WILEY

Model robust profile monitoring for the generalized linear mixed model for Phase I analysis

Keerthi Bandara[1] | **Abdel-Salam G. Abdel-Salam[2]** | **Jeffrey B. Birch[3]**

[1]Head of the Department of Agricultural Systems, Faculty of Agriculture, Rajarata University of Sri Lanka, Puliyankulama, Sri Lanka

[2]Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar

[3]Department of Statistics, College of Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia,

Correspondence
Abdel-Salam G. Abdel-Salam, Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar.
Email: abdo@qu.edu.qa

Abstract

The generalized linear mixed model (GLMM) becomes very popular in profile monitoring, especially when the production processes follow nonnormal distribution. In most of the real-life applications in industry, medicine, biology … and so on researchers assume that the response variable follows a Bernoulli or Binomial distribution. The majority of previous studies in profile monitoring focused on parametric modeling using the logistic regression model, with both fixed or random effects, under the assumption of correct model specification. This research considers those cases where the parametric logistic regression model for the family of profiles is unknown or at least uncertain. Consequently, we propose two mixed model methods to monitor profiles from the exponential family: a nonparametric (NP) regression method based on the penalized spline regression technique and a semiparametric method (model robust profile monitoring for the generalized linear mixed model) which combines the advantages of both the parametric and NP methods. Several Hotelling T2 charts that have been studied for a binary response variable with replicates for Phase I profile monitoring. The performance of the proposed method is evaluated by using mean squares of errors and probability of signals criteria. The results showed satisfactory performance of the proposed control charts.

KEYWORDS

GLMM, logistic regression, nonparametric, profile monitoring, semiparametric

1 | INTRODUCTION

Until recently, univariate and multivariate characteristics were commonly used in statistical process control. However, in some applications detecting the changes in the relationship between the response variable and explanatory variable (s) may be the most important consideration rather than detecting the changes in univariate or multivariate quality characteristics. This relationship between the response variable and one or more explanatory variables is called a profile.[1] The act of using various techniques to statistically monitor the process or product profiles is known as profile monitoring.[2,3]

Different methods have been proposed to monitor different types of profiles. Kang and Albin,[1] Kim et al,[4] Mahmoud and Woodall,[5] Mahmoud et al,[6] Zou et al,[7] Mahmoud,[8] Khedmati and Niaki,[9,10] Noorossana et al,[11]
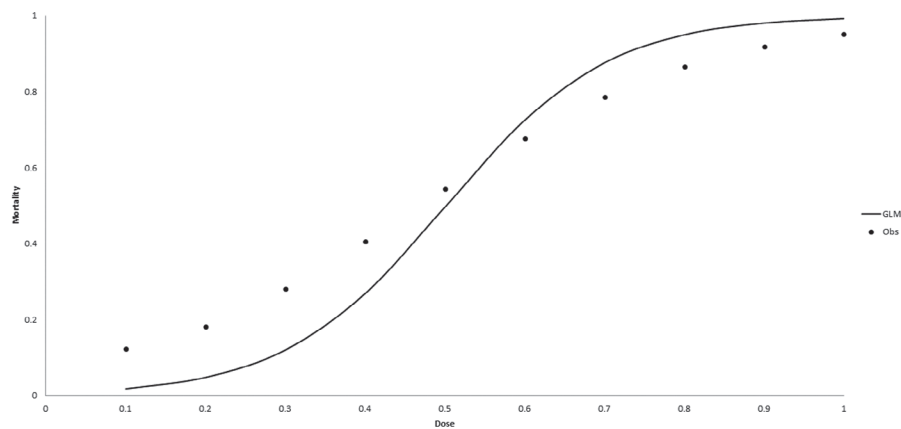
Chiang et al,[12] Kalaei et al,[13] Mahmood et al,[14] Moheghi et al,[15] Touqeer et al[16] proposed methods to monitor linear profiles. Noorossana et al,[17] Eyvazian et al,[18] Soleimani et al,[19] Amiri et al,[20] Ayoubi et al,[21] Paynabar et al,[22] Bahrami et al[23] have studied multivariate profile monitoring. Methods for monitoring nonlinear profiles have been suggested by Vaghefi et al,[24] Williams et al,[25] Steiner et al,[26] Guevara and Vargas,[27] Chen et al,[28] Awad et al,[29] Pan et al,[30] Li et al.[31] Jensen et al[32] proposed a method to monitor the profiles with autocorrelated data using the linear mixed model (LMM). Jensen and Birch[33] have proposed a method to monitor the nonlinear autocorrelated profile using the nonlinear mixed model.

In each of the above-mentioned articles, the response variable is assumed to be normally distributed. However, one can find many examples from almost every subject area where this assumption does not hold. For example, when the response variable is binary, the true distribution is Bernoulli and when the response variable is a count, the underline distribution may be Poisson. In such situations, it is unreasonable to assume a normal distribution for this kind of response variables. Yeh et al[34] proposed five T^2 control charts to monitor the profiles with binary response variables. Amiri et al[35] evaluated two T^2 statistics proposed by Yeh et al[34] for the Poisson response profiles. Koosha and Amiri[36] studied the effect of various link functions on the performances of the T^2 control chart in monitoring the parameters of a logistic profile. Koosha and Amiri[37] studied the effect of the autocorrelation in binary responses on the performances of the T^2 control charts suggested by Yeh et al.[34] Methods for monitoring generalized linear model (GLM) regression profiles have been proposed by Amiri et al.[20] Izadbakhsh et al[38] suggested profile monitoring methods for multinomial responses. Bayer et al,[39] Canterle and Bayer,[40] Lima-Filho et al[41] introduced control charts for monitoring fraction, rate and/or proportion datasets based on beta regression. In each of the above-mentioned articles, one of the assumptions is the response variable can be correctly modeled using the GLM. However, the GLM may be unable to adequately model all the population characteristics contained within the profiles. For example, Figure 1. shows the observations and estimated curve using GLM for dose-response dataset in Nottingham and Birch,[42] where the fitted logistic curve representing the probability of death as a function of drug dosage, consistently underestimates and overestimates the probability of response at low and high doses, respectively.

Thus, the GLM model for profile monitoring can likely be improved by incorporating nonparametric (NP) and semi-parametric (SP) regression methods specifically tailored to distributions contained in the exponential family. Very few studies have been conducted using NP methods for profile monitoring of nonnormal data from the exponential family. Shang et al[43] proposed a profile monitoring method for binary data using NP regression. Piri et al[44] proposed a NP method for profile monitoring of Poisson data using a wavelet approach.

We propose two mixed model methods to monitor profiles from the exponential family; the first one is a NP regression method based on the penalized spline regression technique. The second one is a SP method (model robust profile monitoring for the generalized linear mixed model [MRGLMM]), which combines the advantages of both the parametric and NP methods. A correctly specified parametric (P) model will have the most power in detecting the profile shift,[45] while a NP method can give an improved performance for any type of profile. Splines are recommended for handling sparse data, while polynomial smoothers are preferred for handling dense designs.[44] Our approach is completely general and can be applied to data from any distribution in the exponential family though we used a binary response variable for illustration. Further parametric, NP, and MRGLMM methods will be illustrated using only one regressor, X. However, each method can easily be extended to multiple regressor situation. We introduced versions of the Hotelling's T^2 statistics to determine outlying profiles in the Phase I analysis. The remainder of this article is organized as follows. We first review for the parametric profile monitoring for nonnormal data using the generalized linear mixed model (GLMM) approach.

Then we introduce a NP profile monitoring method for the GLMM using the penalized spline method and a SP profile monitoring method (MRGLMM). The performances of parametric, NP, and MRGLMM methods are compared through a simulation study using binary data.

2 | PARAMETRIC PROFILE MONITORING FOR NONNORMAL DATA USING THE GLMM

The GLMMs extends the GLM by incorporation of random effects into the linear predictor η and can be used to model clustered data when the response has distribution in the exponential family. In general, the expected value of the response variable is modeled when the true response comes from members in this family of distributions.

The expected value of the jth response from the ith cluster ($i = 1, \ldots, m; j = 1, \ldots, n_i$) having any distribution in the exponential family can be modeled using GLMM as

$$g(E[Y_{ij}]) = \eta_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{b}_i, \tag{1}$$

where $g(\cdot)$ is the link function, $E[Y_{ij}]$ is the jth expected value from the ith cluster, η_{ij} is the linear predictor for the jth observation from the ith cluster, $\boldsymbol{\beta}$ is a ($p \times 1$) vector of fixed effect parameters, $\mathbf{x}_{ij}$ is a ($p \times 1$) vector of predictor variables associated with fixed effects, $\mathbf{z}_{ij}$ is a ($q \times 1$) vector of predictor variables associated with random effects, $\mathbf{b}_i$ is a ($q \times 1$) vector of random effects with $\mathbf{b}_i \sim MN(\mathbf{0}, \mathbf{D})$, $\mathbf{D}$ is a ($q \times q$) variance-covariance matrix. Then the parameters of (1) can be estimated by using the pseudo-likelihood method of Wolfinger and O'Connell[46] and the estimated parameter vector and the predicted random effect for the ith cluster can be given as

$$\widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}} = \left(\sum_{i=1}^{m} \boldsymbol{X}_i^T \boldsymbol{V}_i^{-1} \boldsymbol{X}_i \right)^{-1} \left(\sum_{i=1}^{m} \boldsymbol{X}_i^T \boldsymbol{V}_i^{-1} \boldsymbol{v}_i \right), \tag{2}$$

$$\widehat{\boldsymbol{b}}_i = \boldsymbol{D} \boldsymbol{Z}_i^T \boldsymbol{V}_i^{-1} (\boldsymbol{v}_i - \boldsymbol{X}_i \widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}}), \tag{3}$$

where $\boldsymbol{X}_i$ is a ($n_i \times p$) matrix of predictor variables associated with fixed effect, $\boldsymbol{Z}_i$ is a ($n_i \times q$) matrix of predictor variables associated with random effects, $\boldsymbol{V}_i = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i^T + \boldsymbol{\Delta}_i^{-1} \boldsymbol{A}_i^{1/2} \boldsymbol{C}_i \boldsymbol{A}_i^{1/2} \boldsymbol{\Delta}_i^{-1}$ and $\boldsymbol{v}_i \equiv \boldsymbol{\Delta}^{-1} \{ \boldsymbol{y}_i - \mathrm{g}^{-1}(\widehat{\boldsymbol{\eta}}_i) \} + \boldsymbol{X} \widehat{\boldsymbol{\beta}} + \boldsymbol{Z}_i \widehat{\boldsymbol{b}}_i$ is the pseudo-responses vector of ith cluster, $\mathbf{A}_i$ is a diagonal matrix containing the variance functions of the conditional distribution and $\mathbf{C}_i$ is a correlation matrix for the ith cluster, $\boldsymbol{y}_i = [y_{i1}, \ldots, y_{in_i}]^T$, $\widehat{\boldsymbol{b}}_i$s are the estimated best linear unbiased predictors (eblups) and $\boldsymbol{\Delta}_i$ is a diagonal matrix of the first derivatives of the conditional mean evaluated at the fixed effects estimate, $\widehat{\boldsymbol{\beta}}$ and the random effects estimate, $\widehat{\boldsymbol{b}}_i$. Combining pseudo-responses, $\boldsymbol{v}_i$ from all profiles to a ($n \times 1$) vector, $\boldsymbol{v}$, $\boldsymbol{X}_i$s to a ($n \times p$) matrix, $\boldsymbol{X}$, $\boldsymbol{Z}_i$s to a ($n \times q$) matrix, $\boldsymbol{Z}$ and defining $\boldsymbol{V} = \mathrm{diag}(\boldsymbol{V}_i)$, it can be shown that $\widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}}$ can also be expressed as

$$\widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}} = (\boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{X})^{-1} (\boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{v}). \tag{4}$$

The blups of GLMM in stacked form can be written as

$$\widehat{\boldsymbol{b}}_{\mathrm{GLMM}} = \boldsymbol{B} \boldsymbol{Z}^T \boldsymbol{V}^{-1} (\boldsymbol{v} - \boldsymbol{X} \widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}}), \tag{5}$$

where $\boldsymbol{V} = \mathrm{var}(\boldsymbol{v}) = \boldsymbol{Z} \boldsymbol{B} \boldsymbol{Z}^T + \boldsymbol{\Delta}^{-1} \boldsymbol{A}^{1/2} \boldsymbol{R} \boldsymbol{A}^{1/2} \boldsymbol{\Delta}^{-1}$, with $\boldsymbol{R} = \mathrm{diag}(\boldsymbol{C}_i)$ and $\boldsymbol{B} = \mathrm{diag}(\boldsymbol{D})$. Usually, $\boldsymbol{V}$ is unknown and can be estimated by either the maximum likelihood (ML) or the restricted maximum likelihood (REML) method. Then substituting the estimated matrices $\widehat{\boldsymbol{V}}$ and $\widehat{\boldsymbol{B}}$ into (4) and (5) the parameter estimates and the estimated best linear unbiased predictors (eblups) are obtained and the estimated parameter vector for the ith profile, $\widehat{\boldsymbol{\beta}}_i$ can be given by

$$\widehat{\boldsymbol{\beta}}_i = \widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}} + \widehat{\boldsymbol{b}}_i^* \quad i = 1, 2, \ldots, m, \tag{6}$$

where $\widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}}$ represent the profile average (PA) coefficients for all the profiles and $\widehat{\boldsymbol{b}}_i^*$ is a ($p \times 1$) vector containing the elements of $\widehat{\boldsymbol{b}}_i = \widehat{\boldsymbol{D}} \boldsymbol{Z}_i^T \widehat{\boldsymbol{V}}_i^{-1} (\boldsymbol{v}_i - \boldsymbol{X}_i \widehat{\boldsymbol{\beta}}_{\mathrm{GLMM}})$ for the columns of $\boldsymbol{Z}_i$ that are equal to those of $\boldsymbol{X}_i$ and zeros

otherwise. Consequently, $\widehat{\boldsymbol{b}}_i^* = \widehat{\boldsymbol{b}}_i$ when $\boldsymbol{Z}_i = \boldsymbol{X}_i$. Then $g[\widehat{\boldsymbol{y}}] = \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\text{GLMM}}$ is the PA fit which is the estimated PA curve, and $g[\widehat{\boldsymbol{y}}_{cs,i}] = \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\text{GLMM}} + \boldsymbol{Z}_i\widehat{\boldsymbol{b}}_i$ is the fit for the ith profile which is the estimated profile-specific curve for the ith profile.

