Amiri A, Jensen WA, Kazemzadeh RB. A case study on monitoring polynomial profiles in the automotive industry. *Qual Reliab Eng Int*. 2010;26(5):509-520.
7. Amiri A, Eyvazian M, Zou C, Noorossana R. A parameters reduction method for monitoring multiple linear regression profiles. *Int J Adv Manuf Technol*. 2012;58(5-8):621-629.
8. Kim K, Mahmoud MA, Woodall WH. On the monitoring of linear profiles. *J Qual Technol*. 2003;35(3):317-328.
9. Zhang J, Li Z, Wang Z. Control chart based on likelihood ratio for monitoring linear profiles. *Comput Stat Data Anal*. 2009;53(4):1440-1448.
10. Saghaei A, Mehrjoo M, Amiri A. A CUSUM-based method for monitoring simple linear profiles. *Int J Adv Manuf Technol*. 2009;45(11):1252-1260.
11. Mahmoud MA, Morgan JP, Woodall WH. The monitoring of simple linear regression profiles with two observations per sample. *J Appl Stat*. 2010;37(8):1249-1263.
12. Noorossana R, Zerehsaz Y. Effect of measurement error on phase II monitoring of simple linear profiles. *Int J Adv Manuf Technol*. 2015;79(9-12):2031-2040.
13. De Magalhães MS, Von Doellinger ROS. Monitoring linear profiles using an adaptive control chart. *Int J Adv Manuf Technol*. 2016;82(5-8):1433-1445.
14. Kamranrad R, Amiri A. Robust Holt-Winter based control chart for monitoring autocorrelated simple linear profiles with contaminated data. *Sci Iran*. 2016;23(3):1345-1354.
15. Kazemzadeh RB, Amiri A, Kouhestani B. Monitoring simple linear profiles using variable sample size schemes. *J Stat Comput Simul*. 2016;86(15):2923-2945.
16. Khedmati M, Niaki STA. Monitoring simple linear profiles in multistage processes by a MaxEWMA control chart. *Comput Ind Eng*. 2016;98:125-143.
17. Chiang JY, Lio Y, Tsai TR. MEWMA control chart and process capability indices for simple linear profiles with within-profile autocorrelation. *Qual Reliab Eng Int*. 2017;33(5):1083-1094.
18. Zhang Y, Shang Y, He Z, Wang Q. CUSUM schemes for monitoring prespecified changes in linear profiles. *Qual Reliab Eng Int*. 2017;33(3):579-594.
19. Mahmoud MA, Woodall WH. Phase I analysis of linear profiles with calibration applications. *Dent Tech*. 2004;46(4):380-391.
20. Mahmoud MA, Parker PA, Woodall WH, Hawkins DM. A change point method for linear profile data. *Qual Reliab Eng Int*. 2007;23(2):247-268.
21. Noghondarian K, Ghobadi S. Developing a univariate approach to phase-I monitoring of fuzzy quality profiles. *Int J Ind Eng Comput*. 2012;3(5):829-842.
22. Amiri A, Moein E. Some notes on diagnostic procedures in simple and multiple linear regression profiles monitoring. *Commun Stat Simul Comput*. 2013;42(5):981-1002.
23. Yeh A, Zerehsaz Y. Phase I control of simple linear profiles with individual observations. *Qual Reliab Eng Int*. 2013;29(6):829-840.
24. Ghobadi S, Noghondarian K, Noorossana R, Mirhosseini SS. Developing a multivariate approach to monitor fuzzy quality profiles. *Qual Quant*. 2014;48(2):817-836.
25. Riaz M, Touqeer F. On the performance of linear profile methodologies under runs rules schemes. *Qual Reliab Eng Int*. 2015;31(8):1473-1482.
26. Taghipour M, Amiri A, Saghaei A. Phase I monitoring of within-profile autocorrelated multivariate linear profiles. *J Eng Res*. 2018;5(4).
27. Vaghefi A, Tajbakhsh SD, Noorossana R. Phase II monitoring of nonlinear profiles. *Commun Stat—Theory and Methods*. 2009;38(11):1834-1851.
28. Ding Y, Zeng L, Zhou S. Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *J Qual Technol*. 2006;38(3):199-216.
29. Jin J, Shi J. Feature-preserving data compression of stamping tonnage information using wavelets. *Dent Tech*. 1999;41(4):327-339.
30. Jensen WA, Birch JB. Profile monitoring via nonlinear mixed models. *J Qual Technol*. 2009;41(1):18-34.
31. Walker E, Wright S. Comparing curves using additive models. *J Qual Technol*. 2002;34(1):118-129.
32. Steiner S, Jensen WA, Grimshaw SD, Espen B. Nonlinear profile monitoring for oven-temperature data. *J Qual Technol*. 2016;48(1):84-97.
33. Noorossana R, Saghaei A, Amiri A. *Statistical Analysis of Profile Monitoring*. 865 New York: John Wiley & Sons; 2011.
34. Maleki MR, Amiri A, Castagliola P. An overview on recent profile monitoring papers (2008–2018) based on conceptual classification scheme. *Comput Ind Eng*. 2018;126:705-728.
35. Noorossana R, Eyvazian M, Vaghefi A. Phase II monitoring of multivariate simple linear profiles. *Comput Ind Eng*. 2010;58(4):563-570.
36. Adibi A, Montgomery DC, Borror CM. A *P*-value approach for phase II monitoring of multivariate profiles. *Int J Qual Eng Technol*. 2014;4(2):133-143.
37. Zou C, Ning X, Tsung F. LASSO-based multivariate linear profile monitoring. *Ann Oper Res*. 2012;192(1):3-19.