The key feature of profile monitoring is to detect those profiles that may be "different" from the other profiles. Such differences are indicative of a process that maybe, perhaps temporarily, out-of-control when compared with those profiles representing the in-control situation. The T^2 statistic can be used to detect profiles from an out-of-control process. The $\widehat{\boldsymbol{\beta}}_i$ in (6) can be used to calculate the T^2 statistic for the ith profile (T_i^2) as

$$T_{P0,i}^2 = (\widehat{\boldsymbol{\beta}}_i - \bar{\widehat{\boldsymbol{\beta}}}_i)^T \widehat{\boldsymbol{S}}^{-1} (\widehat{\boldsymbol{\beta}}_i - \bar{\widehat{\boldsymbol{\beta}}}_i), \tag{7}$$

where $\bar{\widehat{\boldsymbol{\beta}}}_i = \frac{1}{m}\sum_{i=1}^{m}\widehat{\boldsymbol{\beta}}_i$, and $\widehat{\boldsymbol{S}} = \frac{1}{2(m-1)}\sum_{i=1}^{m-1}(\widehat{\boldsymbol{\beta}}_{i+1} - \widehat{\boldsymbol{\beta}}_i)(\widehat{\boldsymbol{\beta}}_{i+1} - \widehat{\boldsymbol{\beta}}_i)^T$ is the estimator of variance-covariance matrix, $\widehat{\boldsymbol{S}}$ based on the successive difference.[32] Since the $\sum_{i=1}^{m}\widehat{\boldsymbol{b}}_i = 0$, T_i^2 value for GLMM can be given as a function of the eblups by

$$T_{P0,i}^2 = (\widehat{\boldsymbol{b}}_i)^T \widehat{\boldsymbol{S}}^{-1} (\widehat{\boldsymbol{b}}_i), \tag{8}$$

where $\widehat{\boldsymbol{S}} = \frac{1}{2(m-1)}\sum_{i=1}^{m-1}(\widehat{\boldsymbol{b}}_{i+1} - \widehat{\boldsymbol{b}}_i)(\widehat{\boldsymbol{b}}_{i+1} - \widehat{\boldsymbol{b}}_i)^T$ is the successive difference estimator of $\boldsymbol{S}$.[32]

Abdel-Salam et al[45] introduced a new version of the Hotelling T^2 statistic based on the fitted values from the LMM to measure the departures of the ith estimated PS curve from the estimated PA curve. Similar statistics can be used for GLMM also and can be computed using the fitted values in the link scale as follows.

$$T_{P1,i}^2 = (g[\widehat{\boldsymbol{y}}_{cs,i}] - g[\widehat{\boldsymbol{y}}])^T \widehat{\boldsymbol{S}}^{-1} (g[\widehat{\boldsymbol{y}}_{cs,i}] - g[\widehat{\boldsymbol{y}}]), \tag{9}$$

where $g[\widehat{\boldsymbol{y}}_{cs,i}]$ is a $n' \times 1$ vector representing the parametrically estimated expected value (fitted value) curve in the link scale for the ith profile at $l = 1, 2, \ldots, n'$ using GLMM, $g[\widehat{\boldsymbol{y}}]$ is a $n' \times 1$ vector representing the parametrically estimated expected value curve in the link scale for the PA curve and $\widehat{\boldsymbol{S}}$ is an $(n' \times n')$ appropriately estimated variance-covariance matrix for $g[(\boldsymbol{y}_{cs,i})]$ in the link scale, such as either the methods of moments or the successive difference estimator. The n' is the number of distinct locations for $\mathbf{x}_i^T$, the vector of covariates for the fixed-effects portion of the model. In this work, only one covariate is used.

Then outlying profile(s) can be determined by comparing $T_{Pr^*,i}^2(\mathbf{r}^* = 0, 1)$ with values from the χ^2 distribution where the profile will be declared as outlying if $T_{Pr^*,i}^2 \geq \chi_{df,\alpha}^2$ $i = 1, 2, \ldots m$, where α represents the significance level, df represents the degrees of freedom,[32,33,45] computed as the number of random effects in the estimated P model. Similarly T^2 statistic based on the fitted values in the inverse link scale can be computed as follows:

$$T_{P2,i}^2 = [\widehat{\boldsymbol{y}}_{cs,i} - \widehat{\boldsymbol{y}}]^T \widehat{\boldsymbol{S}}^{-1} [\widehat{\boldsymbol{y}}_{cs,i} - \widehat{\boldsymbol{y}}], \tag{10}$$

where $\widehat{\boldsymbol{y}}_{cs,i} = g^{-1}(\eta_{lj})$ is a $n' \times 1$ vector representing the parametrically estimated expected value (fitted value) curve in the inverse link scale (the raw data scale) for the ith profile at $l = 1, 2, \ldots, n'$ using GLMM, $\widehat{\boldsymbol{y}}$ is a $n' \times 1$ vector representing the parametrically estimated expected value curve in the inverse link scale for the PA curve and $\widehat{\boldsymbol{S}}$ is an $(n' \times n')$ appropriately estimated variance-covariance matrix for $\widehat{\boldsymbol{y}}_{cs,i}$ in the inverse link scale. Then outlying profile(s) can be determined by comparing $T_{P2,i}^2$ with simulated cut off values from a simulation study and the profile will be declared as outlying if $T_{P2,i}^2 \geq$ simulated cutoff value, α $i = 1, 2, \ldots m$, where α represents the significance level.

3 | NP PROFILE MONITORING FOR THE GLMM USING THE PENALIZED SPLINE METHOD

The main idea of p-spline regression is to fit the function $f(x_i)$ parametrically using a sufficiently flexible spline basis with a penalty on the spline coefficients to achieve a smooth fit. P-spline coefficients can be considered as the best linear

unbiased prediction (blups) in the mixed model framework[47] and thus the method used for monitoring out-of-control profiles by parametric methods using GLMM can also be extended to the NP framework.

The expected value at the jth regressor value of the ith profile can be modeled nonparametrically using mixed models as

$$g[E(Y_{ij})] = f(x_{ij}) + \xi_i(x_{ij}) \ i = 1, 2, \dots, m \ j = 1, 2, \dots, n_i, \tag{11}$$

where $g(\cdot)$ is the link function, $E[Y_{ij}]$ is the jth expected value from the ith profile, $f(x_{ij})$ is an unknown smooth function, representing the overall PA curve and $\xi_i(x_{ij})$ is the ith profile smooth function, representing the random difference between the ith PS curve and the PA curve. Both $f(x_{ij})$ and $\xi_i(x_{ij})$ can be approximated using a p-spline regression consisting of a parametric component and a spline component as follows.

$$f(x_{ij}) \approx \beta_0 + \sum_{l=1}^{p} \beta_l x_{ij}^l + \sum_{K=1}^{K_1} u_{iK}(x_{ij} - \kappa_K)_+^p \ i = 1, 2, \dots, m \ j = 1, 2, \dots, n_i, \tag{12}$$

$$\xi_i(x_{ij}) \approx b_{i0} + \sum_{l=1}^{p} b_{il} x_{ij}^l + \sum_{K=1}^{K_2} t_{iK}(x_{ij} - \kappa_K)_+^p \ i = 1, 2, \dots, m \ j = 1, 2, \dots, n_i, \tag{13}$$

where $\beta_0 + \sum_{l=1}^{p} \beta_l x_{ij}^l$ is the parametric component and $\sum_{K=1}^{K_1} u_K(x_{ij} - \kappa_K)_+^p$ is the spline component for the PA curve, and $b_{i0} + \sum_{l=1}^{p} b_{il} x_{ij}^l$ is the random parametric component and $\sum_{K=1}^{K_2} t_{iK}(x_{ij} - \kappa_K)_+^p$ is the random spline component for the ith profile. Then (11) can be rewritten using the matrix notation as

$$g[E(y_{cs,i})] = X_i \beta + Z_i u + X_i b_i + E_i t_i \ i = 1, 2, \dots, m, \tag{14}$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$, $u = (u_1, u_2, \dots, u_{K_1})^T$, $b_i = (b_{i0}, b_{i1}, \dots, b_{ip})^T$,

$t_i = (t_{i1}, t_{i2}, \dots, t_{iK_2})^T$, K_1 is the number of knots for the PA curve and K_2 is the number of knots for the PS curve. The matrices X_i, Z_i, and E_i are specified as follows.

$$X_i = \begin{bmatrix} 1 & x_{i1} & \dots & \dots & \dots & x_{i1}^p \\ 1 & x_{i2} & \dots & \dots & \dots & x_{i2}^p \\ . & & & & & . \\ . & & \dots & \dots & \dots & . \\ . & & & & & . \\ 1 & x_{in_i} & \dots & \dots & \dots & x_{in_i}^p \end{bmatrix}, \ Z_i = \begin{bmatrix} (x_{i1} - \kappa_1)_+^p & \dots & (x_{i1} - \kappa_{K_1})_+^p \\ . & & . \\ . & \dots \dots \dots & . \\ . & & . \\ (x_{in_i} - \kappa_1)_+^p & \dots & (x_{ij} - \kappa_{K_1})_+^p \end{bmatrix},$$

$$E_i = \begin{bmatrix} (x_{i1} - \kappa_1)_+^p & \dots & (x_{i1} - \kappa_{K_2})_+^p \\ . & & . \\ . & \dots \dots \dots & . \\ . & & . \\ (x_{in_i} - \kappa_1)_+^p & \dots & (x_{ij} - \kappa_{K_2})_+^p \end{bmatrix}$$

and (14) can be written in stacked matrix notation as

$$g[E(y)] = X\beta + ZB, \tag{15}$$

where $Z = \begin{bmatrix} Z_1 & X_1 & 0 & \cdots & 0 & E_1 & 0 & \cdots & 0 \\ Z_2 & 0 & X_2 & \cdots & 0 & 0 & E_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_m & 0 & 0 & \cdots & X_m & 0 & 0 & \cdots & E_m \end{bmatrix}$, $B = \begin{bmatrix} u \\ b \\ t \end{bmatrix}$,

with

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{K_1} \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \boldsymbol{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{bmatrix}, \text{ and } \mathrm{Cov}(\mathbf{B}) \equiv \boldsymbol{G} = \begin{bmatrix} \sigma_u^2 \boldsymbol{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \text{block diag}_{1 \le i \le m} \Sigma_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_t^2 \boldsymbol{I} \end{bmatrix},$$

where σ_u^2 controls the amount of smoothing to the estimate of $f(x_{ij})$, σ_b^2 measures the between profiles variation and σ_t^2 controls the amount of smoothing to the estimate $\xi_i(x_{ij})$. Then using either the ML or REML method, the estimated parameters $(\widehat{\beta})$ and the predictions $(\widehat{\boldsymbol{B}})$ can be obtained as

$$\widehat{\beta} = (\boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{v}, \tag{17}$$

$$\widehat{\boldsymbol{B}} = \begin{bmatrix} \widehat{\boldsymbol{u}} \\ \widehat{\boldsymbol{b}} \\ \widehat{\boldsymbol{t}} \end{bmatrix} = \boldsymbol{G} \boldsymbol{Z}^T \boldsymbol{V}^{-1} (\boldsymbol{v} - \boldsymbol{X}\widehat{\beta})$$

with $\boldsymbol{V} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^T + \boldsymbol{\Delta}^{-1} \boldsymbol{A}^{1/2} \boldsymbol{R} \boldsymbol{A}^{1/2} \boldsymbol{\Delta}^{-1}$ where $\boldsymbol{\Delta}$ is a block diagonal matrix of the first derivatives of the conditional mean evaluated at the fixed effects estimate, $\widehat{\beta}$ and the random effects estimate, $\widehat{\boldsymbol{B}}$, $\boldsymbol{A}$ is a block diagonal matrix containing the variance functions in the diagonal blocks, $\boldsymbol{R}$ is a correlation matrix and $\boldsymbol{v} \equiv \boldsymbol{\Delta}^{-1} \{\boldsymbol{y}|\boldsymbol{B} - g^{-1}(\widehat{\eta})\} + \boldsymbol{X}\widehat{\beta} + \boldsymbol{Z}\widehat{\boldsymbol{B}}$ is the pseudo-responses vector representing pseudo responses from all clusters.

Hence the estimated expected value of the observation vector for the PA curve using the mixed-effect model can be obtained as

$$g[\widehat{\boldsymbol{y}}_{PA}^{\mathrm{NGLMM}})] = \boldsymbol{X}\widehat{\beta} + \boldsymbol{Z}\widehat{\boldsymbol{u}} \tag{19}$$

and the estimated expected value of the observation vector for the ith profile can be obtained as

$$g[\widehat{\boldsymbol{y}}_{cs,i}^{\mathrm{NGLMM}})] = \boldsymbol{X}_i\widehat{\beta} + \boldsymbol{Z}_i\widehat{\boldsymbol{u}} + \boldsymbol{X}_i\widehat{\boldsymbol{b}}_i + \boldsymbol{E}_i\widehat{\boldsymbol{t}}_i \quad i = 1, 2, \ldots, m, \tag{20}$$

where $g[\widehat{\boldsymbol{y}}_{PA}^{\mathrm{NGLMM}}]$ is the p-spline estimation for the expected values of the PA curve in the link scale using the mixed effect model and $g[\widehat{\boldsymbol{y}}_{cs,i}^{\mathrm{NGLMM}}]$ is the p-spline estimator of the expected values for the ith profile in the link scale. These estimated expected value vectors can be used to calculate the T^2 statistic for the ith profile as

$$T_{NP0,i}^2 = (g[\widehat{\boldsymbol{y}}_{cs,i}^{\mathrm{NGLMM}}] - g[\widehat{\boldsymbol{y}}_{PA}^{\mathrm{NGLMM}}])^T \widehat{\boldsymbol{S}}^{-1} (g[\widehat{\boldsymbol{y}}_{cs,i}^{\mathrm{NGLMM}}] - g[\widehat{\boldsymbol{y}}_{PA}^{\mathrm{NGLMM}}]) \quad i = 1, 2, \ldots, m, \tag{21}$$

where $g[\widehat{\boldsymbol{y}}_{cs,i}^{\mathrm{NGLMM}}]$ is the fitted expected value curve for the ith profile at $l = 1, 2, \ldots, n'$ in the link scale, and $\widehat{\boldsymbol{S}}$ is an $(n' \times n')$ appropriately estimated variance-covariance matrix for $g[\widehat{\boldsymbol{y}}_{cs,i}^{\mathrm{NGLMM}}]$ and $g[\widehat{\boldsymbol{y}}_{PA}^{\mathrm{NGLMM}}]$ is the fitted expected value curve for the PA in the link scale.

It is also possible to calculate T^2 statistic for the ith profile using the estimated predicted random effect vector, $\widehat{\boldsymbol{\phi}}_i$ where

$$\widehat{\boldsymbol{\phi}}_i = \begin{bmatrix} \widehat{\boldsymbol{b}}_i \\ \widehat{\boldsymbol{t}}_i \end{bmatrix} \quad i = 1, 2, \ldots, m$$

with $\widehat{\boldsymbol{b}}_i = \left(b_{i0}, \ldots, b_{ip}\right)^T$ and $\boldsymbol{t}_i = (t_{i1}, t_{i2}, \ldots, t_{iK_2})^T$. Assuming the locations of the regressor values and the number of observations at each location is the same across all m profiles, the T^2 statistic for the ith profile can be calculated as

$$T_{NP1,i}^2 = (\widehat{\boldsymbol{\phi}}_i - \overline{\boldsymbol{\phi}})^T \widehat{\boldsymbol{S}}^{-1} (\widehat{\boldsymbol{\phi}}_i - \overline{\boldsymbol{\phi}}), \tag{23}$$

where $T^2_{NP1,i}$ is the T^2 statistic for the ith profile and $\widehat{S}$ is an $(n' \times n')$ appropriately estimated variance-covariance matrix. For example, using the sample covariance matrix based on the successive difference estimator,[45] $\widehat{S}$ would be calculated as $\frac{\sum_{i=1}^{m-1}(\widehat{\phi}_{i+1}-\widehat{\phi}_i)(\widehat{\phi}_{i+1}-\widehat{\phi}_i)^T}{2(m-1)}$.

Then outlying profile(s) can be determined by comparing $T^2_{NPr^*,i}$ ($r^* = 0,1$), with values from χ^2 distribution where the profile will be declared as outlying if $T^2_{NPr^*,i} \geq \chi^2_{df1,\alpha}$ $i = 1, 2, \ldots m$, where α represents the significance level, $df1$ represents the degrees of freedom which is equal to the number of random effects plus the number of knots in the estimated NP model.[45]

Similarly, the estimated expected value of the observation vector for the PA curve in the link scale is $g[\widehat{y}^{NGLMM}_{PA}] = X\widehat{\beta} + Z\widehat{u}$ and it can be given in inverse link scale as

$$\widehat{y}^{NGLMM}_{PA} = g^{-1}(X\widehat{\beta} + Z\widehat{u}). \tag{24}$$

The estimated expected value of the observation vector for the ith profile in the inverse link scale can be given as

$$\widehat{y}^{NGLMM}_{cs,i} = g^{-1}(X_i\widehat{\beta} + Z_i\widehat{u} + X_i\widehat{b}_i + E_i\widehat{t}_i) \; i = 1, 2, \ldots, m, \tag{25}$$

where $\widehat{y}^{NGLMM}_{PA}$ is the p-spline estimation for the expected values of the PA curve in the inverse link scale using the mixed effect model and $\widehat{y}^{NGLMM}_{cs,i}$ is the p-spline estimation of expected values for the ith profile in the inverse link scale. Thus, a new T^2 statistic for the ith profile based on the nonparamatrically fitted values in the inverse link scale can be given as

$$T^2_{NP3,i} = [\widehat{y}^{NGLMM}_{cs,i} - \widehat{y}^{NGLMM}_{PA}]^T \widehat{S}^{-1}[\widehat{y}^{NGLMM}_{cs,i} - \widehat{y}^{NGLMM}_{PA}] \; i = 1, 2, \ldots, m, \tag{26}$$

where $\widehat{y}^{NGLMM}_{cs,i}$ is a $n' \times 1$ vector representing the nonparametrically estimated expected value (fitted value) curve in the inverse link scale for the ith profile at $l = 1, 2, \ldots, n'$ using p-spline regression, $\widehat{y}^{NGLMM}_{PA}$ is a $n' \times 1$ vector representing the nonparametrically estimated expected value curve in the inverse link scale for the PA curve and $\widehat{S}$ is an $(n' \times n')$ appropriately estimated variance-covariance matrix for $\widehat{y}^{NGLMM}_{cs,i}$ in the inverse link scale.

Then outlying profile(s) can be determined by comparing $T^2_{NP3,i}$ with simulated cut off values from a simulation study and the profile will be declared as outlying if $T^2_{NP3,i} \geq$ simulated cut off value α $i = 1, 2, \ldots m$, where α represents the significance level.

4 | MODEL ROBUST PROFILE MONITORING FOR THE GENERALIZED LINEAR MIXED MODEL

MRGLMM can be considered as an extension of mixed model robust regression to the area of profile monitoring. In this method, estimates are obtained by incorporating a mixed model fit obtained by both the parametric and the NP methods via a convex combination.

The MRGLMM fit in the link scale for the ith profile is given as

$$g[\widehat{y}^{MRGLMM}_{cs,i}] = (1 - \lambda)g[\widehat{y}^{PGLMM}_{cs,i}] + \lambda \, g[\widehat{y}^{NGLMM}_{cs,i}] i = 1, 2, \ldots, m, \tag{27}$$

where $g[\widehat{y}^{MRGLMM}_{cs,i}]$, $g[\widehat{y}^{PGLMM}_{cs,i}]$, and $g[\widehat{y}^{NGLMM}_{cs,i}]$ represents the MRGLMM fit, parametric fit, and NP mixed model fit for the ith profile in the link scale, respectively. The λ is the mixing parameter such that $0 \leq \lambda \leq 1$. Following Mays et al,[48] the mixing parameter λ can be estimated as

$$\widehat{\lambda} = \frac{[g(\widehat{y}^{NPGLMM}_{i,-i}) - g(\widehat{y}^{PGLMM}_{i,-i})]^T - [g(y) - g(\widehat{y}^{PGLMM}_{PA})]}{[g(\widehat{y}^{NPGLMM}_{PA}) - g(\widehat{y}^{PGLMM}_{PA})]^T[g(\widehat{y}^{NPGLMM}_{PA}) - g(\widehat{y}^{PGLMM}_{PA})]}, \tag{28}$$

where $(\widehat{y}^{NPGLMM}_{i,-i})$ and $(\widehat{y}^{PGLMM}_{i,-i})$, respectively, represent the nonparametrically and parametrically estimated expected values for the ith profile without using the observations from the ith profile, $(\widehat{y}^{NPGLMM}_{PA})$ and $(\widehat{y}^{PGLMM}_{PA})$, respectively, represent the nonparametrically and parametrically estimated PA curve, and (y) represents the observation vector. Then to detect

the presence of possible profiles from an out-of-control process, several versions of the MRGLMM T^2 statistics are proposed. One version is the MRGLMM T^2 statistic for the ith profile based on the fitted value in the link scale can be given as

$$T^2_{\text{MRGLMM0},i} = (g[\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{cs,i})] - g[\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{PA})])^T \widehat{\boldsymbol{S}}^{-1} (g[\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{cs,i})] - g[\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{PA})]) \; i = 1, 2, \ldots, m, \tag{29}$$

where $\widehat{\boldsymbol{S}}$ is an appropriate estimate of the variance-covariance matrix and $g[\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{PA})] = \frac{\sum_{i=1}^m g[\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{cs,i})]}{m}$. Here either the pooled sample variance-covariance matrix ($\widehat{\boldsymbol{S}}_P$) or sample covariance matrix based on the successive difference estimator ($\widehat{\boldsymbol{S}}_D$) based on the PS fits can be used. Then the unusual profile(s) can be determined by comparing $T^2_{\text{MRGLMM0},i}$ with values from χ^2 distribution where the profile will be declared as outlying if $T^2_{\text{MRGLMM0},i} \geq \chi^2_{df_2,\alpha} \; i = 1, 2, \ldots m$, where df_2 represents the SP degrees of freedom, computed as $(1 - \widehat{\lambda})df^P + \widehat{\lambda}.df^{NP}$. The df^P is df of parametric model and df^{NP} is df of NP model, and $\widehat{\lambda}$ is the estimated mixing parameter. It is also possible to express the MRGLMM T^2 statistic for the ith profile using a convex combination of the estimated random effects. Defining a new vector, $\widehat{\boldsymbol{\psi}}_i$ where

$$\widehat{\boldsymbol{\psi}}_i = \begin{bmatrix} (1 - \widehat{\lambda})\widehat{\boldsymbol{b}}_i \\ \widehat{\lambda}\widehat{\boldsymbol{\phi}}_i \end{bmatrix} \; i = 1, 2, \ldots, m$$

with $\widehat{\boldsymbol{b}}_i$ is the parametrically estimated random effects as in (3) and $\widehat{\boldsymbol{\phi}}_i$ is the vector of estimated random coefficients (random parametric coefficients and random spline coefficients (knots) from mixed p-spline regression as given in (12). Then the MRGLMM T^2 statistic for the ith profile can be given by

$$T^2_{\text{MRGLMM1},i} = (\widehat{\boldsymbol{\psi}}_i - \overline{\boldsymbol{\psi}})^T \left[\frac{\sum_{i=1}^{m-1}(\widehat{\boldsymbol{\psi}}_{i+1} - \widehat{\boldsymbol{\psi}})(\widehat{\boldsymbol{\psi}}_{i+1} - \widehat{\boldsymbol{\psi}})^T}{2(m-1)} \right]^{-1} (\widehat{\boldsymbol{\psi}}_i - \overline{\boldsymbol{\psi}}) \; i = 1, 2, \ldots, m, \tag{31}$$

where $\overline{\boldsymbol{\psi}} = \frac{\sum_i^m \widehat{\boldsymbol{\psi}}_i}{m}$.

Then the unusual profile(s) can be determined by comparing $T^2_{\text{MMRPM1},i}$ with values from χ^2 distribution where the profile will be declared as outlying if $T^2_{\text{MMRPM1},i} \geq \chi^2_{df_2,\alpha} \; i = 1, 2, \ldots m$.