38. Ayoubi M, Kazemzadeh R, Noorossana R. Estimating multivariate linear profiles change point with a monotonic change in the mean of response variables. *Int J Adv Manuf Technol*. 2014;75(9-12):1537-1556.
39. Noorossana R, Eyvazian M, Amiri A, Mahmoud MA. Statistical monitoring of multivariate multiple linear regression profiles in phase I with calibration application. *Qual Reliab Eng Int*. 2010;26(3):291-303.
40. Eyvazian M, Noorossana R, Saghaei A, Amiri A. Phase II monitoring of multivariate multiple linear regression profiles. *Qual Reliab Eng Int*. 2011;27(3):281-296.
41. Amiri A, Saghaei A, Mohseni M, Zerehsaz Y. Diagnosis aids in multivariate multiple linear regression profiles monitoring. *Commun Stat-Theory and Methods*. 2014;43(14):3057-3079.
42. Ayoubi M, Kazemzadeh RB, Noorossana R. Change point estimation in the mean of multivariate linear profiles with no change type assumption via dynamic linear model. *Qual Reliab Eng Int*. 2016;32(2):403-433.
43. Ghashghaei R, Amiri A. Sum of squares control charts for monitoring of multivariate multiple linear regression profiles in phase II. *Qual Reliab Eng Int*. 2017;33(4):767-784.
44. Hoff PD. *A First Course in Bayesian Statistical Methods*. Berlin: Springer Science and Media; 2009.
45. Ali S, Riaz M. Cumulative quantity control chart for the mixture of inverse Rayleigh process. *Comput Ind Eng*. 2014;73:11-20.
46. Tan MHY, Shi J. A Bayesian approach for interpreting mean shifts in multivariate quality control. *Dent Tech*. 2012;54(3):294-307.
47. Marcellus RL. Bayesian monitoring to detect a shift in process mean. *Qual Reliab Eng Int*. 2008;24(3):303-313.
48. Abbas T, Qian Z, Ahmad S, Riaz M. On monitoring of linear profiles using Bayesian methods. *Comput Ind Eng*. 2016;94:245-268.
49. Abbas T, Ahmad S, Riaz M, Qian Z. A Bayesian way of monitoring the linear profiles using CUSUM control charts. *Commun Stat Simul Comput*. 2017;48:1-24.
50. Abbas T, Qian Z, Ahmad S, Riaz M. Bayesian monitoring of linear profile monitoring using DEWMA charts. *Qual Reliab Eng Int*. 2017;33(8):1783-1812.
51. Riaz S, Riaz M, Hussain Z, Abbas T. Monitoring the performance of Bayesian EWMA control chart using loss functions. *Comput Ind Eng*. 2017;112:426-436.
52. Garthwaite PH, Al-Awadhi SA, Elfadaly FG, Jenkinson DJ. Prior distribution elicitation for generalized linear and piecewise-linear models. *J Appl Stat*. 2013;40(1):59-75.
53. O'Hagan A, Buck CE, Daneshkhah A, et al. *Uncertain Judgements: Eliciting Experts Probabilities*. New York: John Wiley & Sons; 2006.
54. Rencher AC. *Methods of Multivariate Analysis*. 492 John Wiley & Sons; 2003.
55. Holmes DS, Mergen AE. Improving the performance of the T^2 control chart. *Qual Eng*. 1993;5(4):619-625.
56. Williams JD, Woodall WH, Birch JB. *Phase I Monitoring of Nonlinear Profiles*. Paper presented at: quality and productivity research conference. New York: Yorktown Heights; 2003.
57. Rousseeuw PJ. Least median of squares regression. *J Am Stat Assoc*. 1984;79(388):871-880.
58. Rossi PE, Allenby GM, McCulloch R. *Bayesian Statistics and Marketing*. New York: John Wiley & Sons; 2012.
59. Parker P, Morton M, Draper N, Line W. A single-vector force calibration method featuring the modern design of experiments. Paper presented at: 39th Aerospace sciences meeting and exhibit 2001.

AUTHOR BIOGRAPHIES

Ahmad Ahmadi Yazdi is a PhD student of Industrial Engineering at the Isfahan University of Technology (IUT), Iran. His areas of interest include statistical process control, profile monitoring, and acceptance sampling.