Another way to calculate the MRGLMM T^2 statistic for the ith profile is using a random effects vector, $\widehat{\boldsymbol{\omega}}_i$ where

$$\widehat{\boldsymbol{\omega}}_i = \begin{bmatrix} (1 - \widehat{\lambda})\widehat{\boldsymbol{b}}_{0,p,i} \\ (1 - \widehat{\lambda})\widehat{\boldsymbol{b}}_{1,p,i} + \widehat{\lambda}\widehat{\boldsymbol{b}}_{1,np,i} \\ \widehat{\lambda}\widehat{\boldsymbol{t}}_i \end{bmatrix} \; i = 1, 2, \ldots, m,$$

where $\widehat{\boldsymbol{b}}_{0,P,i}$ is the vector containing estimated random effects that appear in the parametric method and do not appear in the NP method, $\widehat{\boldsymbol{b}}_{1,P,i}$ and $\widehat{\boldsymbol{b}}_{1,NP,i}$ are vectors containing the estimated random effects which are common to both parametric and NP methods, $\widehat{\boldsymbol{t}}_i$ is the vector containing specific random effects for NP method (ie, knots coefficients). Then the new MRGLMM T^2 statistic for the ith profile can be given by

$$T^2_{\text{MRGLMM2},i} = (\widehat{\boldsymbol{\omega}}_i - \overline{\boldsymbol{\omega}})^T \left[\frac{\sum_{i=1}^{m-1}(\widehat{\boldsymbol{\omega}}_{i+1} - \widehat{\boldsymbol{\omega}})(\widehat{\boldsymbol{\omega}}_{i+1} - \widehat{\boldsymbol{\omega}})^T}{2(m-1)} \right]^{-1} (\widehat{\boldsymbol{\omega}}_i - \overline{\boldsymbol{\omega}}) \; i = 1, 2, \ldots, m, \tag{33}$$

where $\overline{\boldsymbol{\omega}} = \frac{\sum_i^m \widehat{\boldsymbol{\omega}}_i}{m}$. Then the unusual profile(s) can be determined by comparing $T^2_{\text{MMRPM2},i}$ with values from χ^2 distribution where the profile will be declared as outlying if $T^2_{\text{MMRPM2},i} \geq \chi^2_{df_2,\alpha} \; i = 1, 2, \ldots m$. Similarly as in the cases of parametric profile monitoring using GLMM and NP profile monitoring for the GLMM using the penalized spline methods, T^2 statistic for MRGLMM method also can be computed using the fitted values in the inverse link scale, the raw data scale, and can be given as

$$T^2_{\text{MRGLMM3},i} = [\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{cs,i} - \widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{PA}]^T \widehat{\boldsymbol{S}}^{-1} [\widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{cs,i} - \widehat{\boldsymbol{y}}^{\text{MRGLMM}}_{PA}] \; i = 1, 2, \ldots, m, \tag{34}$$

where $\hat{\boldsymbol{y}}_{cs,i}^{\text{MRGLMM}} = (1 - \lambda)\hat{\boldsymbol{y}}_{cs,i}^{\text{PGLMM}} + \lambda\,\hat{\boldsymbol{y}}_{cs,i}^{\text{NGLMM}}$ represent the MRGLMM fit for the ith profile in the inverse link scale, $\hat{\boldsymbol{y}}_{cs,i}^{\text{PGLMM}}$ represent the parametric fit for ith profile in the inverse link scale, $\hat{\boldsymbol{y}}_{cs,i}^{\text{NGLMM}}$ represent the NP mixed model fits for the ith profile in the inverse link scale, λ represent the mixing parameter, $\hat{\boldsymbol{y}}_{PA}^{\text{MRGLMM}} = \frac{\sum_{i=1}^{m} \hat{\boldsymbol{y}}_{cs,i}^{\text{MRGLMM}}}{m}$ and $\hat{\boldsymbol{S}}$ is an appropriate estimate of the variance-covariance matrix such as pooled sample variance-covariance matrix ($\hat{\boldsymbol{S}}_P$) or sample covariance matrix based on the successive difference estimator ($\hat{\boldsymbol{S}}_D$). Then outlying profile(s) can be determined by comparing $T_{\text{MRGLMM3},i}^2$ with simulated cut off values from a simulation study and the profile will be declared as outlying if $T_{\text{MRGLMM3},i}^2 \geq$ simulated cut off value at α, $i = 1, 2, \ldots m$, where α represents the significance level.

5 | A MONTE-CARLO STUDY

Of the many distributions in the exponential family, we choose to illustrate our methods using the 0/1 response variable, a common response variable in many industrial applications. We consider a variety of replications (r) of the response at each design point, a single covariate (x_j), ranging from $r = 1$ to $r = 10$. We let m equal the total number of profiles and q represent the number of in-control profiles that were generated with the probability of success equal to the pdf,

$$G_{ij}(x_{ij}) = (1 - \gamma)(A) + \gamma(B), \tag{35}$$

where $G_{ij}(x_{ij})$ is the probability of success for observations generated for the ith profile at the jth level of covariate (x_{ij}), γ represents the misspecification parameter and satisfies $0 \leq \gamma \leq 1$, A is a logistic pdf defined as $A = \left[\frac{1}{1+e^{-((-5)+(10+b_{i1})x_{ij})}}\right]$ and B is a mixture of two logistics pdfs defined as $B = \left[0.5\left(\frac{1}{1+e^{-((-5)+(20+b_{i2})x_{ij})}}\right) + 0.5\left(\frac{1}{1+e^{-((-15)+(20+b_{i3})x_{ij})}}\right)\right]$, and b_{i1}, b_{i2}, b_{i3} are independent random effects generated from the normal distribution and i is the profile number, $i = 1, \ldots, q, j = 1, \ldots, k$. To represent the out-of-control profiles we consider profiles subject to a sustained shift. Consequently, the observations for the m-q out-of-control profiles were generated by shifting the intercepts of each components of (35) by the amount d_0. Thus, the mean function used to generate observations for shifted profile can be given as

$$G_{ij}(x_{ij}) = (1 - \gamma)(A) + \gamma(B), \tag{36}$$

where A and B are defined as

$$A = \left[\frac{1}{1 + e^{-5+d_0} + (10 + b_{i1})x_{ij}}\right],$$

and

$$B = \left[0.5\left(\frac{1}{1 + e^{-((-5+d_0)+(20+b_{i2})x_{ij})}}\right) + 0.5\left(\frac{1}{1 + e^{-((-15+d_0)+(20+b_{i3})x_{ij})}}\right)\right].$$

Three settings for k ($k = 10, 20, 30$) the number of distinct locations of the single regressor "X," were used with the values for X determined as equally spaced design points on the interval from zero to one. The number of binary replications (r) at each value of X was set at five different values ($r = 1, 2, 3, 5,$ and 10). The random effects b_{i1}, b_{i2}, b_{i3} were generated independently from the normal distribution with mean zero and variances 5, 20, 20, respectively.

We considered five different misspecification levels (γ) ranging from zero to one ($\gamma = 0, 0.25, 0.5, 0.75, 1$) and the users model equals the true model at $\gamma = 0$ where there is no misspecification and the true profile is simply the logistic curve. The degree of profile misspecification increases with γ and the highest level of misspecification occurs at $\gamma = 1$. Figure 2 shows the effect of γ, the misspecification parameter, on the true PA curves generated from the above model (5.3) at $\gamma = 0$, 0.25, 0.5, 0.75, and 1. We used 1000 replicates for each setting (m, k, r, γ, and different value of shifts) in order to obtain results with acceptable precision. The short version for SAS code is given in the Appendix.

The parametric and NP fits for these simulated profiles were obtained using the GLIMMIX procedure of SAS software with the logit link function. The MRGLMM fits were obtained by using a SAS macro written by the author and is available upon request. The parametric fits were calculated based on two random effects (random slope and random intercept). The NP fits were obtained using a random slope and a number of knots sufficient, determined by simulation, to fit the profiles at each misspecification level.
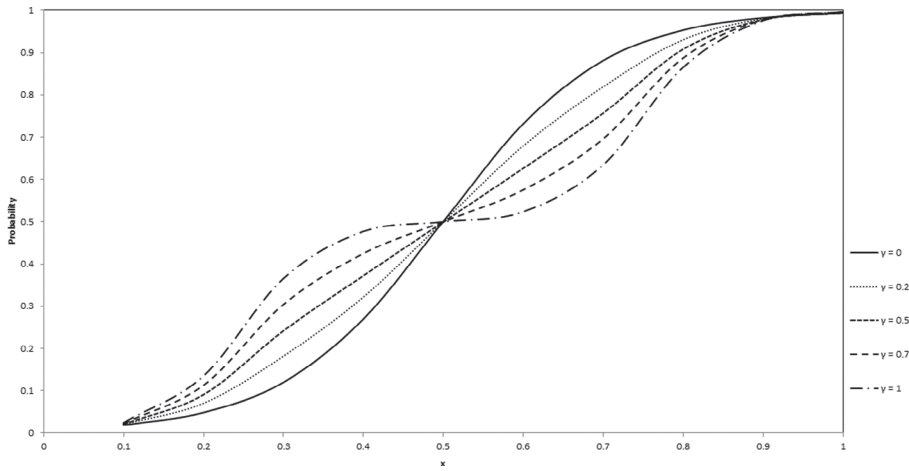
FIGURE 2 Plots of profile average curves at different misspecification levels

5.1 | Metrics of interest

The performance of the parametric, NP, and MRGLMM methods were compared using the simulated integrated mean square error (SIMSE) and the simulated probability of signal (POS). The SIMSE was used as a measure of the goodness of fit. It was calculated using the differences between the estimates for the ith profile and the true values at 50 equally spaced design points between 0 and 1, assuming that these 50 points would properly represent the features in entire curve. The SIMSE was computed as

$$\text{SIMSE} = \frac{1}{m} \sum_{i=1}^{m} (\hat{\boldsymbol{y}}_{cs,i} - \boldsymbol{\mu}_{cs,i})^T (\hat{\boldsymbol{y}}_{cs,i} - \boldsymbol{\mu}_{cs,i}), \tag{37}$$

where SIMSE is the simulated integrated mean square error occurring in estimating m profiles, m is the number of profiles, $\hat{\boldsymbol{y}}_{cs,i}$ is a vector of estimated values for the ith profile using either the parametric, NP, or MRGLMM methods, $\boldsymbol{\mu}_{cs,i}$ represents the vector of true probabilities generated from (35) for the ith profile.

The simulated POS is the probability of at least one of m calculated T^2 statistics exceeds the control limits in a simulation study. Ten T_i^2 statistics, two each for parametric, NP, and MRGLMM methods based on fitted values in link scale and fitted values in inverse link scale, and four T_i^2 statistics, one each for the parametric and NP methods, and two for the MRGLMM method based on estimated random-effects were computed for each Monte Carlo replication. Since these T_i^2 statistics ($i = 1, 2, \ldots, m$) computed from each method are based on the same estimated mean and variance-covariance matrix, the m T_i^2 statistics obtained for each method are correlated; hence it is very difficult to determine their joint distribution. Hence the upper control limit (UCL) of the computed T_i^2 statistics was calculated assuming the T_i^2 statistics are independent.[5,25,32,45] Then the approximate overall probability of false alarm for a sample of m independent statistics is $\alpha_{\text{overall}} = 1 - (1 - \alpha)^m$, where the probability of a false alarm for any individual T_i^2 statistic for a given overall probability of a false alarm, α, is $\alpha = 1 - (1 - \alpha_{\text{overall}})^{\frac{1}{m}}$. Thus, the UCL using the asymptotic chi-squared distribution of each T_i^2 is given by $\text{UCL}_{\chi^2} = \chi^2_{(1-\alpha,df)}$ where $\chi^2_{(1-\alpha,df)}$ is the $(1 - \alpha)$ quantile of a $\chi^2_{(df)}$ distribution.

In this study, the nominal α value, the probability of signal (POS) for in-control dataset, was set at 0.05. The actual simulated POS was calculated as the proportion of simulated datasets that exceed the 95th percentile of computed T_i^2 statistics using in-control model with no misspecification.

5.2 | Simulation results for profile estimation

Table 1 gives the estimated average mixing parameter $(\overline{\hat{\lambda}})$ and the SIMSE values at different misspecification levels for various combinations of k and r. The bolded values represent the smallest SIMSE values among the three methods and the values in brackets are Monte-Carlo standard errors from 1000 replications. All of the results in Table 1 are as

TABLE 1 SIMSE and average $\bar{\hat{\lambda}}$ across k, r, and γ

k	r	γ	$\bar{\hat{\lambda}}$	P	NP	MRGLMM
10	1	0.00	0.08	0.0103 (0.00273)	0.0174 (0.00406)	**0.0100** (0.00263)
		0.25	0.11	0.0097 (0.00260)	0.0168 (0.00440)	**0.0095** (0.00248)
		0.50	0.14	0.0095 (0.00209)	0.0174 (0.00430)	**0.0092** (0.00199)
		0.75	0.25	0.0116 (0.00206)	0.0171 (0.00367)	**0.0108** (0.00195)
		1.00	0.35	0.0166 (0.00236)	0.0204 (0.00362)	**0.0150** (0.00223)
	5	0.00	0.08	**0.0029** (0.00075)	0.0041 (0.00094)	**0.0029** (0.00074)
		0.25	0.24	0.0040 (0.00085)	0.0054 (0.00100)	**0.0039** (0.00080)
		0.50	0.42	0.0055 (0.00087)	0.0062 (0.00104)	**0.0048** (0.00078)
		0.75	0.66	0.0080 (0.00092)	0.0065 (0.00097)	**0.0058** (0.00080)
		1.00	0.84	0.0120 (0.00118)	0.0077 (0.00118)	**0.0075** (0.00104)
	10	0.00	0.06	**0.0016** (0.00041)	0.0023 (0.00051)	**0.0016** (0.00040)
		0.25	0.37	0.0028 (0.00058)	0.0032 (0.00054)	**0.0025** (0.00049)
		0.50	0.57	0.0045 (0.00063)	0.0040 (0.00061)	**0.0034** (0.00050)
		0.75	0.83	0.0070 (0.00074)	0.0041 (0.00059)	**0.0040** (0.00054)
		1.00	0.97	0.0101 (0.00099)	**0.0048** (0.00069)	**0.0048** (0.00068)
20	1	0.00	0.07	0.0063 (0.00170)	0.0103 (0.00230)	**0.0062** (0.00167)
		0.25	0.14	0.0067 (0.00158)	0.0102 (0.00226)	**0.0065** (0.00151)
		0.50	0.25	0.0075 (0.00136)	0.0109 (0.00219)	**0.0071** (0.00128)
		0.75	0.41	0.0098 (0.00133)	0.0112 (0.00204)	**0.0086** (0.00127)
		1.00	0.56	0.0143 (0.00166)	0.0136 (0.00223)	**0.0117** (0.00160)
	5	0.00	0.06	**0.0016** (0.00040)	0.0023 (0.00049)	**0.0016** (0.00040)
		0.25	0.37	0.0027 (0.00055)	0.0032 (0.00054)	**0.0024** (0.00048)
		0.50	0.58	0.0045 (0.00066)	0.0040 (0.00060)	**0.0033** (0.00052)
		0.75	0.84	0.0070 (0.00068)	0.0041 (0.00056)	**0.0039** (0.00051)
		1.00	0.97	0.0103 (0.00099)	**0.0047** (0.00069)	**0.0047** (0.00067)
	10	0.00	0.05	**0.0008** (0.00022)	0.0012 (0.00027)	**0.0008** (0.00022)
		0.25	0.52	0.0019 (0.00036)	0.0019 (0.00029)	**0.0016** (0.00027)
		0.50	0.69	0.0036 (0.00050)	0.0026 (0.00037)	**0.0023** (0.00033)
		0.75	0.92	0.0056 (0.00056)	**0.0026** (0.00034)	**0.0026** (0.00033)
		1.00	0.99	0.0084 (0.00070)	**0.0029** (0.00040)	**0.0029** (0.00040)
30	1	0.00	0.07	**0.0045** (0.00111)	0.0074 (0.00153)	**0.0045** (0.00110)
		0.25	0.17	0.0054 (0.00122)	0.0078 (0.00155)	**0.0053** (0.00117)
		0.50	0.31	0.0065 (0.00110)	0.0086 (0.00154)	**0.0060** (0.00102)
		0.75	0.53	0.0089 (0.00110)	0.0087 (0.00147)	**0.0073** (0.00104)
		1.00	0.7	0.0132 (0.00137)	0.0105 (0.00164)	**0.0097** (0.00129)

(Continues)

k	r	γ	$\bar{\hat{\lambda}}$	P	NP	MRGLMM
	5	0.00	0.06	**0.0011** (0.00028)	0.0016 (0.00034)	**0.0011** (0.00028)
		0.25	0.46	0.0022 (0.00042)	0.0023 (0.00039)	**0.0019** (0.00035)
		0.50	0.65	0.0040 (0.00056)	0.0031 (0.00045)	**0.0027** (0.00039)
		0.75	0.89	0.0063 (0.00065)	**0.0031** (0.00042)	**0.0031** (0.00040)
		1.00	0.99	0.0091 (0.00081)	**0.0035** (0.00049)	**0.0035** (0.00049)
	10	0.00	0.04	**0.0006** (0.00014)	0.0008 (0.00017)	**0.0006** (0.00014)
		0.25	0.61	0.0016 (0.00030)	0.0014 (0.00021)	**0.0012** (0.00020)
		0.50	0.73	0.0031 (0.00042)	0.0020 (0.00032)	**0.0018** (0.00025)
		0.75	0.94	0.0050 (0.00047)	**0.0020** (0.00026)	**0.0020** (0.00025)
		1.00	1	0.0079 (0.00065)	**0.0022** (0.00030)	**0.0022** (0.00030)

Note: Monti-Carlo standard errors are in parenthesis. Best values in Bold.

Abbreviations: MRGLMM, model robust profile monitoring for the generalized linear mixed model; NP, nonparametric; SIMSE, simulated integrated mean square error.

expected. That is, as the number of observations per profile (r or k) increases, the SIMSE of each method decreases, as more observations will result in more precise estimates. The average value of the estimated mixing parameter ($\bar{\hat{\lambda}}$) for the correctly specified model ($\gamma = 0$) varies from 0.03 to 0.08 and these values are close to the true value of 0, especially that when there are a higher number of observations per profile, $\bar{\hat{\lambda}}$ is very close to zero. Thus, when $\gamma = 0$, MRGLMM results are identical to the parametric results (or nearly so) since the MRGLMM fit is mostly or all composed of the parametric fit.

For the highest level of model misspecification ($\gamma = 1$), the average value of the estimated mixing parameter ($\bar{\hat{\lambda}}$) ranges from 0.35 to 1 and is somewhat different from the ideal value of one, especially when the number of replications per profile is small, as in the $k = 10$ and $r = 1$ case. But $\bar{\hat{\lambda}}$ is very close to the ideal value when there is an adequate number of replications per profile. Thus, MRGLMM results are identical to the NP results (or nearly so), since MRGLMM fit is mostly or all composed of the NP fit when the estimated mixing parameter is close to one.

A similar observation can be made for other levels of model misspecification also. That is, when there are a small number of observations per profile, the average value of the estimated mixing parameter ($\bar{\hat{\lambda}}$) varies between 0 and 1, as one would expect. For intermediate values of γ, the mild model misspecification case, the MRGLMM method performs better than either the parametric or NP methods. Hence the main advantage of the MRGLMM method over either the parametric or the NP methods in fitting a curve to the data occurs when the user's model is partially correct.

Thus, these simulation results support our claim that the MRGLMM method is robust to model misspecification. All these results are similar to the findings of other researchers when evaluating the performance of the MRR method compared with the parametric or NP methods for many different and diverse statistical situations, see for example, Mays et al,[49] Robinson et al,[50] Pickle et al,[51] Wan and Birch,[52] and Abdel-Salam et al.[45]

5.3 | Simulation results for control limits

Table 2 gives the simulated control limits for each T^2 statistic with a different number of design points (k) and a different number of replications (r) for the $m = 30$ profile scenario. The control limits computed from the T_i^2 statistics based on the eblups and fitted value in the link scale for the parametric method as well as NP methods give identical results, and the control limits computed from the second T_i^2 (T^2_{MRGLMM2}) based on eblups and fitted value in link scale (T^2_{MRGLMM0}) for MRGLMM also identical numerically confirming the theoretical result of Bandara.[53]

TABLE 2 Simulated control limits (95% cut off values) and χ^2_{df} control limits for each T^2 statistics with different number of design points (k) 86 and different number of replications (r) for $m = 30$ profile scenario

k	r	T^2 statistic based on eblups				T^2 statistic based on fitted value in link scale		
		T^2_{P0}	T^2_{NP2}	$T^2_{MRGLMM1}$	$T^2_{MRGLMM2}$	T^2_{P1}	T^2_{NP0}	$T^2_{MRGLMM0}$
χ^2_{df}		**12.8**	**15.1**	**19.3**	**17.3**	**12.8**	**15.1**	**17.3**
30	10	11.15018	14.56320	22.25727	15.25522	11.15018	14.56320	15.25522
	5	11.46921	14.36956	23.24260	15.63807	11.46921	14.36956	15.63807
	3	11.82638	15.16028	23.01403	16.61421	11.82638	16.16028	16.61421
	2	12.55043	15.83710	23.68356	17.41661	12.55043	15.83710	17.41661
	1	13.09435	15.03447	23.80980	19.51177	13.09435	15.03447	19.51177
20	10	11.24573	14.09540	21.77913	14.86252	11.24573	14.09540	14.86252
	5	11.77257	14.98102	24.14715	17.00617	11.77257	14.98102	17.00617
	3	12.08321	14.91906	22.82393	17.07817	12.08321	14.91906	17.07817
	2	12.20828	14.47795	24.20207	18.75467	12.20828	14.47795	18.75467
	1	14.13219	16.63540	26.57542	21.22500	14.13219	16.63540	21.22500
10	10	11.35678	14.91077	23.88632	17.24173	11.35678	14.91077	17.24173
	5	12.49234	16.09698	25.48290	19.80107	12.49234	16.09698	19.80107
	3	13.25530	14.88940	24.37941	19.16707	13.25530	14.88940	19.16707
	2	13.97497	15.76509	25.74915	20.40299	13.97497	15.76509	20.40299
	1	15.07575	22.10641	28.64565	25.00528	15.07575	22.10641	25.00528

Abbreviations: MRGLMM, model robust profile monitoring for the generalized linear mixed model; NP, nonparametric.

Another important observation is that the simulated control limits increase when the number of replications per design point is decreased for all T^2 statistics. However, this is not prominent for the control limits for the T^2 statistics based on the fitted value in inverse link scale for $k = 10$ case. Furthermore, there is an inverse relationship between simulated control limits and the number of design points. These simulated control limits for T^2 based on the fitted value in the link scale are very similar to those in Abdel-Salam et al,[45] especially when k and r are large.

The comparison of simulated control limits from each method with corresponding critical values from χ^2 distribution shows that these values are comparable for T^2 based on either the eblups or the fitted values in the link scale except for the control limits for $T^2_{MRGLMM1}$, especially when there are a moderate number of replications per design point ($k = 30$, $r = 2, 3$; $k = 20$, $r = 3, 5$, and $k = 10$, $r = 10$). But those critical values are far away for the T^2 based on the fitted values in the inverse link scale. Thus, it is clear from this study that some adjustment for the number of design points and number replications per design point is needed to obtain proper control limits from χ^2 distribution. Hence further studies are needed to determine these proper critical values. Perhaps this is a moot point, however, since the T^2 based on the fitted values in the inverse link scale is not competitive with the other methods in detecting the presence of true out-of-control profiles. Consequently, additional results for the T^2 based on the fitted values in inverse link scale is not presented.

Table 3 shows the proportion of the 1000 datasets that had a signal on the control charts for various T^2 statistics from the correctly specified model ($\gamma = 0$) without a shift in model parameters (in-control scenario). The UCL were calculated from the χ^2 distribution with corresponding df chosen as described above. Results reveal that POSs for each T^2_i based on the fitted values in inverse link scale are always one (and thus not given here) but those computed from the T^2_i based on the eblups and the fitted values in link scale are closer to α level except for the POS of $T^2_{MRGLMM1}$, especially when there is a moderate number of replications per design point. As we suggested earlier, this also reveals that adjustment is needed in order to obtain better cut off values for the T^2 statistics. Since the estimated POS based on χ^2 distribution is not always 0.05 for in-control data due to the simulated random errors, the study of control charts considered in this work is based

TABLE 3 POS for in-control scenario using the Chi-squared distribution based on df and the nominal value is 0.05

		T^2 statistic based on eblups				T^2 statistic based on fitted value in link scale		
k	r	T^2_{P0}	T^2_{NP2}	$T^2_{MRGLMM1}$	$T^2_{MRGLMM2}$	T^2_{P1}	T^2_{NP0}	$T^2_{MRGLMM0}$
30	10	0.021	0.042	0.101	0.016	0.021	0.042	0.016
	5	0.028	0.035	0.123	0.025	0.028	0.035	0.025
	3	0.033	0.083	0.088	0.042	0.033	0.083	0.042
	2	0.041	0.064	0.128	0.053	0.041	0.064	0.053
	1	0.056	0.050	0.167	0.086	0.056	0.050	0.086
20	10	0.021	0.031	0.091	0.018	0.021	0.031	0.018
	5	0.029	0.050	0.154	0.043	0.029	0.050	0.043
	3	0.035	0.045	0.106	0.048	0.035	0.045	0.048
	2	0.034	0.035	0.163	0.073	0.034	0.035	0.073
	1	0.088	0.091	0.210	0.127	0.088	0.091	0.127
10	10	0.022	0.047	0.144	0.049	0.022	0.047	0.049
	5	0.045	0.077	0.181	0.084	0.045	0.077	0.084
	3	0.064	0.048	0.171	0.080	0.064	0.048	0.080
	2	0.077	0.075	0.210	0.121	0.077	0.075	0.121
	1	0.109	0.311	0.312	0.248	0.109	0.311	0.248

Abbreviations: MRGLMM, model robust profile monitoring for the generalized linear mixed model; NP, nonparametric; POS, probability of signal.

on the simulated control limits to ensure that the POS for in-control data will be the same for all charts and equal to the nominal 0.05 level.

5.4 | Simulation results for detecting shift

A power study was conducted to evaluate the performances of each T^2 statistic on detecting an out-of-control situation by introducing a step change, drift shift and transient shift in the mean vector, β.