Ali Zeinal Hamadani is a professor of statistics in the Department of Industrial Engineering in Isfahan University of Technology. He was selected as a distinguished teaching professor by Isfahan University of Technology and as a distinguished statistics professor by Iranian Statistical society. He received his BSc in Statistics and Computer Science from the College of Statistics and Information (Tehran) in 1975 and MSc and PhD from Universities of Sussex and Bedford in 1977 and 1980, respectively. His main research interests include applied statistics, statistical quality control, reliability and availability modelling of complex systems, dependency and reliability analysis, regression analysis of experiments (DOE), and data mining.

Amirhossein Amiri is an associate professor at Shahed University in Iran. He holds a BS, MS, and PhD in Industrial Engineering from Khajeh Nasir University of Technology, Iran University of Science and Technology, and Tarbiat Modares University in Iran, respectively. He is now vice chancellor of education in the Faculty of Engineering at Shahed University in Iran and a member of the Iranian Statistical Association. His research interests are

statistical quality control, profile monitoring, and Six Sigma. He has published many papers in the area of statistical process control in high-quality international journals such as *Quality and Reliability Engineering International*, *Communications in Statistics, Computers and Industrial Engineering*, *Journal of Statistical Computation and Simulation*, and *Soft Computing*. He has also published a book with John Wiley and Sons in 2011 entitled *Statistical Analysis of Profile Monitoring*.

Marco Grzegorzcyk is an assistant professor for computational statistics at the Bernoulli Institute of Groningen University (the Netherlands). His research expertise is on Bayesian modelling, and he is particularly interested in network reconstruction methods for learning cellular regulatory pathways from post-genomic data. He received his German Diploma and his PhD in Statistics from TU Dortmund University (Germany), and in between, he was a postdoctoral fellow at Edinburgh University (Scotland). He has published most of his papers in the following fields: bioinformatics and machine learning. He is in the editorial board of the two journals: *Computational Statistics* and the *Journal of Applied Statistics*. In 2018, he was guest editor for a special issue on Statistical Modelling, launched by the journal *Statistica Neerlandica*.

How to cite this article: Ahmadi Yazdi A, Zeinal Hamadani A, Amiri A, Grzegorzcyk M. A new Bayesian multivariate exponentially weighted moving average control chart for phase II monitoring of multivariate multiple linear profiles. *Qual Reliab Engng Int*. 2019;35:2152–2177. <https://doi.org/10.1002/qre.2496>

APPENDIX

A: COVARIANCE MATRIX OF BAYESIAN ESTIMATOR OF REGRESSION PARAMETERS

Assume that $\mathbf{y}_{kj}$ is the $n \times 1$ vector of the j th response variable for the k th sample. Then Σ_{hj} is equal to $\text{cov}(\hat{\beta}_{kh}^b, \hat{\beta}_{kj}^b)$ and

$$\begin{aligned} \text{cov}(\hat{\beta}_{kh}^b, \hat{\beta}_{kj}^b) &= \text{cov}\left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} (\mathbf{X}^T \mathbf{y}_{kh} + \Lambda_0 \mathbf{B}_0), (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} (\mathbf{X}^T \mathbf{y}_{kj} + \Lambda_0 \mathbf{B}_0)\right] \\ &= (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \text{cov}(\mathbf{y}_{kh}, \mathbf{y}_{kj}) \left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T\right]^T \\ &= (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T (\sigma_{hj} \mathbf{I}) \left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T\right]^T \\ &= \sigma_{hj} (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \mathbf{I} \left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T\right]^T \\ &= \sigma_{hj} (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T\right]^T \\ &= \sigma_{hj} (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \mathbf{X} \left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1}\right]^T, \end{aligned} \quad (\text{A1})$$

Consequently, we have

$$\text{cov}(\hat{\beta}_{kh}^b, \hat{\beta}_{kj}^b) = \sigma_{hj} (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \mathbf{X} \left[(\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1}\right]^T. \quad (\text{A2})$$

Hence,

$$\Sigma_{\hat{\beta}_k^b} = \begin{bmatrix} \sigma_{11}\Delta & \sigma_{12}\Delta & \cdots & \sigma_{1p}\Delta \\ \sigma_{21}\Delta & \sigma_{22}\Delta & \cdots & \sigma_{2p}\Delta \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}\Delta & \sigma_{p2}\Delta & \cdots & \sigma_{pp}\Delta \end{bmatrix}, \quad (\text{A3})$$

or equivalently,

$$\Sigma_{\hat{\beta}_k^{bb}} = \Sigma_{\epsilon} \otimes \Delta, \quad (\text{A4})$$

where $\Delta = (\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1} \mathbf{X}^T \mathbf{X} ((\mathbf{X}^T \mathbf{X} + \Lambda_0)^{-1})^T$ is a symmetric $(q+1) \times (q+1)$ matrix.