5.4.1 | Detecting step shift

Here, we studied the effect of the size of shift (d_0) in the mean vector, β degree of misspecification, number of design points (k) per profile and number of replications (r) per design point on the simulated POS. In step shift the first 20 profiles of 30 profiles (m) were generated from the in-control distribution using the model (35) and last 10 profiles were generated from the out-of-control model (36) in which intercepts have been shifted by d_0 amount ($d_0 = 1$, 2, 3, and 4) where $d_0 = 1$ referring to small (20%) shift and $d_0 = 4$ representing the large (80%) shift. Table 4 gives the simulated POS computed using our seven T^2 statistics at different degrees of misspecification (γ) and different amounts of shift (d_0) for $m = 30$, $k = 30$, and $r = 10$ setting. In order to obtain a proper comparison of the methods, the in-control POS (POS at $d_0 = 0$) was controlled at $\alpha = 0.05$ by using the simulated 0.95 cutoff values seen in Table 2.

Table 4 indicates that the simulated POS of all seven T^2 statistics increase as size of the shift increases for all degree of misspecification (γ), indicating that all these T^2 statistics are sensitive to step shift at this experimental settings ($m = 30$, $k = 30$, $r = 10$). The performances of each parametric T^2 statistics at $\gamma = 0$ (the properly specified model) is superior to corresponding NP and MRGLMM T^2 statistics.

TABLE 4 POS for out-of-control scenario using T^2 statistics based on eblups, fitted value in link scale and fitted value in inverse link scale with different degree of misspecification (γ) and different sizes of step shifts for $m = 30$, $k = 30$, $r = 10$, and $q = 20$

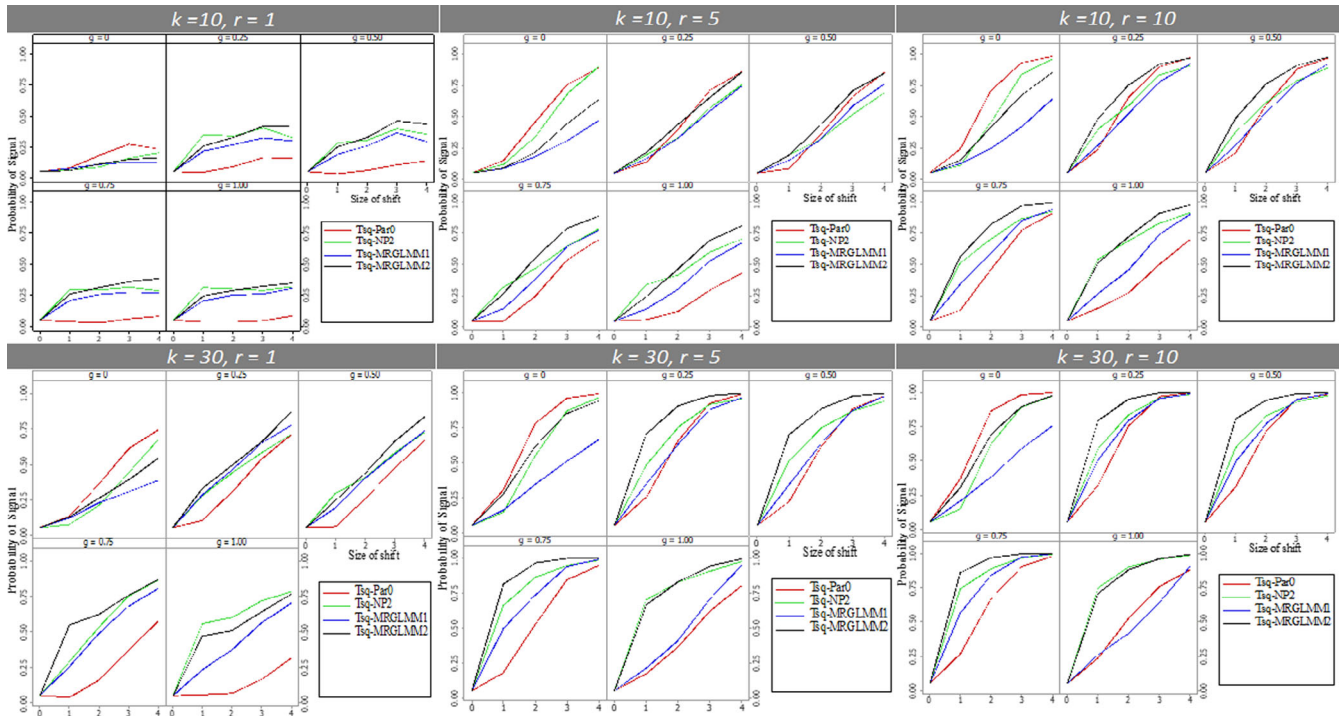|		T^2 statistic based on eblups				T^2 statistic based on fitted value in link scale		
γ	shift	T^2_{P0}	T^2_{NP2}	$T^2_{MRGLMM1}$	$T^2_{MRGLMM2}$	T^2_{P1}	T^2_{NP0}	$T^2_{MRGLMM0}$
0.00	0	0.050	0.050	0.050	0.050	0.050	0.050	0.050
	1	**0.368**	0.148	0.205	0.304	**0.368**	0.148	0.304
	2	**0.864**	0.622	0.380	0.691	**0.864**	0.622	0.691
	3	**0.983**	0.894	0.587	0.896	**0.983**	0.894	0.896
	4	**1.000**	0.974	0.753	0.975	**1.000**	0.974	0.975
0.25	1	0.321	0.588	0.505	**0.794**	0.321	0.588	**0.794**
	2	0.754	0.837	0.795	**0.952**	0.754	0.837	**0.952**
	3	0.968	0.959	0.955	**0.999**	0.968	0.959	**0.999**
	4	0.997	0.988	0.991	**0.999**	0.997	0.988	**0.999**
0.50	1	0.309	0.597	0.494	**0.809**	0.309	0.597	**0.809**
	2	0.718	0.829	0.771	**0.943**	0.718	0.829	**0.943**
	3	0.951	0.931	0.947	**0.991**	0.951	0.931	**0.991**
	4	0.988	0.976	0.990	**1.000**	0.988	0.976	**1.000**
0.75	1	0.270	0.743	0.567	**0.862**	0.270	0.743	**0.862**
	2	0.670	0.890	0.841	**0.969**	0.670	0.890	**0.969**
	3	0.903	0.969	0.974	**1.000**	0.903	0.969	**1.000**
	4	0.982	0.995	0.997	**1.000**	0.982	0.995	**1.000**
1.00	1	0.238	**0.749**	0.261	0.709	0.238	**0.749**	0.709
	2	0.527	**0.902**	0.418	0.882	0.527	**0.902**	0.882
	3	0.758	0.959	0.643	**0.964**	0.758	0.959	**0.964**
	4	0.882	0.988	0.907	**0.994**	0.882	0.988	**0.994**

Note: Best value in bold.

Abbreviations: MRGLMM, model robust profile monitoring for the generalized linear mixed model; NP, nonparametric; POS, probability of signal.

T^2 statistics based on eblups shows that $T^2_{MRGLMM2}$ often gives a higher POS than $T^2_{MRGLMM1}$ and is better for detecting step shift. The performances of MRGLMM methods ($T^2_{MRGLMM2}$, $T^2_{MRGLMM0}$, and $T^2_{MRGLMM3}$) are superior to corresponding parametric and NP methods for partially misspecified models. The performances of NP T^2 statistics based on eblups and fitted value in the link scale are better than corresponding parametric and MRGLMM methods for small shifts in the totally misspecified model ($\gamma = 1$) but for large shifts those MRGLMM methods are better. Furthermore, we observe that the performance of those T^2 statistics based on the eblups and the fitted values in link scale are equal.

Figure 3 shows the simulated POS of T^2 statistics based on eblups, with different sizes of shifts, different degrees of misspecification and different numbers of binary observations (r) for $k = 10$ and 30 scenarios. Results indicate that the simulated POSs of all T^2 statistics based on eblups increase with an increasing number of replications (r) per design point for all values of design points used (k) in this study. These plots show a similar pattern to those plots given in Abdel-Salam et al[45] for normal data cases, especially when there is an adequate number of observations per profile ($k = 30$ and $r = 10$).

The T^2 statistics based on the fitted value in probability scale also indicates that the POS of all T^2 statistics increase with an increasing number of replications (r) per design point for $k = 20$ and 30, but for $k = 10$, there

FIGURE 3 Colored: probability of signal of T^2 statistics based on eblups, with shifts, degree of misspecification ($\gamma = g$) and number of binary replications (r) for $m = 30$, $k = 10$ and 30 [Color figure can be viewed at wileyonlinelibrary.com]

is a small increase in the simulated POSs of parametric method and no significant increase in NP and MRGLMM methods.

Furthermore, it was noticed that the T^2 statistics based on the fitted value in probability scale are highly sensitive to the number of design points (k) than those T^2 statistics based on ebulps and perform well with a large number of design points.

5.4.2 | Detecting drift shift

To study the performances of T^2 statistics under drift shift in the mean vector, β, was shifted introducing a constantly increasing rate from the second profile up to the last profile by multiplying the shift ($d_0 = 1, 2, 3$, and 4) in the out of control model (36) by $\left(\frac{q-1}{m-1}\right)$ amount. Table 5 shows a variation of the simulated POS computed using different T^2 statistics at different degrees of misspecification (γ) and different amounts of shift (d_0) for $m = 30$, $k = 30$, and $r = 10$ setting. In order to obtain a proper comparison of the methods, the in-control POS (POS at $d_0 = 0$) was controlled at $\alpha = 0.05$ by using the simulated 0.95 cutoff values seen in Table 2. The results reveal that the performances of each T^2 statistics in detecting drift shift are similar to step shift. The performances of each parametric T^2 statistics at $\gamma = 0$ is superior to corresponding NP and MRGLMM T^2 statistics; and the performances of MRGLMM methods (T^2_{MRGLMM2} and T^2_{MRGLMM0}) are superior to corresponding parametric and NP methods for partially misspecified models, whereas NP T^2 are better for a totally misspecified model.

5.4.3 | Detecting transient shift

To evaluate the performances of different T^2 statistics under the transient shift, the first and last 10 profiles of 30 profiles (m) were generated from the in-control distribution using the model (35) and the middle 10 profiles were generated from the out-of-control model (36) in which intercepts have been shifted by d_0 amount ($d_0 = 1, 2, 3$, and 4) where $d_0 = 1$ referring to small (20%) shift and $d_0 = 4$ representing the large (80%) shift. Table 6 shows the variation of the

TABLE 5 POS for out-of-control scenario using T^2 statistics based on eblups, fitted value in link scale, and fitted value in inverse link scale with different degree of misspecification (γ) and different sizes of drift shifts for $m = 30$, $k = 30$, and $r = 10$

γ	shift	T^2_{P0}	T^2_{NP2}	$T^2_{MRGLMM1}$	$T^2_{MRGLMM2}$	T^2_{P1}	T^2_{NP0}	$T^2_{MRGLMM0}$
				T^2 statistic based on eblups		T^2 statistic based on fitted value in link scale		
0.00	0	0.050	0.050	0.050	0.050	0.050	0.050	0.050
	1	**0.405**	0.181	0.189	0.312	**0.405**	0.181	0.312
	2	**0.781**	0.489	0.287	0.583	**0.781**	0.489	0.583
	3	**0.929**	0.713	0.362	0.707	**0.929**	0.713	0.707
	4	**0.972**	0.788	0.421	0.750	**0.972**	0.788	0.750
0.25	1	0.250	0.170	0.187	**0.394**	0.250	0.170	**0.394**
	2	0.599	0.417	0.441	**0.715**	0.599	0.417	**0.715**
	3	0.841	0.624	0.568	**0.858**	0.841	0.624	**0.858**
	4	**0.905**	0.668	0.639	0.896	**0.905**	0.668	0.896
0.50	1	0.145	0.155	0.112	**0.265**	0.145	0.155	**0.265**
	2	0.508	0.505	0.346	**0.626**	0.508	0.505	**0.626**
	3	0.888	0.873	0.787	**0.947**	0.888	0.873	**0.947**
	4	0.995	0.995	0.981	**0.998**	0.995	0.995	**0.998**
0.75	1	0.222	0.205	0.048	**0.242**	0.222	0.205	**0.242**
	2	0.477	0.478	0.203	**0.528**	0.477	0.478	**0.528**
	3	0.831	0.821	0.543	**0.853**	0.831	0.821	**0.853**
	4	0.981	0.980	0.889	**0.989**	0.981	0.980	**0.989**
1.00	1	0.263	**0.266**	0.048	0.254	0.263	**0.266**	0.254
	2	0.402	**0.422**	0.103	0.366	0.402	**0.422**	0.366
	3	0.613	**0.675**	0.267	0.652	0.613	**0.675**	0.652
	4	0.846	**0.884**	0.578	0.876	0.846	**0.884**	0.876

Note: Best value in bold.

Abbreviations: MRGLMM, model robust profile monitoring for the generalized linear mixed model; NP, nonparametric; POS, probability of signal.

simulated POS computed using different T^2 statistics at different degrees of misspecification (γ) and different amounts of shift (d_0) for $m = 30$, $k = 30$, and $r = 10$ case. The performances parametric T^2 statistics, MRGLMM T^2 and NP T^2 statistics are similar to step and drift shifts for correctly specified ($\gamma = 0$) model to moderately misspecified models ($\gamma = 0.25, 0.5$). However, with highly misspecified models ($\gamma = 0.75, 1$), the performances of parametric T^2 statistics is better.

6 | A REAL CASE

In this section, we compare the performance of the proposed method using the warranty claim dataset given in Reference 43. Figure 4 illustrates changes in the probability of warranty claim with the month in service for 10 lots (profiles) of automotives manufactured during 10 months of production (MOP).

We first removed the data of 10th MOP since it looks unusual and deviate from other nine profiles. Then GLMM was fitted to data of remaining nine profile using model in (1) with binomial link function. The model estimated mean vector, $\boldsymbol{\beta}$ (fixed effect parameters) as $(-5.331, 0.1603)^{\cdot}$ and random effects vector, $\boldsymbol{b}$ as $(0.008211, 0.000012)^{\cdot}$. Then 30 profiles were generated using these estimate to represent in control profiles and control

TABLE 6 POS for out-of-control scenario using T^2 statistics based on eblups, fitted value in link scale and fitted value in inverse link scale with different degree of misspecification (γ) and different sizes of transient shifts for $m = 30$, $k = 30$, and $r = 10$

		T^2 statistic based on eblups				T^2 statistic based on fitted value in link scale		
γ	shift	T^2_{P0}	T^2_{NP2}	$T^2_{MRGLMM1}$	$T^2_{MRGLMM2}$	T^2_{P1}	T^2_{NP0}	$T^2_{MRGLMM0}$
0.00	0	0.050	0.050	0.050	0.050	0.050	0.050	0.050
	1	**0.405**	0.181	0.189	0.312	**0.405**	0.181	0.312
	2	**0.781**	0.489	0.287	0.583	**0.781**	0.489	0.583
	3	**0.929**	0.713	0.362	0.707	**0.929**	0.713	0.707
	4	**0.972**	0.788	0.421	0.750	**0.972**	0.788	0.750
0.25	1	0.250	0.170	0.187	**0.394**	0.250	0.170	**0.394**
	2	0.599	0.417	0.441	**0.715**	0.599	0.417	**0.715**
	3	0.821	0.589	0.547	**0.839**	0.821	0.589	**0.839**
	4	0.905	0.668	0.639	**0.896**	0.905	0.668	**0.896**
0.50	1	0.202	0.173	0.118	**0.297**	0.202	0.173	**0.297**
	2	0.569	0.435	0.332	**0.669**	0.569	0.435	**0.669**
	3	0.789	0.565	0.527	**0.849**	0.789	0.565	**0.849**
	4	0.905	0.686	0.630	**0.942**	0.905	0.686	**0.942**
0.75	1	0.253	0.217	0.072	**0.281**	0.253	0.217	**0.281**
	2	**0.550**	0.449	0.193	0.527	**0.550**	0.449	0.527
	3	**0.792**	0.606	0.417	0.769	**0.792**	0.606	0.769
	4	0.862	0.686	0.554	**0.875**	0.862	0.686	**0.875**
1.00	1	**0.305**	0.286	0.043	0.266	**0.305**	0.286	0.266
	2	**0.451**	0.409	0.084	0.393	**0.451**	0.409	0.393
	3	**0.655**	0.534	0.172	0.561	**0.655**	0.534	0.561
	4	**0.759**	0.590	0.264	0.650	**0.759**	0.590	0.650

Note: Best value in bold.
Abbreviations: MRGLMM, model robust profile monitoring for the generalized linear mixed model; NP, nonparametric; POS, probability of signal.

limits for each of T^2 statistics described above were calculated from 1000 replications of these 30 in control profiles.

To study the practical performances of each T^2 statistics under the step shift the first 20 profiles of 30 profiles were generated as above and treated as in-control profiles and last 10 profiles were generated by shifting the intercept of the estimated mean vector, β by d_0 amount ($d_0 = 1, 2, 3,$ and 4) and treated as out-of-control profiles. The POS of each method was computed using 1000 replication and results for $m = 30$, $k = 12$, and $r = 100$ is presented in Figure 5.

The results indicate that parametric T^2 statistics is superior in detecting the step shift in this practical scenario and MRGLMM2 T^2 is comparatively better than MRGLMM1 T^2 and NP T^2.

7 | CONCLUSION

The study showed that, the SIMSE of the SP (MRGLMM) method is less than or equal to that of the parametric method for correctly specified model and it is always less than that of the parametric and NP method for moderately misspecified models. For severe misspecification's levels, the SIMSE of the NP method and MRGLMM are identical. Thus, we can

FIGURE 4 The profile samples for automotive warranty claims data [Color figure can be viewed at wileyonlinelibrary.com]
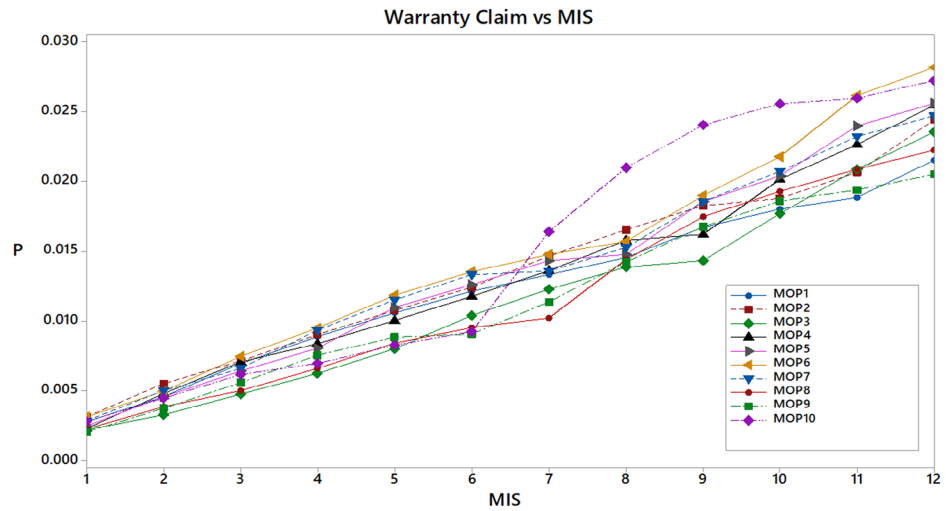
FIGURE 5 Probability of signal of T^2 statistics with step shifts generated using profile samples for automotive warranty claims data
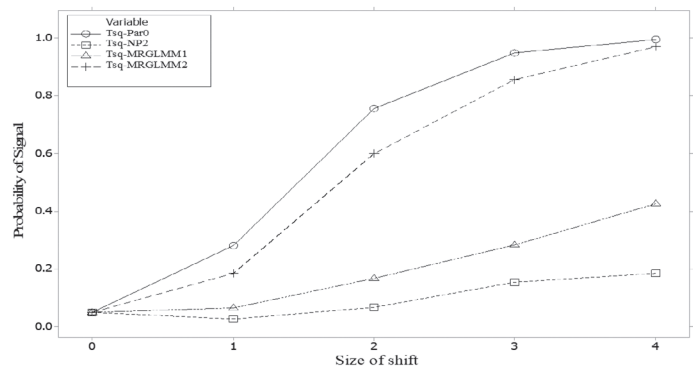
conclude that MRGLMM fit is robust to the degree of model misspecification and it combines the advantage of both parametric and NP methods. Furthermore, we observed that SIMSE of all methods decreases as the number of observations per profile increases (r or k). Thus, the model fitting ability by the parametric and NP methods can be improved by using more observations per profile (k or/and r). Thus, we advise to increase the k and/or r whenever possible.

The power study's results show that, the T^2 statistics based on eblups (fitted value in link scale) are superior to the corresponding T^2 statistic based on the fitted value in the probability scale in detecting out-of-control profiles. In addition, a comparison of two SP T^2 statistics based on eblups shows that $T^2_{\text{MRGLMM2},i}$ often gives higher POS than $T^2_{\text{MRGLMM1},i}$ and is thus better for detecting out-of-control situation. Thus, we conclude that, the user, when dealing with a 0/1 response variable, needs to transform the observed probabilities into the link scale and use the T^2 control chart based on the fitted value in the link scale or that based on eblups rather than using a control chart based on observed probabilities. This recommendation may or may not be true for other distributions within the exponential family of distributions. It is our observation that the poor performance of the T^2 method based on the probability scale is due to numerical inaccuracies involved in the computation of T^2. We speculate that such inaccuracies may not be an issue when dealing with responses from other distributions in the exponential family of distribution such as the Poisson or the Gamma distributions. Further work is needed to address this concern.

Furthermore, the probability of out-of-control signal of MRGLMM for the correctly specified model case is greater than that of NP method and closer to POS of parametric method. For totally misspecified model, the POS of MRGLMM is greater than that of parametric method and closer to POS of NP method. For moderate misspecification levels, the MRGLMM method results in higher POS values than either the parametric and NP methods. These results suggest that the MRGLMM method is robust to the level of model misspecification in detecting out-of-control conditions. Thus, we recommended using the MRGLMM method for Phase I profile monitoring when the response variable is a 0/1 variable and no parametric model can be found that provides an adequate fit to the data over the entire range of the explanatory variables.

ORCID

Abdel-Salam G. Abdel-Salam https://orcid.org/0000-0003-4905-6489

REFERENCES

1. Kang L, Albin SL. On-line monitoring when the process yields a linear profile. *J Qual Technol*. 2000;32:418-426.
2. Woodall WH, Spitzner DJ, Montgomery DC, Gupta S. Using control charts to monitor process and product quality profiles. *J Qual Technol*. 2004;36:309-320.
3. Woodall WH. Current research in profile monitoring. *Production*. 2007;17:420-425.
4. Kim K, Mahmoud MA, Woodall WH. On the monitoring of linear profiles. *J Qual Technol*. 2003;35:317-328.
5. Mahmoud MA, Woodall WH. Phase I analysis of linear profiles with calibration applications. *Technometrics*. 2004;46:380-391.
6. Mahmoud MA, Parker PA, Woodall WH, Hawkins DM. A change point method for linear profile data. *Qual Reliab Eng Int*. 2007;23:247-268.
7. Zou C, Wang Z, Tsung F. A self-starting control chart for linear profiles. *J Qual Technol*. 2007;39:364-375.
8. Mahmoud MA. Phase I analysis of multiple linear regression. *Commun Stat Simul Comput*. 2008;37:2106-2130.
9. Khedmati M, Niaki S. Phase-I monitoring of general linear profiles in multistage processes. *Commun Stat Simul Comput*. 2015;46(6):4465-4489. https://doi.org/10.1080/03610918.2015.1118507.
10. Khedmati M, Niaki S. Phase-I robust parameter estimation of simple linear profiles in multistage processes. *Commun Stat Simul Comput*. 2019;48:1-26. https://doi.org/10.1080/03610918.2019.1653916.
11. Noorossana R, Aminmadani M, Saghaei A. Effect of Phase I estimation error on the monitoring of simple linear profiles in Phase II. *Int J Adv Manuf Technol*. 2016;84:873-884. https://doi.org/10.1007/s00170-015-7078-2.
12. Chiang JY, Lio YL, Tsai TR. MEWMA control chart and process capability indices for simple linear profiles with within profile autocorrelation. *Qual Reliab Eng Int*. 2016;33:1083-1094.
13. Kalaei M, Atashgar K, Niaki S. Phase I monitoring of simple linear profiles in multistage processes with cascade property. *Int J Adv Manuf Technol*. 2018;94:1745-1757. https://doi.org/10.1007/s00170-016-9691-0.
14. Mahmood T, Riaz M, Hafidz Omar M, Xie M. Alternative methods for the simultaneous monitoring of simple linear profile parameters. *Int J Adv Manuf Technol*. 2018;97:2851-2871. https://doi.org/10.1007/s00170-018-2149-9.
15. Moheghi HR, Noorossana R, Ahmadi O. Phase I and Phase II analysis of linear profile monitoring using robust estimators. *Commun Stat Theory Methods*. 2020;49:1-18. https://doi.org/10.1080/03610926.2020.1758724.
16. Touqeer F, Mahmood T, Riaz M, Abbas N. On developing linear profile methodologies: a ranked set approach with engineering application. *J Eng Res*. 2020;8:203-225.
17. 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;26:291-303.
18. Eyvazian M, Noorossana R, Saghaei A, Amiri A. Phase II monitoring of multivariate multiple linear regression profiles. *Qual Reliab Eng Int*. 2011;27:281-296.
19. Soleimani P, Noorossana R, Niaki S. Monitoring autocorrelated multivariate simple linear profiles. *Int J Adv Manuf Technol*. 2013;67:1857-1865. https://doi.org/10.1007/s00170-012-4614-1.
20. Amiri A, Saghaei A, Mohseni M, Zerehsaz Y. Diagnosis aids in multivariate multiple linear regression profiles monitoring. *Commun Stat Theory Methods*. 2014;43(14):3057-3079. https://doi.org/10.1080/03610926.2012.694543.
21. Ayoubi M, Kazemzadeh R, 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. https://doi.org/10.1002/qre.1760.
22. Paynabar K, Zou C, Qiu P. A change-point approach for Phase-I analysis in multivariate profile monitoring and diagnosis. *Technometrics*. 2016;58(2):191-204. https://doi.org/10.1080/00401706.2015.1042168.
23. Bahrami H, Niaki S, Khedmati M. Monitoring multivariate profiles in multistage processes. *Commun Stat Simul Comput*. 2019;48:1-29. https://doi.org/10.1080/03610918.2019.1626882.
24. Vaghefi A, Tajbakhsh SD, Noorossana R. Phase II monitoring of nonlinear profiles. *Commun Stat Theory Methods*. 2009;38:1834-1851.
25. Williams JD, Woodall WH, Birch JB. Statistical monitoring of nonlinear product and process quality profiles. *Qual Reliab Eng Int*. 2007;23:925-941.
26. Steiner S, Jensen WA, Grimshaw SD, Espen B. Nonlinear profile monitoring for oven-temperature data. *J Qual Technol*. 2016;48(1):84-97. https://doi.org/10.1080/00224065.2016.11918153.
27. Guevara RD, Vargas JA. Evaluation of process capability in multivariate nonlinear profiles. *J Stat Comput Simul*. 2015;86(12):2411-2428. https://doi.org/10.1080/00949655.2015.1112391.
28. Chen PW, Peng WS, Cheng CS, Wang YT. Monitoring of nonlinear profile variations using support vector machines. *J Qual*. 2017;24(5):313-323.
29. Awad MI, AlHamaydeh M, Faris A. Fault detection via nonlinear profile monitoring using artificial neural networks. *Qual Reliab Eng Int*. 2018;34:1195-1210.
30. Pan JN, Li CI, Lu MZ. Detecting the process changes for multivariate nonlinear profile data. *Qual Reliab Eng Int*. 2019;35:1890-1910. https://doi.org/10.1002/qre.2482.
31. Li CI, Pan JN, Liao CH. Monitoring nonlinear profile data using support vector regression method. *Qual Reliab Eng Int*. 2018;35(1):127-135.
32. Jensen WA, Birch JB, Woodall WH. Monitoring correlation within linear profiles using mixed models. *J Qual Technol*. 2008;40:167-183.
33. Jensen WA, Birch JB. Profile monitoring via nonlinear mixed models. *J Qual Technol*. 2009;41:18-34.

34. Yeh AB, Huwang L, Li YM. Profile monitoring for a binary response. *IIE Trans*. 2009;41:931-941.

35. Amiri A, Koosha M, Azhdari A. Profile monitoring for Poisson response. Paper presented at: Proceedings of IEEM Conference; 2011; IEEE, Singapore.

36. Koosha, M. and Amiri, A. (2011). The effect of link function on the monitoring of logistic regression profiles. In Proceedings of the World Congress on Engineering; London, U.K. (Vol. 1, pp. 6–8).

37. Koosha M, Amiri A. Generalized linear mixed model for monitoring autocorrelated logistic regression profiles. *Int J Adv Manuf Technol*. 2012;64:487-495.

38. Izadbakhsh H, Noorossana R, Niaki S. Monitoring multinomial logistic profiles in Phase I using log-linear models. *Int J Qual Reliab Manag*. 2018;35(3):678-689. https://doi.org/10.1108/IJQRM-04-2017-0068.

39. Bayer FM, Tondolo CM, Müller FM. Beta regression control chart for monitoring fractions and proportions. *Comput Ind Eng*. 2018;119:416-426. https://doi.org/10.1016/j.cie.2018.04.006.

40. Canterle DR, Bayer FM. Variable dispersion beta regressions with parametric link functions. *Stat Pap*. 2019;60:1541-1567. https://doi.org/10.1007/s00362-017-0885-9.

41. Lima-Filho LMA, Pereira TL, Souza TC, Bayer FM. Process monitoring using inflated beta regression control chart. *PLoS One*. 2020;15(7):e0236756. https://doi.org/10.1371/journal.pone.0236756.

42. Nottingham QJ, Birch JB. A semiparametric approach to analyzing dose-response data. *Stat Med*. 2000;19:389-404.

43. Shang Y, Wand Z, He Z, He S. Nonparametric change-point detection for profiles with binary data. *J Qual Technol*. 2017;49(2):123-135. https://doi.org/10.1080/00224065.2017.11917984.

44. Piri S, Abdel-Salam A-SG, Boone EL. A wavelet approach for profile monitoring of Poisson distribution with application. *Commun Stat Simul Comput*. 2019;48:1-12.

45. Abdel-Salam GA, Birch JB, Jensen WA. Nonparametric and semiparametric mixed model methods for Phase I profile monitoring. *Qual Reliab Eng Int*. 2013;29:555-569. https://doi.org/10.1002/qre.1405.

46. Wolfinger R, O'Connell M. Generalized linear models: a pseudo likelihood approach. *J Stat Comput Simul*. 1993;48:233-243.

47. Ruppert D, Wand MP, Carroll RJ. *Semiparametric Regression*. Cambridge, NY: Cambridge University Press; 2003;12.

48. Mays JE, Birch JB, Einsporn RL. An overview of model-robust regression. *J Stat Comput Simul*. 2000;66:79-100.

49. Mays J, Birch JB, Starnes B. Model robust regression: combining parametric, nonparametric and semiparametric methods. *J Nonparametric Stat*. 2001;13:245-277.

50. Robinson TJ, Birch JB, Starnes A. A semiparametric approach to dual modeling when no replication exists. *J Stat Plan Inference*. 2010;140:2860-2869.

51. Pickle SM, Robinson TJ, Birch JB, Anderson-Cook C. A semiparametirc approach to robust parameter design. *J Stat Plan Inference*. 2008;138:114-131.

52. Wan W, Birch JB. A semiparametric technique for the multi-response optimization problem. *Qual Reliab Eng Int*. 2011;27:47-59.

53. Bandara AMKR. Nonparametric and Semiparametric Profile Monitoring Methods for Non-Normal Data [unpublished doctoral thesis]. Peradeniya, Sri Lanka: Postgraduate Institute of Agriculture, University of Peradeniya; 2016.

APPENDIX: SAS CODE A

```
*Following is the main macro, here r is the number of binary observations at particular point
m is the number of profiles, k is the number of design points for x variable,
s1, s2, s3 are variance used to generate random effects b1, b2 and b3 respectively
q is the last in-control profile
nr is the number of simulation per sample*/
%macro mse (r=10, m=30, k=30, s1=5, s2=20,  s3=20, nr=1000, q=20, g=0, nk=3);
%do kt=1 %to 1 %by 1;
%do dd0=%SysEvalF(0) %to %SysEvalF(40) %by %SysEvalF(10);
%let d0= %SysEvalF( &dd0 / 10 );
%let outpa=1;
%let tot=1;
%let notest=0;
%let notcon_p=0;
```

```
%let notcon_np=0;
%do %until (&outpa > &nr);
proc iml;
dm 'clear out';dm 'clear log';
d1=0;
xk = j(&k,1,1)||(t(1:&k)*(.0333333));
run;
xr = xk @ j(&r,1,1);
x = repeat (xr[,2],&m,1);
run;
y1= j(&m*&k*&r,1,0);* print y1;
PA= j(&m,&k,0);*print PA;
CS= j(&m,&k,0);*print CS;
b11 = root(&s1) * normal (j(&m,1,&tot));
run;
b1 = t(b11)@(j(&k,1,1));*contribution of random effect b1 to each obs;
run;
b22 = root(&s2) * normal (j(&m,1,&tot+10##6));
run;
b2 = t(b22)@(j(&k,1,1));
run;
b33 = root(&s3) * normal (j(&m,1,&tot+10##7));
run;
b3 = t(b33)@(j(&k,1,1));
run;
%let tot=%eval(&tot+1);
do repm = 1 to &q by 1;
    do repn = 1 to &k by 1; *———-5 in following equation is beta0
        PA[repm,repn] =(1-&g)*(1/(1+exp(-((-5)+(10)*xk[repn,2]))))+

                &g*(.5*(1/(1+exp(-((-5)+(20)*xk[repn,2]))))
                +.5*(1/(1+exp(-((-15)+(20)*xk[repn,2]))))));
        CS[repm,repn] =(1-&g)*(1/(1+exp(-((-5)+(10+b1[repn,repm])*xk[repn,2]))))+
                &g*(.5*(1/(1+exp(-((-5)+(20+b2[repn,repm])*xk[repn,2]))))
                +.5*(1/(1+exp(-((-15)+(20+b3[repn,repm])*xk[repn,2])))))); end; end; run; prpa= colvec(pa); Genpa = prpa
@ j(&r,1,1); pcs= rowvec(cs);L = j(&r,&k*&m,0);
do i = 1 to &k*&m by 1;
do j = 1 to &r;
L[j,i]=rand('BERN',pcs[1,i]);
end; end;
data cs; set cs; rename col1=prof; rename col2=x; rename col3=cs;rename col4=PA;rename col5=y;rename col6=obsp;
run;
ods listing;
proc glimmix data = cs outdesign = z;
ods output SolutionR=parcs ParameterEstimates = parp;
class prof;
model y(event='1')=x / solution link=logit dist=binary s;
random int/subject=prof s;
random x/subject=prof s;
output out = gmxpara pred(blup ilink)=predpcs pred(blup)=predpcsL pred(noblup ilink)=predparapa;
nloptions tech = newrap; run; quit;
data _null_;   slept= sleep(6,.1); run;
%let exist_p=%sysfunc(exist(work.parp)); %if &exist_p=0 %then %do;
```

```
 %put dataset P does not exist;   %let notcon_p=%eval(&notcon_p+1); proc datasets; delete cs;run; %end;
%else %if &exist_p=1 %then %do;
%put dataset p exist;
ods listing;
proc glimmix data = cs outdesign = z;
ods output SolutionR =concs ParameterEstimates = parnp knotInfo = knotcs;
title 'Estimated CS profiles using Rsmoother';
class prof;
model y(event='1')=x / solution link=logit dist=binary s;
random x/ type = rsmooth subject = prof knotmethod=data(knotdata&kt) knotinfo g s;
random x/ subject = prof s; random _RESIDUAL_;
output out = gmxnpara pred(blup ilink) = prednpcs pred(blup) = prednpcsL pred(noblup ilink) = prednppa;
nloptions tech = newrap; run; quit;
%let exist_np=%sysfunc(exist(work.parnp));
%if &exist_np=0 %then %do; %put dataset NP does not exist;
%let notcon_np=%eval(&notcon_np+1);  proc datasets;                    delete cs parp;  run;    %end;
%else %if &exist_Np=1 %then %do; %put dataset NP exist;
tsqP_fit = j(&m,1,0); tsqNP_fit = j(&m,1,0);tsqsP_fit = j(&m,1,0);
tsqP_fitL = j(&m,1,0); tsqNP_fitL = j(&m,1,0);tsqsP_fitL = j(&m,1,0);
 do w1 = 1 to &m by 1;
tsqP_fit[w1,1] = (mtypar[w1,]-meanmtypar)*sefitp*t(mtypar[w1,]-meanmtypar);
tsqNP_fit[w1,1] = (mtynpar[w1,]-meanmtynpar)*sefitnp*t(mtynpar[w1,]-meanmtynpar); tsqSP_fit[w1,1] = (mtyspar
[w1,]-meanmtyspar)*sefitsp*t(mtyspar[w1,]-meanmtyspar);
tsqP_fitL[w1,1] = (mtyparL[w1,]-meanmtyparL)*sefitpL*t(mtyparL[w1,]-meanmtyparL);
tsqNP_fitL[w1,1] = (mtynparL[w1,]-meanmtynparL)*sefitnpL*t(mtynparL[w1,]-meanmtynparL);
tsqSP_fitL[w1,1] = (mtysparL[w1,]-meanmtysparL)*sefitspL*t(mtysparL[w1,]-meanmtysparL); end;
Htsq_fit = max(tsqP_fitL)||max(tsqNP_fitL)||max(tsqSP_fitL)||max(tsqP_fit)||max(tsqNP_fit)||max(tsqSP_fit);
tsquarecsebNP = j(&m,1,0);
 do w = 1 to &m by 1;
     tsquarecsebNP[w,1] = (csmteb[w,]-meancseb)*seb2*t(csmteb[w,]-meancseb); end;
HtscsebPNP = max(tsqcsebp[,1])||max(tsquarecsebNP);
do l = 1 to &m-1 by 1;
     vdebsp[l,]=(csebsp[l+1,]-csebsp[l,]);
         vdebsp_new[l,]=(csebsp_new[l+1,]-csebsp_new[l,]);
end;
 do w = 1 to &m by 1;
     tsqSP[w,1] = (csebsp[w,]-meancsebsp)*sebsp2*t(csebsp[w,]-meancsebsp);
         tsqSP_new[w,1] = (csebsp_new[w,]-meancsebsp_new)*sebsp2_new*t(csebsp_new[w,]-meancsebsp_new); end;
%mend; %mse
```